Glycosaminoglycan Degradation - Structural and Kinetic Studies of Heparinase II and Chondroitinase ABC

David Shaya, M.Sc.

Doctor of Philosophy

Department of Biochemistry

McGill University
Montreal, Quebec, Canada
August 2008

A thesis submitted to McGill University
in partial fulfillment of the requirements of the degree of
Doctor of Philosophy

ABSTRACT

Glycosaminoglycans (GAGs) are linear, heterogeneous, negatively charged polysaccharides, common within the extracellular matrix (ECM) and at the cell-surfaces of all metazoan cells. Apart from their structural importance to the integrity of the ECM, GAGs are also fundamental modulators of various biological processes at the level of cells (i.e. adhesion, signaling and proliferation), tissue (i.e. inflammation, wound repair, tissue morphogenesis and organogenesis) and organism (i.e. cancer and developmental processes). The biological activities of GAGs are intricately linked to their turnover.

Niche-adapted microorganisms express GAG-degrading enzymes, allowing them to process GAGs for nutritional purposes, both for themselves and for their mammalian hosts. In particular, bacterial GAG lyases cleave the glycosidic bond next to the uronic acid present in GAGs, through a β-elimination mechanism that yields disaccharide products with a saturated 4-deoxy-α-L-*threo*-hex-4-enopyranosyluronic acid (ΔUA). Among the lyasses, heparinase II (HepII) and chondroitinase ABC (ChonABC) possess the striking ability to degrade GAGs regardless of their uronic acid epimer. HepII degrades both heparin and heparan sulfate, whereas ChonABC degrades both chondroitin sulfate and dermatan sulfate. This research is aimed at providing insight into the catalytic mechanism and the substrate specificity of these two enzymes at the molecular level. The applied methodologies used were: (1) X-ray crystallography, aimed at obtaining the crystal structures of *P. heparinus* HepII and *B. thetaiotaomicron* ChonABCII as well as enzyme-substrate/product complexes, (2) site-directed mutagenesis of key residues identified in the structures, and (3) the enzymologic characterization of these mutants, to provide insight into the mechanistic roles of the postulated active site residues necessary for degradation of GAGs by HepII and ChonABCII.

These studies resulted in the first structure of a heparin/HS-degrading enzyme, namely *P. heparinus* heparinase II, as well as the structure of *B.*

thetaiotaomicron chondroitin lyase ABCII. They demonstrate that both enzymes utilize a single active site for cleavage next to the uronic acid, with different subset of residues assuming catalytic roles for L-iduronic acid *vs.* D-glucoronic acid. Based on the accumulated experimental data, an understanding of the observed substrate specificities of these lyases and their modes of action is proposed. Finally, the *B. thetaiotaomicron* chondroitin lyase was biochemically characterized in order to lay the groundwork for its clinical applications.

ABRÉGÉ

Les glycoaminoglycans (GAGs) sont des polysaccharides, hétérogènes, linéaires et négativement chargés, communs de la matrice extracellulaire (ECM) et de la surface cellulaire chez tous les métazoaires. En plus de leurs importances structurales dans l'intégrité de la ECM, les GAGs sont aussi des modulateurs fondamentaux de divers processus biologiques au niveau cellulaire (dans l'adhésion, la signalisation et la prolifération), au niveau tissulaire (dans l'inflammation, la cicatrisation, la morphogénie des tissus et l'organogenèse) ou dans les organismes (dans le cancer et dans les processus du développement). Les activités biologiques de GAGs sont donc fortement liées à leurs « turnover » au cours du métabolisme.

Les microorganismes adaptés aux niches expriment des enzymes dégradant les GAGs afin de les utiliser dans leurs nutritions et dans celles de leurs hôtes. Plus particulièrement les GAGs lyases bactériennes dégradent, par un mécanisme de β-élimination, les liaisons glycosidiques contiguës à l'acide uronique présent dans les GAGs, créant ainsi comme produits des disaccharides saturés (ΔUA). Parmi ces lyases, l'heparinase II (HepII) et la chondroitinase ABC (ChondABC) possèdent toutes deux l'étonnante habileté de dégrader les GAGs quel que soit l'épimérisation de l'acide uronique (HepII dégrade l'héparine et l'heparan sulfate (HS), alors que ChonABC dégrade la choindroitine sulfate et le dermatan sulfate). Le but de ce travail fut de comprendre le mécanisme catalytique et la spécificité des substrats des ces enzymes à un niveau moléculaire. Pour ce faire, les méthodes utilisées ont été (1) la cristallographie et la diffraction des rayons X qui ont permis l'obtention des structures cristallographiques des enzymes *P. heparinus* HepII et *B. thetaiotaomicron* ChonABCII en complexe avec substrats et produits, (2) la mutagenèse dirigée des résidus identifiés dans les structures et (3) la caractérisation enzymatique des ces mutants donnant un aperçu des rôles mécaniques des supposés résidus catalytiques pendant la dégradation des GAGs par HepII et BactnABCII.

Cette étude présente donc la première structure d'un enzyme dégradant l'héparine et HS, nommé *P. heparinus* heparinase II, ainsi que la structure de *B. thetaiotaomicron* chondroitine lyase ABCII. Il a été démontré que ces deux enzymes utilisent un seul site actif pour le clivage de chaque épimère de l'acide uronique, mais mettent en jeu divers sous-ensembles de résidus impliqués dans des rôles catalytiques différents selon épimérisation. De plus cette recherche apporte une meilleure compréhension de la spécificité, pour les substrats observés de ces lyases, et de leurs modes d'action. Finalement, *B. thetaiotaomicron* chondroitine lyase a été biochimiquement caractérisée posant ainsi les bases en vue d'une utilisation clinique.

ACKNOWLEDGMENTS

I would first like to thank my supervisor **Dr. Miroslaw Cygler** for providing me the opportunity to work in his group. I am very grateful for the guidance, intellectual freedom and encouragement he provided throughout my research. I would especially like to thank **Dr. Allan matte** and **Dr. Joseph Schrag** for their assistance, and interest in my projects. I have greatly benefited from our many long discussions. I thank **Mm. Yunge Li** for all her patience and assistance during my first year as a graduate student. Many thanks are extended to all members of the lab, both past and present, **Dr. Ante Tocilj**, **Dr. Marie-Line Garron, Dr. Tonje M. Bjerkan, Dr. Stephane Raymond, Dr. Ming-Ni Hung, Dr. Rong Shi, Dr. Rangarajan Erumbi, Dr. Michael Sacher, Dr. Vladimir Lunin, Mm. Eunice Ajamian, Mm. Christine Munger, Mm. Ariane Proteau, Mr. Guy Nadeau,** and **Mr. John Wagner** for creating an exciting and positive environment to work in. I would also like to thank my collaborators: at the Rensselaer Polytechnic Institute, NY, **Prof. Robert J. Linhardt**, at the National Product Research Institute, Seul, **Prof. Yeong Shik Kim** and **Dr. Bam-Soo Hahn** and at Momenta Pharmaceuticals, Boston, **Dr. James R. Myette**, who supplied me with material for my research and worked side by side with me overcoming together the research hurdles I have encountered. Finally, I give my heartfelt thanks to my parents and to my sweet wife, **Zeng Shang Ping**, for their continuing support and patience.

CONTRIBUTION OF AUTHORS

Text tables and figures from four published research articles are included in this thesis. Chapters 2-4 and 6 are based on manuscripts where I am the first author (for the manuscript in chapter 6 the second author has contribution equal to mine). Chapter 5 is based on a manuscript in preparation of which I am the first author. The thesis has been written in a manuscript-based format, and the texts have been re-formatted according to the general guidelines for thesis submission. The references and abbreviations of all chapters have been combined in one section; the figures and tables have been adapted from the printed versions according to the general guidelines for thesis submission. The contributions of the authors for each of the manuscripts are as follows:

Chapter 2:

Shaya, D., Li, Y., and Cygler, M. (2004) Crystallization and preliminary X-ray analysis of heparinase II from Pedobacter heparinus. *Acta Crystallogr. D.* **60:** 1644-1646.

I preformed all the experiments, analyzed the data and wrote the draft to the manuscript. Y.L. instructed me with the crystallization steps. M.C. supervised the project and provided comments and suggestions thought the course of the work. The final manuscript for publication was prepared by M.C. and myself.

Chapter 3:

Shaya, D., Tocilj, A., Li., Y., Myette, J., Venkataraman, G., Sasisekharan, R., and Cygler, M. (2006) Crystal structure of heparinase II from Pedobacter heparinus and its complex with a disaccharide product. *J Biol Chem.* **22**:15525-35.

I preformed all the experiments and analyzed the data for native heparinase II. I performed all the crystallographic experiments and the structure determination with the selenomethionine labeled heparinase II and wrote the draft of the manuscript. A.T. and Y.L. instructed me in crystallization steps and computational structure determination. Cloning, expression, purification and selenomethionine labeling of heparinase II in E. coli were preformed by J.M., and supervised by

G.V. and R.S. MC supervised the project, orchestrated the collaboration, and prepared the final manuscript for publication with my aid.

Chapter 5:

Shaya, D., Hahn, B. S., Bjerkan, T. M., Kim, W. S., Park, N. Y., Sim, J. S., Kim Y. S., and Cygler, M. (2008) Composite Active Site of Chondroitin Lyase ABC Accepting Both Epimers of Uronic Acid. *Glycobiology*. 18: 270-277.

I performed all the experiments and analyzed the data leading to the structure of chondroitin lyase ABC. I designed and constructed the active site mutants, proposed the kinetic experiments for characterization of these mutants and for defining the metal dependency of chondroitin lyase ABC activity. I have analyzed the data and wrote the draft of the manuscript. T.M.B. assisted in expression of selenomethionine labeled chondroitin lyase ABC. The kinetic characterization of chondroitin lyase ABC, its mutants and its dependence on metal cations, was proposed by me and carried out by W.S.K. and N.Y.P. and supervised by B.S.H. Cloning of chondroitin lyase ABC was carried out by J.S.S. under the supervision of Y.S.K. M.C supervised the project, orchestrated the collaboration, and prepared the final manuscript for publication with my aid.

Chapter 6:

Shaya, D.*, Hahn, B. S.*, Park, N. Y., Sim, J.S., Kim Y. S., and Cygler, M. (2008) Characterization of Chondroitin Sulfate Lyase ABC from *Bacteroides Thetaiotaomicron* WAL2926. *Biochemistry* 2008 May 31.

* These authors contributed equally to this paper

I performed the thermal stability experiments characterizing B.thetaiotaomicron *chondroitin lyase ABC in comparison to the* P.vulgaris *enzyme, designed and constructed the mutant proteins, and proposed their kinetic characterization experiments. I have proposed the experiments characterizing chondroitin lyase ABC, analyzed the data and wrote the draft of the manuscript. The kinetic characterization of chondroitin lyase ABC and its mutants was carried out by N.Y.P. and J.S.S under the supervision of B.S.H. and Y.S.K. M.C supervised the project, orchestrated the collaboration, and prepared the final manuscript for publication with my aid.*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ArthroAC	*Arthrobacter aurescens* chondroitin lyase AC
BactnABCII	*Bacteroides thetaiotaomicron* chondroitin lyase ABCII
CCD	charge-coupled device (detector)
ChonABC	chondroitin lyase ABC
ChonAC	chondroitin lyase AC
ChonB	chondroitin lyase B
CNS	central nervous system
CS	chondroitin sulfate
CS-A	chondroitin 4-sulfate
CS-B	dermatan sulfate
CS-C	chondroitin 6-sulfate
dATP	2'-deoxyadenosine 5'-triphosphate
DNA	Deoxyribonucleic acid
DS	dermatan sulfate
ECM	extracellular matrix
EDTA	ethylenediaminetetracetic acid
FOM	figure of merit
GAG	glycosaminoglycan
GlcA	β-D-glucopyranosyluronic acid
GlcN	2-amino-2-deoxy-α-D-glucopyranose
GlcNAc	2-amino-acetyl-2-deoxy-α-D-glucopyranose
GlcNS	2-amino-sulfo-2-deoxy-α-D-glucopyranose
HepI	heparin lyase I
HepII	heparin lyase II
HepIII	heparin lyase III (heparitinase)
HS	heparan sulfate
HSGAG	heparan sulfate glycosaminoglycan
HSPG	heparan sulfate proteo-glycan
IdoA	α-L-idopyranosyluronic acid
IMAC	Immobilized Metal Ion Affinity Chromatography

IPTG	isopropyl-beta-D-thiogalactopyranoside
MALDI-MS	matrix assisted laser desorption ionization mass spectrometry
MAD	multiple-wavelength anomalous diffraction
MR	molecular replacement
MIR	multiple isomorphous replacement
nHepII	native heparin lyase II
PEG	polyethylene glycol
PG	proteo-glycan
PL	polysaccharide lyase
PedAC	*Pedobacter heparinus* chondroitin lyase AC
PedHepII	*Pedobacter heparinus* heparin lyase II
PNN	perineuronal nets
PvulABCI	*Proteus vulgaris* chondroitin lyase ABCI
PvulABCII	*Proteus vulgaris* chondroitin lyase ABCII
rHepII	recombinant heparin lyase II
SAD	single anomalous diffraction
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SeMet	seleno-methionine
TLS	translation libration screw
ΔUA	4-deoxy-α- L-*threo*-hex-4-enopyranosyluronic acid

Chapter 1 General Introduction

1.1 Glycosaminoglycans and Proteoglycans

Glycosaminoglycans (GAGs) are linear, heterogeneous, highly negatively charged, acidic polysaccharides, abundant in the extracellular matrix (ECM) and on the cell surface. GAGs are made of a basic disaccharide repeating unit composed of two, six-member sugar rings, a hexosamine (D-glucosamine, GlcNAc, or D-galactosamine, GalNAc) linked to a uronic acid (D-glucuronic acid, GlcA, or L-iduronic acid, IdoA) (Ernst *et al.* 1995b; Bulow & Hobert 2006). Both sugars of the repeat unit can be variably *N*- and *O*-sulfated, resulting in complex, chemically versatile macromolecules. Apart from their structural importance to the integrity of the ECM, GAG chains serve as fundamental modulators of a wide variety of biological processes (Perrimon & Bernfield 2001). The high negative charge associated with GAGs facilitates their interaction with a large array of extracellular proteins (Capila & Linhardt 2002), while their linear structure promotes the sliding movement of bound proteins along the chains. In this manner, GAGs facilitate molecular encounters between proteins in the assembly of multicomponent complexes, reducing the three dimensional search problem to a one dimensional search (Knudson & Knudson 2001).

This thesis research is focused on the enzymatic mechanism of GAG degradation, in particular, that of the prokaryotic GAG lyases. I have studied the structures and catalysis of two broad substrate-specificity GAG-lyases, heparinase II from the soil bacterium *Pedobacter heparinus* (PedHepII) and chondroitin lyase ABCII from the human gut isolate *Bacteroides thetaiotaomicron* BactnABCII, using a combination of X-ray crystallography and kinetic analysis of active-site mutants.

To provide the background for my research I will introduce general topics regarding the biological roles, classification, structures, biosynthesis and degradation of GAGs.

GAGs are classified into four main classes based on their biosynthetic pathways and chemical composition: heparan sulfate (HS) and heparin, chondroitin sulfate (CS) and dermatan sulfate (DS), keratan Sulfate (KS) and hyaluronic Acid (HA). Two classes HS/heparin, and CS/DS are linked through a common tetrasaccharide (xylose-galactose-galactose-glucuronic acid) to serine residues of a protein core forming a proteoglycan (PG; Ruoslahti 1988; Oldberg *et al.* 1990; Kjellen & Lindahl 1991; Iozzo 1998; Esko &

Selleck 2002). The third class, KS, is attached to the core protein through an *N*-acetylglucosamine by *N*-linkage to asparagine or *O*-linkage to serine (Funderburgh 2000). The fourth class, HA is produced as a free GAG (no protein attachment) and forms high molecular weight assemblies in the ECM (Knudson & Knudson 1993).

PGs form a large and diverse group of macromolecules found in various tissues in animals. Their characteristics are defined both by their protein cores and carbohydrate constituents. PG structures are complex and versatile due to variations in the primary sequence and domain arrangement of their protein core, the type, size and composition of their GAG chains, as well as the degree and distribution of substitutions of their GAG chains (Esko & Lindahl 2001). Low molecular weight PGs (e.g, decorin and biglycan) normally contain 1-3 GAG chain constituents (Fransson *et al.* 2000), whereas multi-domain high molecular weight PGs, such as aggrecan, can contain more than a hundred GAG chains (Kiani *et al.* 2002). The size of GAG chains also varies, from a polysaccharide chain of just 10 kDa to as much as 400,000 kDa. Finally, variation in the extent and pattern of epimerization (conversion of the GlcA to IdoA) and the sulfation of the GAG chains, all contribute to the tremendous structural heterogeneity of PGs (Bulow & Hobert 2004; Lamanna *et al.* 2007).

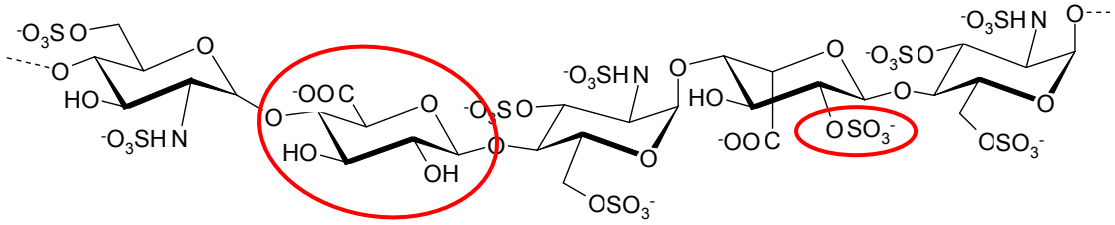
The GAG chains account for a large portion of the PGs, and through their interaction with a broad range of GAG binding proteins govern many of the diverse activities of PGs (Bandtlow & Zimmermann 2000). PGs reside on the basement membranes (i.e. syndecan, decorin, and glypican) and form a negatively charged canopy that coats virtually all metazoan cells and acts as the glue connecting cells to the ECM (Bernfield *et al.* 1992). PGs are critical environmental modulators, playing important roles in cell adhesion, migration, and differentiation, tissue morphogenesis and phenotypic stabilization of all tissue and organ systems via cell-matrix adhesiveness, cell-cell interactions, or binding to growth regulators and differentiation factors (Bandtlow & Zimmermann 2000; Funderburgh 2000; Blackhall *et al.* 2001; Filmus 2001; Kresse & Schonherr 2001; Bulow & Hobert 2006). Besides their localization on the cell surfaces, PGs are also secreted into the ECM (Iozzo & Murdoch 1996) or localized in intracellular secretion granules (Kolset & Gallagher 1990).

1.1.1 The Biological Roles of Glycosaminoglycans

GAGs serve as key biological response modifiers by acting as stabilizers, cofactors, or co-receptors for growth factors, cytokines, and chemokines (Lander & Selleck 2000; Bulow & Hobert 2006; Taylor & Gallo 2006). GAGs regulate enzyme activities and signaling molecules in response to cellular damage, such as wounding, infection and tumorigenesis (Varki 1993). Through their properties and interactions, GAGs play important roles in intercellular communication of metazoan cells (Linhardt & Toida 2004). GAGs are targets for bacterial, viral, and parasitic virulence factors for attachment, invasion, and immune system evasion (Rostand & Esko 1997; Schmidtchen *et al.* 2001). GAGs have appeared early in evolution and their prominent extracellular location have ensured their conserved roles throughout the animal kingdom.

The best characterized activity of HS and heparin is in the regulation of hemostasis through their anti-coagulation activity in the plasma-based coagulation cascade (Bourin & Lindahl 1993). A specific pentasaccharide sequence (Fig. 1-1) found in HS and heparin binds to antithrombin, enhancing its activity as an inhibitor of thrombin and several other serine proteases in the coagulation cascade (Hirsh *et al.* 2001).

Figure 1-1. AT-binding pentasaccharide.
The rare modifications that were found to be crucial for thrombin binding to the pentasaccharide are highlighted in red, namely GlcA as the 2nd sugar moiety, and 2 *O*-sulfation of IdoA the 4th sugar moiety of the pentasaccharide.



HS GAGs are also implicated as scaffold structures, designed to accommodate proteins through noncovalent binding to their GAG substituents (Kjellen & Lindahl 1991; Esko & Selleck 2002; Raman *et al.* 2005; Zhang *et al.* 2007). Binding of proteins to HS chains serve a variety of functional purposes, from simple immobilization or protection against proteolytic degradation to distinct modulation of biological activities. Protein-GAG interactions govern the critical involvement of HS-PGs in many biological phenomena at various levels of complexity, including organogenesis and tissue remodeling in embryonic development, involvement in angiogenesis, neovasculation, and metastasis in cancer processes, regulation of growth factor and cytokine action governing cell adhesion, differentiation and maturation, regulation of the immune system complement cascade, anti-inflammatory actions, and wound healing and repair (Lindahl *et al.* 1994; Nelson *et al.* 1995; Salmivirta *et al.* 1996; Rosenberg *et al.* 1997; Bali *et al.* 2001; Lander *et al.* 2002; Saadi *et al.* 2002; Sasisekharan *et al.* 2002; Arya *et al.* 2003; Peterson *et al.* 2004; Ibrahimi *et al.* 2004; Bulow & Hobert 2006; Vlodavsky *et al.* 2006; Taylor & Gallo 2006). Viruses, such as the herpes simplex virus, human immunodeficiency virus, Dengue virus, and Hepatitis C virus target HS cell surface PG receptors for attachment and invasion (Rider 1997; Chen *et al.* 1997; Barth *et al.* 2003). Bacterial pathogens including *Bordetella pertussis, Borrelia burgdorferi, Chlamydia trachomatis, Helicobacter pylori, Listeria monocytogenes, Neisseria gonorrheae,* and *Streptococcus pyogenes* bind to HS on the cell surface, which mediates their invasion (Zhang & Stephens 1992; Sawitzky 1996; Chen & Stephens 1997; Alvarez-Dominguez *et al.* 1997; Freissler *et al.* 2000; Frick *et al.* 2003). Finally, it was shown that treatment with heparin interferes with the adhesion of parasites such as *Leishmania* (an intracellular protozoan parasite that cause a variety of illnesses in humans), *Trypanosoma cruzi* (the parasite causing Chagas' disease), and *Plasmodium falciparum* sporozoites (the parasite causing malaria) with their target cells (Butcher *et al.* 1992; Frevert *et al.* 1993; Rathore *et al.* 2001; Dong *et al.* 2001), proving that these parasites have evolved the capacity to attach and invade cells through HS cell surface PGs as well.

Heparin is released from specialized storage granules in mast cells upon de-granulation in an allergic response and it is believed to have a unique and possibly defensive role in response to infectins and parasitic diseases (Nader *et al.* 1999).

CS-PGs are abundant within the ECM and are expressed in the central nervous system (CNS), particularly in the developing embryonic brain (Condic *et al.* 1999; Bandtlow & Zimmermann 2000; Asher *et al.* 2001; Sugahara & Mikami 2007; Purushothaman *et al.* 2007). CS-PGs can either promote neurite outgrowth (Faissner *et al.* 1994; Brittis & Silver 1995; Li *et al.* 2007), or inhibit axonal sprouting and neurite outgrowth (McKeon *et al.* 1991; Condic *et al.* 1999; Tropea *et al.* 2003) in various contexts, they were also shown to stabilize new synapse formation (Dityatev & Schachner 2003). CS-PGs are up-regulated in CNS lesions (Johnson 1993; Morgenstern *et al.* 2002; Sugahara & Mikami 2007; Galtrey & Fawcett 2007) and their degradation by chondroitin lyase ABC (ChonABC) enhances axon regeneration (Yick *et al.* 2000; Moon *et al.* 2001; Bradbury *et al.* 2002) suggesting a role in recovery from injury as well. CS-PGs were reported to posses a hallmark role in the establishment of condense ECM structures in the CNS termed *Perineuronal Nets* (PNN). PNNs serve lead roles in maintaining the plasticity of the CNS manifested during postnatal development and adult CNS-learning and memory processes (Galtrey & Fawcett 2007).

Arthritis in the connective tissues, including cartilage and skeletal systems, has been associated with degradation of the ECM and a resulting lack of DS-/CS-PGs, which are important components of the articular cartilage (Aigner & McKenna 2002). Furthermore, models of experimentally induced PG deficiencies have linked mutant mice lacking various PGs with abnormal development and pathological phenotypes (Schwartz & Domowicz 2002; Ameye & Young 2002).

DS adds biological flexibility to many normal and pathological responses, such as development, growth, wound repair, infection, and tumorigenesis (Trowbridge & Gallo 2002; Taylor & Gallo 2006). Biological activities specifically assigned to DS include anti-coagulative activity in hemostasis through heparin cofactor II (HCII) mediated thrombin inhibition (Liaw *et al.* 1999; Liaw *et al.* 2001), stimulation of cell growth in vitro through promotion of fibroblast growth factor-2 (FGF2), FGF7, and hepatocyte growth factor/scatter factor activity (Lyon *et al.* 1998), and tenascin-X mediated stability of the connective tissue critical to establishing the normal tensile strength of the skin (Elefteriou *et al.* 2001). DS is the predominant GAG expressed in the skin and is released at high concentration during wound repair (Penc *et al.* 1998; Taylor & Gallo 2006).

Differential expression of specific DS-PGs have been associated with increased scarring (Gallo *et al.* 1996). Bacteria such as *Streptococcus pyogenes* and *Borrelia burgdorferi* (the bacterial agent of Lyme disease) were shown to bind DS-PGs, facilitating their adhesion and penetration through skin (Brown *et al.* 2001; Frick *et al.* 2003). Pathogenesis of *Pseudomonas aeruginosa*, *Enterococcus faecalis*, and *Streptococcus pyogenes* involves the release of proteases that degrade DS-PGs (Schmidtchen *et al.* 2001); the free DS binds to and inactivates the neutrophil-derived cationic antimicrobial peptide, α-defensin (short cationic peptide essential for resistance to infection), which in turn provides a defense mechanism and increases the virulence of this pathogen.

The high abundance of KS in cornea was shown to maintain tissue hydration critical for the corneal transparency (Bettelheim & Plessy 1975). When the GAG balance (KS-DS) in the corneal stroma is altered (i.e. macular corneal dystrophy, an autosomal recessive hereditary eye disease) the tissue becomes opaque, leading eventually to the loss of vision (Hassell *et al.* 1980). Recent studies have presented data suggesting that KS is also an active participant in cellular processes regulating cellular recognition of protein ligands, cells adhesiveness and motility, axonal guidance, and embryo implantation (Burg & Cole 1994; Graham *et al.* 1994; Olsson *et al.* 1996; Funderburgh *et al.* 1997; Greenwood & Murphy-Ullrich 1998; Davies *et al.* 1999). KS attachment to the CS/DS-PG aggrecan blocks the immune response to this protein, suppressing development of antigen-induced osteoarthritis (Guerassimov *et al.* 1999).

Hyaluronan (hyaluronic acid, HA), a unique, high molecular weight, viscous GAG, orders water in the extracellular environment providing both structure and flexibility to tissues (Delpech *et al.* 1997). In addition to its mechanical function related to its tensile properties, low molecular weight HA (short sequences of HA released after injury and during inflammation) is involved in the complex carbohydrate metabolism related to fundamental biological processes such as cell migration and proliferation, tissue remodeling and reconstruction during embryogenesis, angiogenesis and tumor invasion (Knudson & Knudson 1993; Toole 2001; Toole 2004; Toole & Slomiany 2008). A hallmark role for HA involvement in the granulating tissue formed in the process of wound repair has been described (Weigel *et al.* 1986). More recently, it was suggested that fetal scarless wound repair is linked to prolonged existence of the granulation tissue

enriched by HA (Gailit & Clark 1994). As HS and CS GAGs, HA presents versatile roles in the immune system during inflammation (Siegelman *et al.* 1999; Johnson *et al.* 2000; Termeer *et al.* 2003; Mummert 2005).

1.1.2 The Chemical Compositions and Structures of Glycosaminoglycans

1.1.2.1 Heparan Sulfate and Heparin

HS polysaccharides are composed of alternating hexuronic acid (GlcUA) and glucosamine residues, which may be *N*-sulfated ($GlcNSO_3$), *N*-acetylated (GlcNAc), or in rare cases unmodified ($GlcNH_3$) (Gallagher & Turnbull 1992; Yanagishita & Hascall 1992). Another prominent modification found in HS-GAGs is epimerization of GlcA to IdoA. HS disaccharides are further modified by *O*-sulfation at position 6 of the glucosamine and position 2 of the uronic acids. As much as 23 chemically distinct disaccharides have been identified to date in HS and its biosynthetic intermediates (Esko & Lindahl 2001).

Compositional analysis of HS-GAGs revealed that chain modifications are arranged in macroscopic domains, resulting in *N*-sulfated oligosaccharide stretches of variable length (NS domains) flanked by regions of alternating *N*-acetylated and *N*-sulfated units (NA/NS domains). Those domains are interspersed among long, contiguous tracts of unmodified *N*-acetylated disaccharides (the NA regions), devoid of IdoA units (Maccarana *et al.* 1996; Murphy *et al.* 2004). Since other modifications, such as *O*-sulfation and epimerization of D-GlcA to L-IdoA, depend on prior *N*-sulfation of GlcN units, the modified disaccharide units tend to cluster in the NS or NA/NS domains (Lindahl *et al.* 1998). The NS domains present constant amounts of IdoA as well as 2-*O*-sulfation, regardless of the PG core or the parent tissue the GAG was isolated from. On the other hand, the spacing between the NS-domains, their specific *O*-sulphation patterns, as well as the relative abundance of the NA/NS domains appear to be tissue-specific, characteristic of the cells from which the HS was obtained (Edge & Spiro 1990; Lyon *et al.* 1994; Kato *et al.* 1994; Tekotte *et al.* 1994; Lindahl *et al.* 1995; Lindahl *et al.* 1998; Feyzi *et al.* 1998; Dietrich *et al.* 1998; Esko & Selleck 2002).

Heparin, considered as a specialized HS-PG (Esko & Selleck 2002; Linhardt 2003), is uniquely found within intracellular granules of mast cells, linked to the core

protein serglycin (Kolset & Gallagher 1990; Humphries *et al.* 1999; Feyerabend *et al.* 2006). Heparin is the most highly sulfated GAG, with an average of 2.7 sulfate groups per disaccharide. The glucosamine moiety in heparin is typically *N*- sulfated and bears as well 6-*O*- sulfation, while the uronic acid is typically 2-*O*-sulfated. Heparin is also characterized by its unusually high content of IdoA (up to ~90%) in comparison to HS.

1.1.2.2 Chondroitin Sulfate and Dermatan Sulfate

CS and DS are characterized by galactosamine, as their hexosamine component, linked to uronic acid. CS consists of a GlcA(1→3)GalNAc disaccharide-repeating unit connected through β(1→4) glycosidic linkages (Galtrey & Fawcett 2007). CS is regio-selectively *O*-sulfated at position 4 (chondroitin 4-sulfate, chondroitin sulfate A) or 6 (chondroitin 6-sulfate, chondroitin sulfate C) of the GalNAc (Habuchi 2000).

DS, considered as the third type of CS (chondroitin sulfate B), contains variable amounts of IdoA and somewhat higher levels of sulfation in position 4 or 6 of the GalNAc, and position 2 of the IdoA (Halldorsdottir *et al.* 2006). Typical DS (isolated from porcine skin) contains ~1.1 sulfate groups per disaccharide and as much as 80% IdoA (Mascellani *et al.* 1993). It is widely accepted that the variable chain lengths of DS, disaccharide composition, and sulfation, determine the binding affinity and control the functional interactions with potential protein partners (Bernfield & Sanderson 1990; Kim *et al.* 1994; Litwack *et al.* 1998).

1.1.2.3 Keratan Sulfate

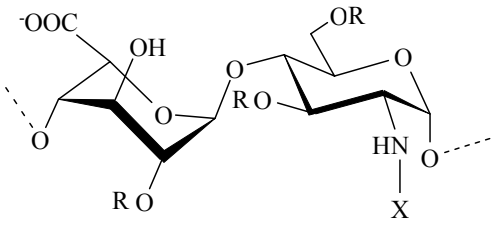
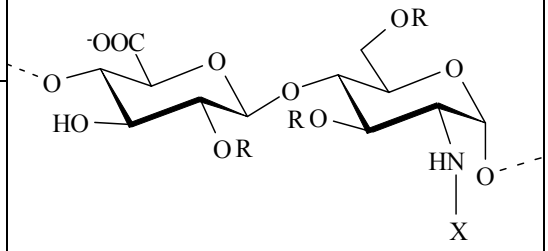
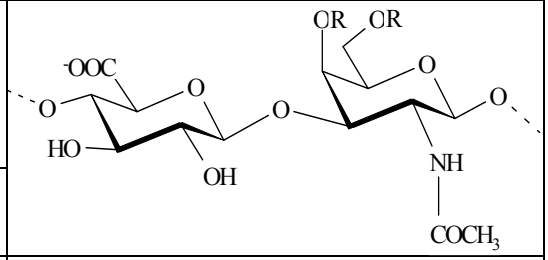
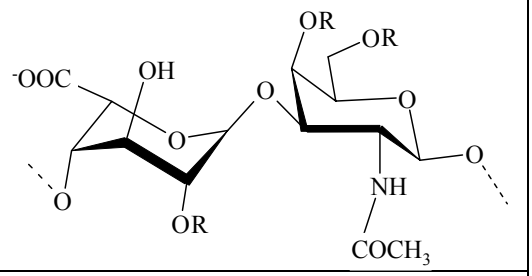
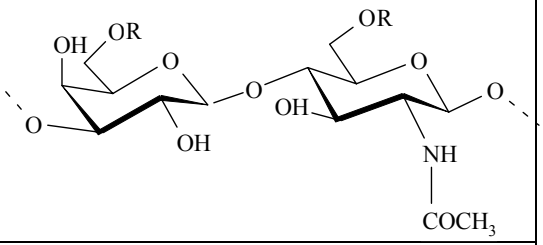
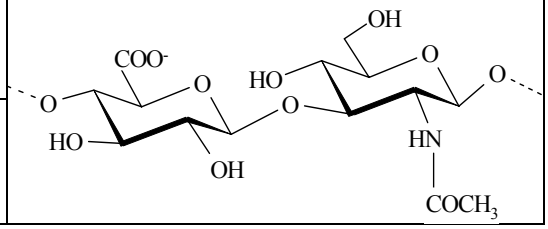
KS-PGs differ from the other GAGs by their lack of uronic acid; rather they are composed of *N*-acetyllactosamine →3) Galβ (1→4) GlcNAcβ (1→, as their repeating dissacharide unit with a galactose moiety in place of the uronic acid, and by their linkage to the protein cores. KS-GAGs are classified into three types, KSI, which includes all Asn-linked KS molecules, KSII, which refers to all KS linked to protein through GalNAc-O-Ser/Thr (mucin type "core 2" oligosaccharide), and KSIII molecules, which are linked via Man-O-Ser to their protein cores (Krusius *et al.* 1986; Funderburgh 2000; Bulow & Hobert 2006). All linkage types present complex biantennary structures (two-chains). Some repots described two KS chains extended from the same oligosaccharide

linkage, with position 6 of the third mannose group normally extended by KS chains, and the second chain extended from position 3 of the same mannose group (Nilsson *et al.* 1983)). The nonreducing termini of KS chains are capped with a variety of structures. About 70% of KS chains are capped with neuraminic acid, the remainder with β-GalNAc or α-Gal (Tai *et al.* 1996; Tai *et al.* 1997). KSII includes *O*-linked Lewis x structures bearing sulfation only at the non-reducing terminus and fucosylation at the terminal GlcNAc (Hemmerich *et al.* 1995). Other modifications found in KSII include partial sialylation of Gal linked to C-3 of the linkage GalNAc, capping by sialic acid at C-3 or C-6 of the terminal GlcNAc, and fucosylation at C-3 of the sulfated GlcNAc (Brown *et al.* 1996).

1.1.2.4 Hyaluronan

Hyaluronan a simple polymer of unmodified GlcNAc and GlcA disaccharides is the only known GAG produced as a free polysaccharide with no attachment to a protein core. It is made of disaccharide repeated units of →3) D-GlcNac (β1→4) D-GlcUA (β1→ (Knudson & Knudson 1993; Bulow & Hobert 2006). HA polysaccharides are abundant in the ECM, as part of high molecular weight protein-GAG complex networks forming the physical structure of the ECM. HA polysaccharides are extremely long and can reach the size of 25,000 disaccharide units or 10^7 Da (Knudson & Knudson 2001; Bulow & Hobert 2006).

Table 1-1. GAG disaccharide-repeat building blocks

GAG	Hexauronic acid[a]	Hexosamine	Disaccharide repeat
Heparin	L-iduronic acid, IdoA	D-glucosamine, GlcNAc	
	→4)IdoAα(1→4)GlcNXα(1→		
Heparan sulfate	D-glucoronic acid, GlcA	D-glucosamine, GlcNAc	
	→4)GlcAα(1→4)GlcNXα(1→		
Chondroitin sulfate	D-glucoronic acid, GlcA	D-galactosamine, GalNAc	
	→4)GlcAβ(1→3)GalNAc β(1→		
Dermatan sulfate	L-iduronic acid, IdoA / D-glucoronic acid, GlcA	D-galactosamine, GalNAc	
	→4)IdoAα(1→3)GalNAcβ(1→		
Keratan sulfate	Galactose, Gal	D-glucosamine, GlcNAc	
	→3)Galβ(1→4)GlcNAcβ(1→		
Hyaluronic acid	D-glucoronic acid, GlcA	D-glucosamine, GlcNAc	
	→4)GlcAβ(1→3)GlcNAc β(1→		

[a] In the case of KS it is hexose (galactose) rather than hexauronic acid; R = SO_3, H

1.1.3 Glycosaminoglycans, GAG-lyases and Human Health

Over the years GAGs and their lyases have found numerous uses as tools in research and as clinical agents. The most prevalent use of GAGs in the clinic is the use of heparin as an anti-coagulation agent. Its utilities range from a prophylactic drug used in prevention of post-operative thrombosis, to maintenance of blood in anti-coagulative form during extra-corporal treatments (Hirsh *et al.* 2001; Hirsh & Raschke 2004). In this context, prior to restoration of the blood back into circulation, heparin levels are reduced using heparinase I treatment (Langer *et al.* 1982). This enzyme is used, as well, to monitor heparin levels in the blood during medical procedures when heparin is administered (Ameer *et al.* 1999; Miller *et al.* 2000). The heparinases were shown to inhibit neovasculation and the proliferation of capillary endothelial cells and are under clinical investigations (Sasisekharan *et al.* 1994; Liu *et al.* 2002).

Due to the high similarity between heparin and HS it was shown that in many cases heparin is able to slow and even inhibit infections by pathogens that use HS receptors for adhesion to human cells (HIV, HSV, malaria, etc.). Severe hemorrhages limit heparin administration to prevent those infections, yet in specific cases heparin is still used to alleviate life threatening parasitic infection, such as malaria. Research in the clinical field is directed toward producing heparin / HS derivatives that will lack the anti-coagulation effects but will still coat HS receptors in an efficient manner to prevent infections (Shriver *et al.* 2004; Rusnati *et al.* 2005).

Chondroitin, the non sulfated derivative of CS is widely used as a therapeutic material for the prevention or alleviation of symptoms of diseases, such as rheumatoid and osteo-arthritis (Johnson *et al.* 2001). It is used as well in food supplements (McAlindon *et al.* 2000).

A potential medical application of ChonABC has been demonstrated during the past few years in the field of neurology. Treatment with ChonABC enhances axonal regeneration following spinal cord injury in model systems by effectively removing CS from CS-PGs enriched tissue in the glial scar that develops after injury, thus indicating that CS-PGs inhibit axon regeneration. Researches have reported upregulation of regeneration-associated proteins in the injured neurons, axons regeneration promotion, restored post-synaptic activity, and promoted functional recovery of locomotor and

proprioceptive behaviors, in response to treatment with ChonABC (Bradbury *et al.* 2002; Chau *et al.* 2004; Caggiano *et al.* 2005). The goal of healing paralysis generated by spinal cord injury and nerve damage may be far, yet, understanding the complex processes involved in tissue scarring / healing, and axon guiding will be the first achievements resulting from this research direction.

1.2 Biosynthesis and Degradation of Glycosaminoglycans

GAG turnover is intricately connected to their biological activities, (and hence the biological activities of PGs), and was shown to play a fundamental role in processes such as cell mobility, inflammation, metastasis and angiogenesis, where GAGs and their degradation products play regulatory roles (Taylor & Gallo 2006).

1.2.1 Glycosaminoglycan Synthesis

GAG assembly, apart from HA, is initiated in the endoplasmic reticulum and ends in the Golgi apparatus, where the various modifications take place. The *O*-linked GAG chains (CS/DS and HS/heparin) are initiated by xylose addition after the extrusion of the core protein into the lumen of the ER is completed (Sugahara & Kitagawa 2000). The xylose residue is covalently attached to the hydroxyl of a serine residue embedded in a specific peptide sequence defined by Ser-Gly residues flanked by one or more acidic residues (Bourdon *et al.* 1987; Esko & Zhang 1996). The synthesis continues through the assembly of the linkage tetrasaccharide GlcAβ(1→3)Galβ(1→3)Galβ(1→4)Xylβ(1→Ser on the core polypeptide (Sugahara & Kitagawa 2000). Elongation of the tetrasaccharide linkage region is catalyzed by distinct glycosyltransferases, each specific with respect to its cognate acceptor, donor and linkage formed. After assembly of the linkage region, GlcNAc transferases add a single (1→4)-linked GlcNAc unit to the chain; this step results in commitment of the intermediate to the assembly of HS/heparin or CS/DS. There is a competition between addition of α or β (1→4)-linked GalNAc. The α type link commits the chain to HS/heparin elongation whereas the β type link commits the chain for CS/DS formation. It appears that amino acid sequence determinants lying close to the GAG attachment site or structural domains at some distance away regulate this process, with CS representing the default pathway (Esko & Zhang 1996).

1.2.1.1 Chondroitin Sulfate and Dermatan Sulfate

Upon completion of the linkage region and addition of the first GalNAc, CS polymerization continues in the Golgi apparatus by the alternate addition of GlcA and GalNAc residues. Both glycosyl transfers are preformed by a single enzyme, chondroitin synthase, and require a chondroitin polymerizing factor, with both proteins cooperatively extending the growing polysaccharide chain (Kitagawa *et al.* 2001; Kitagawa *et al.* 2003). After the polymerization is complete, chondroitin chains are sulfated in several positions, including GalNAc 4-sulfate (chondroitin 4-sulfate, CS-A), GalNAc 6-sulfate (Chondroitin 6-Sulfate, CS-C), CS with both GalNAc 2-sulfate and GalNAc 6-sulfate in the same GAG chain (CS-D), and CS with occasional GalNAc 4,6-disulfate (CS-E). Sulfation of GlcA at position 2 or 3 is found in all CS species in variable amounts, although is less frequent than found in DS. There are also variable amounts of unsulfated GalNAc residues in all CS-PGs (Sugahara *et al.* 1996; Nadanaka & Sugahara 1997) .

DS is derived from chondroitin by epimerization of GlcA residues to IdoA residues during, or subsequent to, the formation of chondroitin and concomitant with its 4-sulfation at the GalNAc residue (Galtrey & Fawcett 2007). In addition, many of the IdoA residues in DS are 2-sulfated, while *O*-sulfation at position 6 of GalNAc also occurs, but at lower frequency. IdoA 2-*O*-sulfatation adjacent to GalNAc 4-sulfate in DS was shown to be the last step of DS biosynthesis, and is presumed to occur in a relatively late Golgi compartment (Kobayashi *et al.* 1999).

1.2.1.2 Heparan Sulfate and Heparin

The diverse structures of HS are generated through the alternating addition of GlcAβ(1→4) and GlcNAcα(1→4) residues, catalyzed by proteins that were initially identified as members of the exostosin family of tumor suppressors (McCormick *et al.* 2000; Senay *et al.* 2000; Duncan *et al.* 2001; Busse *et al.* 2007). The precursor polysaccharide, *N*-acetyl heparosan, has the chemical formula $[GlcA\text{-}GlcNAc]_n$. This precursor is subsequently modified by different enzymes (Rosenberg *et al.* 1997). The modifications are initiated with *N*-deacetylation/*N*-sulfation typically converting 40–50% of the GlcNAc residues to *N*-sulfo-glucosamine (GlcNS). Subsequently, GlcA moieties adjacent to the GlcNS units are frequently epimerized to IdoA. The resulting

GlcNS/IdoA-rich domains then undergo *O*-sulfation at position 2 of the IdoA and positions 6 and 4 of the GlcNS. Not all monosaccharides undergo *O*-sulfation, leading to a considerable variation in sulfation patterns within these NS domains (Gallagher & Turnbull 1992; Maccarana *et al.* 1993). Furthermore, it was recently demonstrated that extracellular sulfatases modify HS by removing 6-O sulfate groups, which in turn confer changes in its biological activities on the cell surfaces (Dai *et al.* 2005; Lamanna *et al.* 2006). Experimental evidence points to the coupling of the various modification steps (Pinhal *et al.* 2001; Holmborn *et al.* 2004; Presto *et al.* 2008; Carlsson *et al.* 2008), leading to speculation of the existence of a "GAGosome", a physical complex of enzymes (and potentially other proteins) committed to the assembly of HS chains (Esko & Selleck 2002; Lamanna *et al.* 2007). The concerted action of the membrane-bound modifying enzymes catalyzing these reactions results in the formation and organization of NA, NS, and NA/NS domains in a tissue specific manner (Lindahl *et al.* 1998). Moreover, the modifications follow unique patterns, creating specialized sequences embedded within the GAG chains, designed to specifically interact with target proteins responsible for the wide-ranging biological influence of HS-GAGs (Maccarana *et al.* 1996).

Heparin is primarily synthesized and stored in mast cells linked to a specialized protein forming the PG serglycin. Most of the enzymatic machinery used for HS biosynthesis is used for the synthesis of heparin as well, while some modifying enzymes are specific for heparin (Ledin *et al.* 2006). Bioactive short heparin oligosaccharides, termed heparin GAGs, are released by the action of heparin hydrolases in response to external signals and circulate through the blood system effecting the immune system and biological activities such as inflammation.

1.2.1.3 Keratan Sulfate

KS is elongated *via* the action of glycosyltransferases that alternately add Gal and GlcNAc to the growing polymer. There is evidence that two different glycosyl transferases, β-1,4-galactosyltransferase and N-acetylglucosaminyltransferase, are responsible for the polymerization of KS GAGs. The exact identity of the enzymes has yet to be determined. Sulfation of KS is carried out by several distinct sulfotransferase

enzymes (Ruter & Kresse 1984; Habuchi *et al.* 1996; Fukuta *et al.* 1997). One of these will sulfate GalNAc moieties of CS and also Gal in KS. It was shown that chain elongation and sulfation are coordinated during biosynthesis of KS, and further suggested that KS GlcNAc sulfation may occur simultaneously with the elongation at the terminus of the growing chain (Degroote *et al.* 1997; Uchimura *et al.* 1998).

1.2.1.4 Hyaluronan

HA is synthesized at the cellular membrane and is then released as a soluble polysaccharide and distributed within the ECM (Prehm 1983a; Prehm 1983b). HA synthases (HASs), the bi-functional glycosyl transferases responsible for HA polymerization, are divided into two classes. Class I, the streptococcal, vertebrate, and viral HASs, are integral membrane proteins with four to six predicted trans-membrane domains and one to two membrane-associated domains (Weigel *et al.* 1997). HASs contain a single glycosyltransferase-2 module, GT2 (www.cazy.org/), which is similar to the structures of chitin synthases. It was shown that the protein active sites are intracellular (Heldermon *et al.* 2001). HASs appear to both catalyze sugar transfer and facilitate HA transfer across the membrane (Weigel & DeAngelis 2007). Class II HAS has so far been found only in *Pasteurella multocida* (livestock and avian pathogen) and shown to be a peripheral membrane protein with two GT2 modules. This HAS elongates HA on the nonreducing terminus of the polysaccharide (as opposed to all the other known HASs, (Weigel & DeAngelis 2007)).

1.2.1.5 Glycosaminoglycans in Prokaryotes

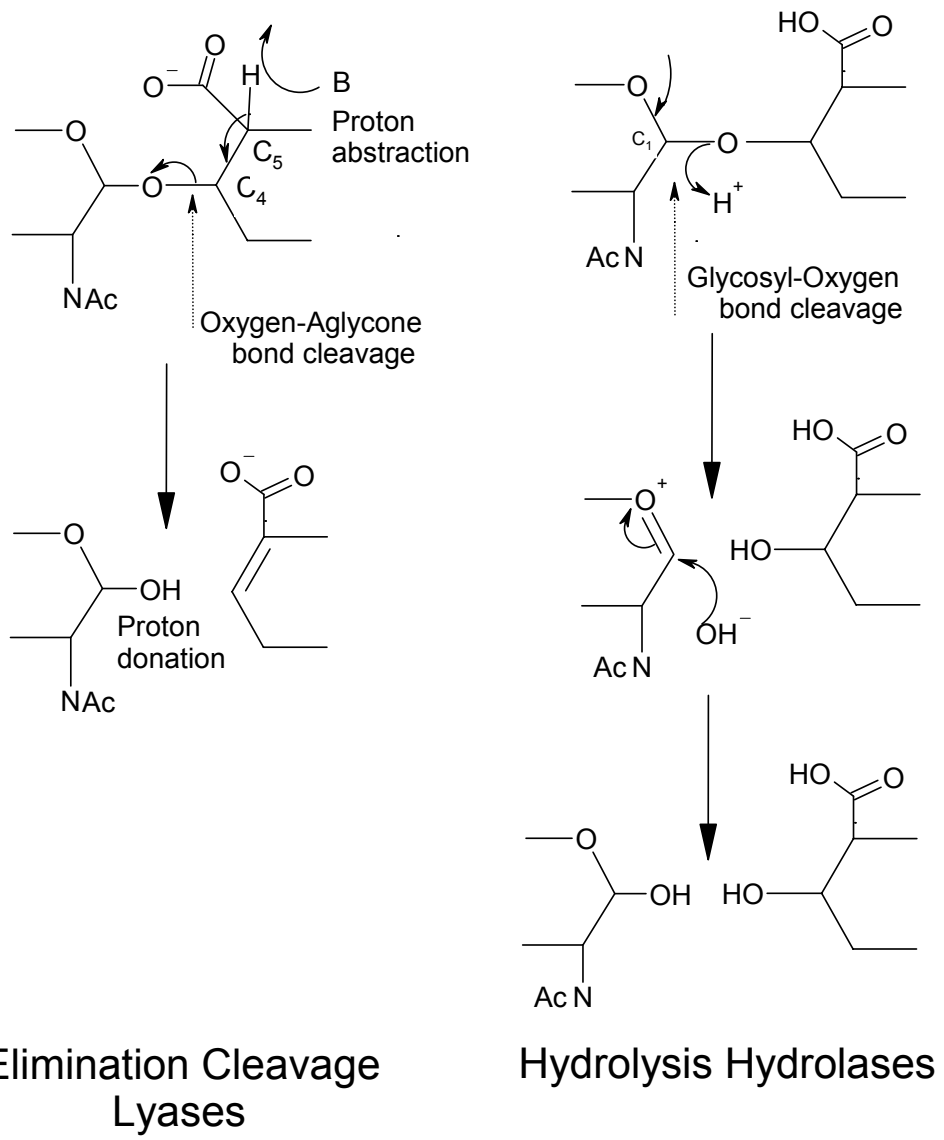
Certain pathogenic bacteria produce extracellular coatings, called capsules, which are composed of polysaccharides mimicking the metazoan GAGs (Ovodov 2006a; Ovodov 2006b). Capsules increase bacterial virulence during infection by shielding the microbe from the host immune defenses or by modulating the host's physiology (Bouchet *et al.* 2003a; Bouchet *et al.* 2003b; Vimr *et al.* 2004; Comstock & Kasper 2006). The bacterial "GAG" capsules are important for the survival of bacteria by preventing desiccation and sequestering compounds from the environment; they also adhere to surfaces or to other cells (DeAngelis 2002). In contrast to animal GAGs, the known

bacterial “GAGs” are not linked to a protein core and they lack the sulfation and epimerization commonly found in the animal GAGs.

1.2.2 Glycosaminoglycan Degradation

Two distinct chemical mechanisms have evolved in order to break down GAGs: hydrolysis and β-elimination. Mammalian enzymes employ solely hydrolytic mechanisms to cleave the glycosyl-oxygen (C1–O) bond. This double displacement mechanism involves the net addition of a water molecule, affording saturated oligosaccharide products (Lairson *et al.* 2008). Bacterial enzymes, on the other hand, degrade GAGs either through hydrolysis or by a β-elimination mechanism (lyases). GAG lyases, performing β-elimination, act exclusively on the oxygen-aglycone (O-C4) bond at the non-reducing end of the uronic acid and yield products with an unsaturated C4-C5 double bond at the uronic acid. From a chemical point of view, the proposed elimination is a proton acceptance and donation mechanism (PAD) involving four main steps. Upon binding of the polysaccharides to the protein, the negative charge of the uronic acid carboxyl group is neutralized, presumably by interaction with a positively charged amino acid or a metal ion, thereby reducing the acidity of the C5 proton. Next, a side-chain acting as a general base abstracts the proton from the C5 position of the uronic acid (also termed Cα), which leads to the elimination of the 4-*O*-glycosidic bond (elimination at position C4, also termed Cβ). Concomitantly or in a concerted manner, a double bond between C4 and C5 of the uronic acid is formed. At the last step, a general acid from the protein donates a proton to the O^- leaving group of a glucosamine, reconstituting the OH functional group at the reducing end of the cleaved bond and the products are released.

Figure 1-2 GAG degradation, β-elimination vs. hydrolysis



1.3 Bacterial GAG Lyases

Bacteria have evolved the capacity to degrade animal GAGs in order to facilitate their invasion of host tissues. Other types of bacteria, mutualists residing in animal digestive tracts, utilize GAG degradation as a regular digestion process providing a source of carbon for nutritional purposes, both for themselves and for their metazoan hosts (Comstock & Coyne 2003). GAG degradation is carried out by polysaccharide degrading enzymes utilizing two modes of polysaccharide-degradation, hydrolysis or β-elimination. Polysaccharide hydrolases are better understood than the lyases, and their

reaction mechanism is well characterized (Davies & Henrissat 1995). The lyases have been characterized as to their substrate specificities and modes of action (Jandik *et al.* 1994; Gu *et al.* 1995; Ernst *et al.* 1998; Rhomberg *et al.* 1998b; Pritchard *et al.* 2000), but their reaction mechanisms are poorly understood at the molecular level.

1.3.1 GAG Lyases Categorization

GAG lyases belong to the polysaccharide lyase super family (PL, E.C No. 4.2.2.-). Based on their primary sequences, PLs have been categorized into 18 families in the Carbohydrate Active Enzymes (CAZy) database (Coutinho & Henrissat 1999). GAG lyases have been assigned to five PL families: chondroitin lyase B (ChonB, E.C. No. 4.2.2.4) belongs to PL6, hyaluronan lyase (Hyal, E.C. No. 4.2.2.1), chondroitin AC Lyase (ChonAC, E.C. No. 4.2.2.5), and chondroitin ABC Lyase (ChonABC, E.C. No. 4.2.2.20) are all assigned to PL8, heparinase III (HepIII, E.C. No. 4.2.2.8) belongs to PL12, heparinase I (HepI, E.C. No. 4.2.2.7) is part of PL13, whereas Heparinase II (HepII, No assigned E.C. No.) is among the non-classified sequences due to its poor sequence homology to the other polysaccharide lyases.

1.3.2 GAG Lyases - Stringent Specificity vs. GAG Class Specificity

The majority of GAG lyases degrade glycosidic bonds with absolute uronic acid epimer specificity towards either GlcA or IdoA. Thus, ChonAC specifically degrades the glucoronic acid-containing glycosidic bonds present in CS-A and CS-C (Capila *et al.* 2002), ChonB degrades solely DS (Michelacci & Dietrich 1974; Michelacci & Dietrich 1975; Jandik *et al.* 1994; Gu *et al.* 1995), HepI is specific toward heparin, cleaving next to IdoA of a tri-sulfated disaccharide (LeBrun & Linhardt 2001), and HepIII degrades the next to a monosulfated GlcA bound to an unsulfated GlcNAc (LeBrun & Linhardt 2001). On the other hand, ChonABC and HepII, both degrade all GAGs belonging to one class/type of GAG. Thus ChonABC shows activity towards C4S, C6S and DS (Hamai *et al.* 1997), whereas HepII degrades all possible glycosidic bonds appearing in HS and heparin (Nader *et al.* 1990). Most striking is the ability of both enzymes to process the glycosidic bond regardless of the uronic acid epimer present (GlcA as well as IdoA). As mentioned above, two crucial steps initiate the β-elimination-catalyzed degradation,

acidic charge neutralization and proton abstraction. Through accepting both uronic acid epimers as substrates, ChonABC and HepII, posses a unique stereochemical ability to neutralize the charge and abstract the proton from either side of the C5 chiral center.

Continuing research efforts have resulted in crystal structures of enzyme-substrate/product complexes, combined with site-directed mutagenesis data to generate a complete description for polysaccharide binding and degradation mechanisms for pectate lyases C (Scavetta *et al.* 1999) and B (Herron *et al.* 2000), ChonB (Huang *et al.* 1999; Michel *et al.* 2004), ChonAC (Fethiere *et al.* 1999; Huang *et al.* 2001; Capila *et al.* 2002; Lunin *et al.* 2004a; Rye *et al.* 2006), and hyaluronate lyases (Li *et al.* 2000; Ponnuraj & Jedrzejas 2000; Pritchard *et al.* 2000). The following text will summarize the findings regarding the epimer-specific lyases, ChonAC and ChonB, as models for two general strategies evolved by GAG lyases to degrade GAGs.

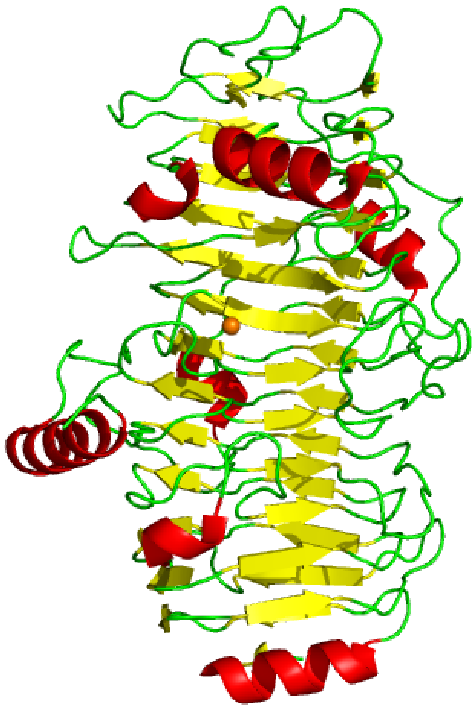
Throughout the text, the nomenclature of Davies *et al.* (Davies *et al.* 1997) has been adopted for the sugar-binding subsites, where "+" subsites refer to those toward the reducing end of the substrate, and "–" subsites indicate those toward the nonreducing end of the polysaccharide chain. Cleavage occurs between sugars occupying the "+1" and "–1" subsites.

1.3.3 Variety of GAG Lyase Folds

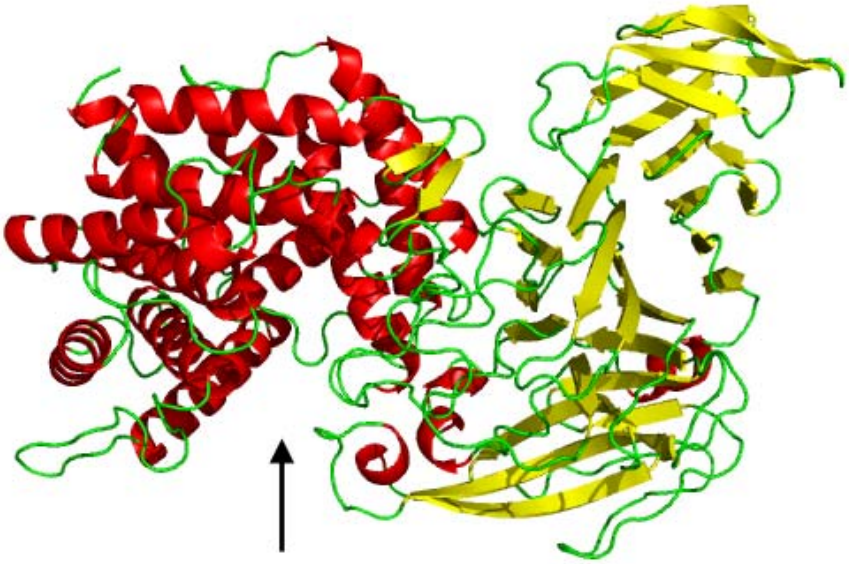
The structures of PL8 proteins *Pedobacter heparinus and Arthrobacter aurescens* ChonAC (Fethiere *et al.* 1999; Lunin *et al.* 2004), *Streptococcus pneumonia* (Jedrzejas *et al.* 1998; Li *et al.* 2000; Ponnuraj & Jedrzejas 2000) and *Streptococcus agalactiae* Hyal (Jedrzejas & Chantalat 2000; Mello *et al.* 2002), *Proteus vulgaris* ChonABCI (Huang *et al.* 2003), and PL6 *Pedobacter heparinus* ChonB (Huang *et al.* 1999) were solved by X-ray crystallography. In topological terms, they revealed two protein-fold families. PL6 proteins such as ChonB and Pectin/Pectate lyases belong to the right-handed parallel β-helix family, which also contains polysaccharide hydrolases (Jenkins *et al.* 1998). The right-handed parallel β-helices are characterized by a series of helical turns each composed of three β-strands and connecting loops, stacked parallel to each other along one axis (Fig. 1-3 (Huang *et al.* 1999)). Topologically, the right-handed β-helices contain substrate-binding sites in roughly the same location. The second topological family,

termed the α+β lyase fold, includes the PL8 proteins ChonAC, ChonABC, Hyal, and also alginate lyase A1-III from *Sphingomonas* species A1 (Yoon *et al.* 1999). The α+β lyase architecture is characterized by 3-4 structural domains, prominent among them is an (α/α) incomplete double-layered toroid connected to an anti-parallel β-sandwich. Some of these proteins also include N-terminal auxiliary domains that have an overall architecture typical for glycosyl processing enzymes / modules (Huang *et al.* 2003b). The toroidal motif varies between α_6/α_6, α_6/α_5 or α_5/α_5 topology, differing in the number of helices and their orientation toward each other. This structural motif was also identified in the structures of alginate lyases belonging to the PL5 family, (represented by *Sphingomonas sp.* alginate lyase A1-III, PDB codes 1QAZ and 1HV6 (Yoon *et al.* 1999)). The helices of the α-helical double-layered toroid are packed in such a way that one end of this domain is wider than the other. At the wider end, an elongated cleft is formed harboring the substrate-binding and active site (Fig. 1-3B-D).

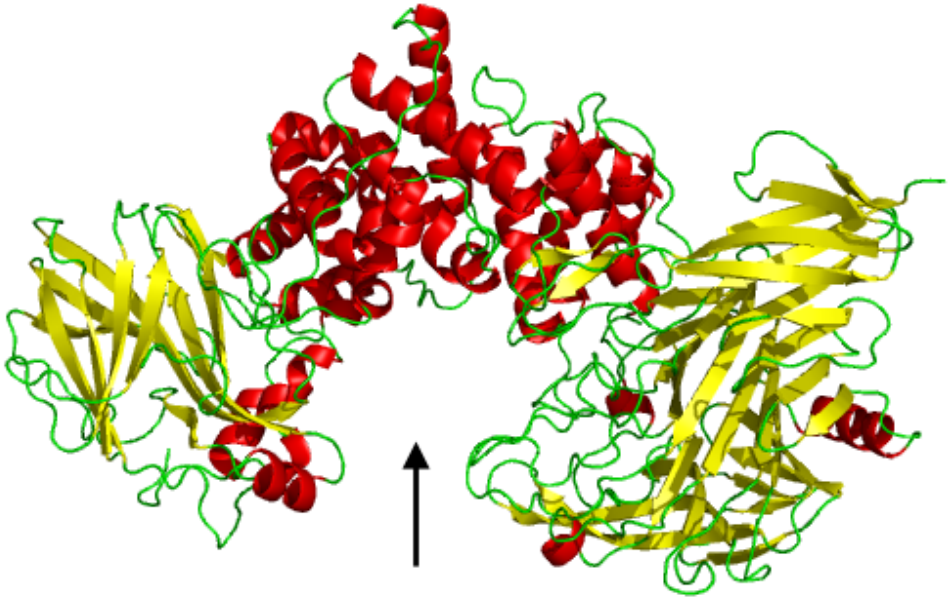
Figure 1-3. Representative structures of PL6 and 8 enzymes.
A. *Pedobacter heparinus* ChonB (PDB accession code 1DBO, (Michel *et al.* 2004) representing the right-handed β-helix polysaccharide lyase fold. The orange sphere represents the Ca^{2+} ion, originating from the complex structure (enzyme-metal ion-substrate) and denotes the active site location.



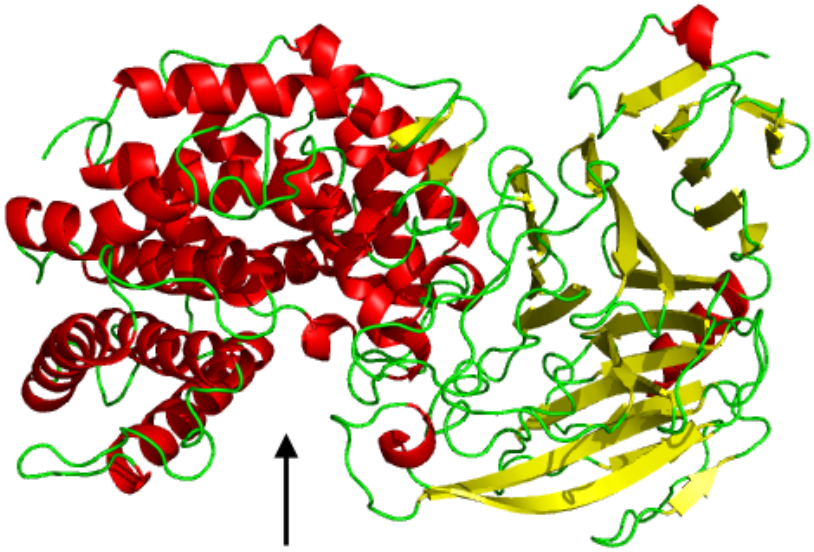
B. *Arthrobacter aurescens* ChonAC (PDB accession code 1RWG, (Lunin *et al.* 2004a)). Black arrow denotes the substrate-binding cleft.



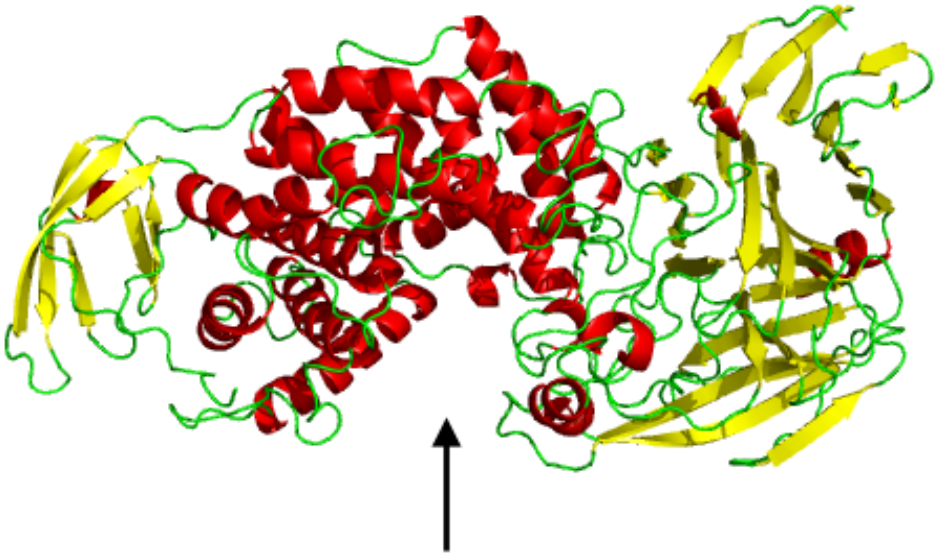
C. *Proteus vulgaris* ChonABCI (PDB accession code 1HN0, (Huang *et al.* 2003b)). Black arrow denotes the substrate-binding cleft.



D. *Streptococcus pneumoniae* Hyal (PDB accession code 1EGU, (Li *et al.* 2000)). Black arrow denotes the substrate-binding cleft.



C. *Streptococcus agalactiae* Hyal (PDB accession code 1F1S, (Li & Jedrzejas 2001)). Black arrow denotes the substrate-binding cleft.



1.3.4 GAG Lyase Substrate Binding Characteristics

The substrate binding site of PL8 proteins is located within a surface exposed cleft (Fig. 1-3). The substrate-binding area shows positive electrostatic potential displaying charge complementarity with the negatively charged GAG substrates. The interactions found in the binding clefts are typical for carbohydrate-binding proteins (Hileman *et al.* 1998) and are dominated by both polar and positively charged residues, particularly arising from arginine / asparagine side-chains, oriented by hydrogen bonding to glutamate side-chains. Aromatic stacking interactions between tryptophan residues and the hydrophobic faces of GalNAc rings are prominent features. In *A. aurescens* ChonAC, Trp126 stacks against the "-1" GalNAc and Trp465 stacks against the "+2" GalNAc from the opposite face of the oligosaccharide chain (Fig. 1-4). Similar aromatic stacking has been identified in lectins and is typical of enzymes acting on neutral polysaccharides (Vyas *et al.* 1991; Nagy *et al.* 1998; Ponyi *et al.* 2000). Aromatic stacking has also been observed in the interaction between hyaluronate lyase and hyaluronan disaccharide (Ponnuraj & Jedrzejas 2000).

Substrate binding to *A. aurescens* ChonAC is associated with a deformation of the GlcA sugar ring at the "+1" site from a chair to a distorted boat conformation (Lunin *et al.* 2004). Such a change to a higher-energy conformation of the sugar ring is not unusual, and has been observed in other carbohydrate-processing enzymes (Strynadka & James 1991; Kuroki *et al.* 1993; Strynadka & James 1996). Distortion of the ring brings the hydroxyl groups into equatorial positions, and the C5 carboxylate group to pseudo-axial position, coplanar with the amide group of an asparagine side chain. This asparagine side chain in turn, forms a strong hydrogen bond to the acidic group, which requires protonation and results in charge neutralization. This higher energy conformation of the uronic acid ring was argued to serve in lowering the energy barrier for bond cleavage.

Figure 1-4. The binding site of *Arthrobacter aurescens* ChonAC.
CS tetrasaccharide is shown in yellow/blue/red sticks, amino acids involved in binding of the tetrasaccharide are shown in white/blue/red sticks, the relevant H-bonds are shown in cyan dashed lines.

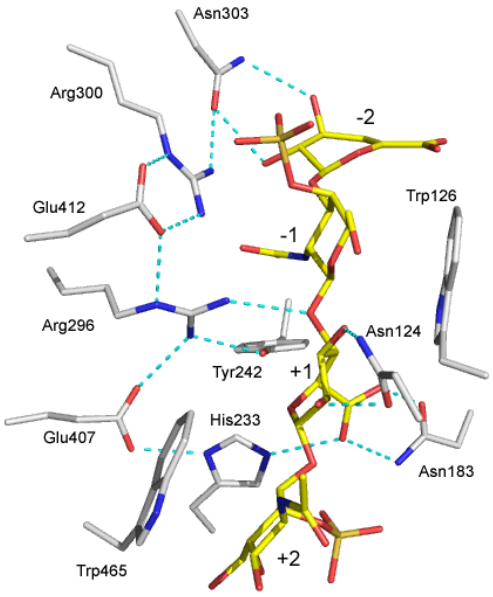
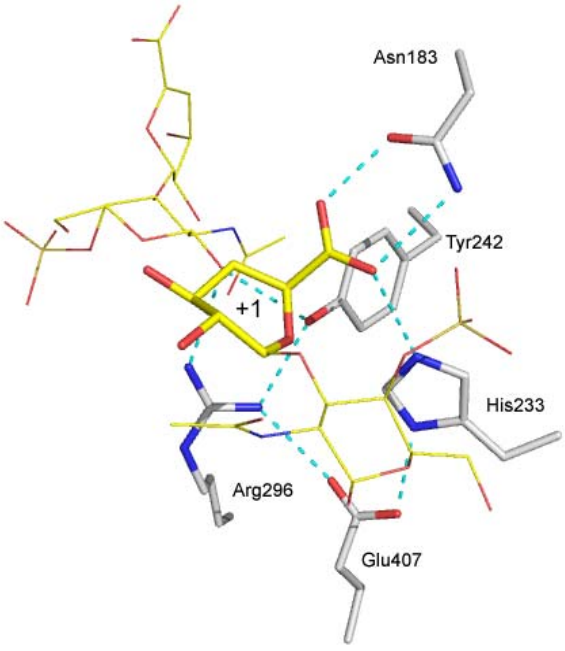


Figure 1-5. High-energy sugar conformation in *A. aurescens* ChonAC.
The CS tetrasaccharide is shown in yellow thick sticks and thin lines, Asn183 (charge neutralization) and the catalytic tetrad residues are shown in white sticks, relevant H-bonds are shown in cyan dashed lines, the glucoronic acid at subsite "+1" adopts a distorted boat conformation.



The substrate-binding site of ChonB (Huang *et al.* 1999) was located in a "U" shaped depression on the convex surface of the enzyme. The enzyme binds DS essentially through ionic interactions involving basic residues (arginines and lysines), and specific hydrogen bonds to the hydroxyl groups of the polysaccharide (Fig. 1-6). The specificity towards DS is conferred by electrostatic interactions between two arginine side chains, a glutamine and a lysine binding the sulfate groups at the C4 position of the GalNAc residues at subsite "+1", and the formation of a ternary complex including the IdoA carboxylate oxygen atom, a Ca^{2+} ion, and three amino acids, namely an asparagine and two glutamates (Fig. 1-7). A distorted boat conformation was argued from a model of ChonB binding its substrate (Michel *et al.* 2004), as was also observed in the complex structure of *A. aurescens* ChonAC-substrate complex (Lunin *et al.* 2004).

Figure 1-6. *Pedobacter heparinus* ChonB active site.
The binding site around the "+" subsites; White sticks represent amino acids that bind the disaccharide occupying the "+" subsites, the Ca^{+2} ion is shown as an orange sphere, DS di-saccharide occupying "–" subsites is shown in pale brown, and the relevant H-bonds are shown as cyan dashed lines.

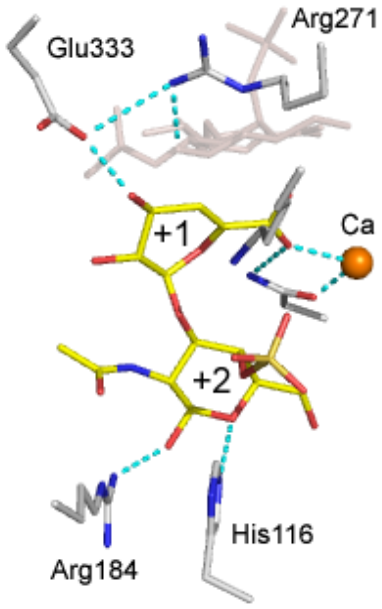
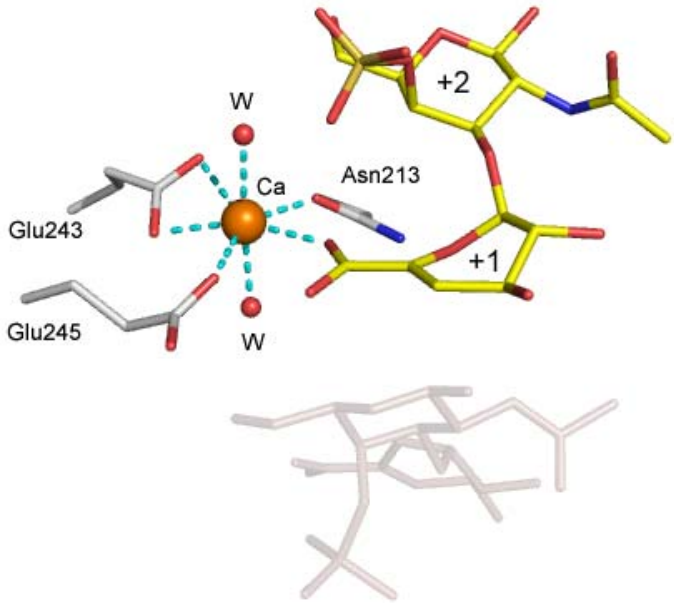


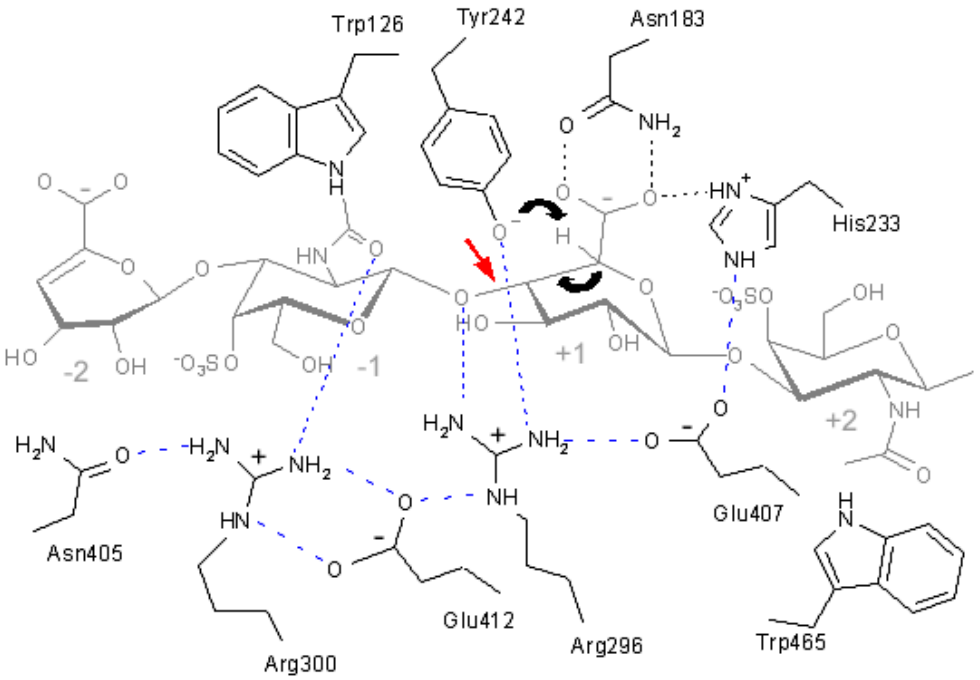
Figure 1-7. *Pedobacter heparinus* ChonB Ca^{2+} coordination
The Ca^{+2} ion (shown as an orange sphere) coordinated by Asn213, Glu243, Glu245 (represented by white/blue/red sticks), 2 water molecules (represented by red spheres), and the carboxylate oxygen of the iduronic acid in subsite "+1". The dissacharide occuping the "+" sites is shown by yellow/blue/red sticks, whereas the disaccharide occuping the "–" sites is shown in pale brown sticks.



1.3.5 The Catalytic Mechanism of ChonAC

Four residues, a histidine, a tyrosine, an arginine and a glutamate were shown to form a catalytic tetrad. The glutamate residue of this tetrad is bridging the histidine and arginine side-chains and is therefore important in maintaining the three-dimensional arrangement of the active site residues. The arginine side-chain was shown to be primarily involved in binding the glycosidic oxygen (forms the glycosidic bond to be cleaved), and is important for proper positioning of the substrate in the catalytic center. The role of the general base removing the proton bound to C5 of the glucuronic acid was assigned to a deprotonated tyrosine side chain. The structural (Lunin *et al.* 2004) and chemical (Rye *et al.* 2006) data argued for the same tyrosine to also protonate the leaving group, acting as a general acid. Charge neutralization of the acidic group was shown to take place by hydrogen bonds between the histidine side-chain together with another side chain, an asparagine. Figure 1-8 summarizes schematically the proposed catalytic mechanism for ChonAC based on the enzyme-substrate complex structure of *A. aurescens* ChonAC.

Figure 1-8. The putative catalytic mechanism of ChonAC.
The substrate is portrayed in gray whereas active / binding site residues are in black.
Step 1: The negative charge of the GlcA is neutralized by H-bonding to His233 and Asn183.
Step 2: Proton abstraction preformed by deprotonated Y242 (denoted by a black arrow).
Step 3: the double bond in position C4-C5 of the glucoronic acid is formed (denoted by a black arrow) and the glycosidic bond is eliminated (denoted by a red arrow).
Step 4: the products are released by protonation of the leaving sugar by Y242. Blue dashed lines denote hydrogen bonds identified in the structure of the active site.

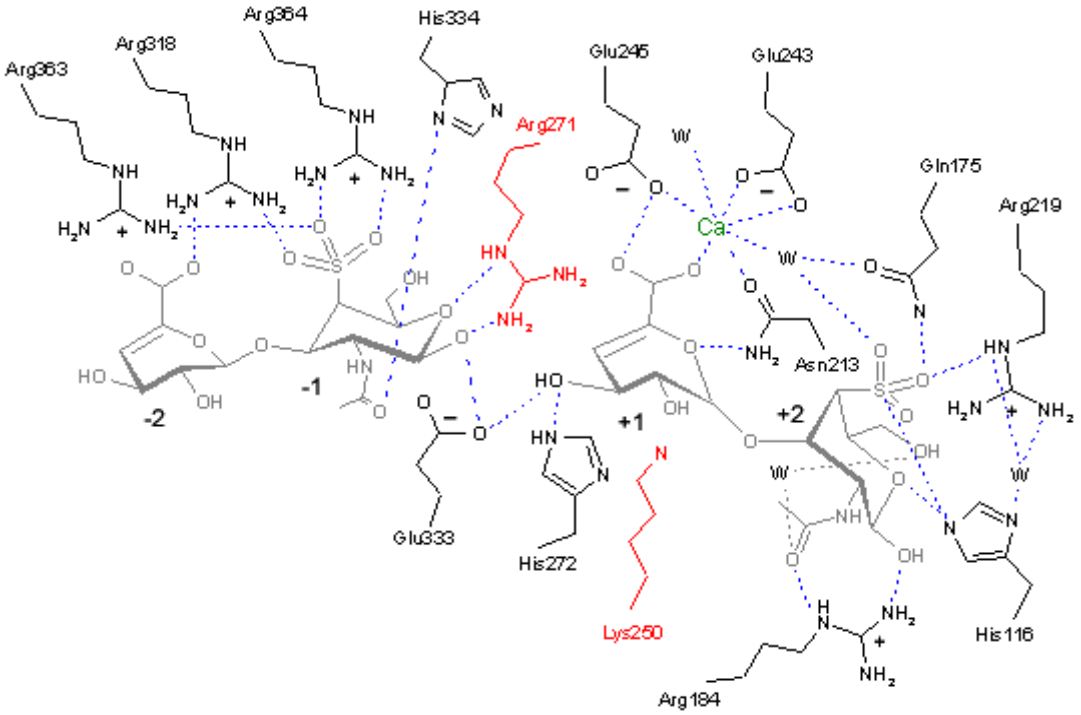


1.3.6 The Catalytic Mechanism of Chondroitin Lyase B

A calcium ion was identified in the structure of the enzyme co-crystallized with DS, but not in the native structure or that containing bound disaccharide product, PDB codes 1DBO, 1OFM respectively (Huang *et al.* 1999; Michel *et al.* 2004)). The calcium ion was chelated by three amino acids – asparagine and two glutamates, two water molecules and by the iduronic acid at the "+1" subsite. Enzyme kinetics combined with site-directed mutagenesis of these residues revealed an absolute requirement for calcium during catalysis (Michel *et al.* 2004). It was argued that calcium plays a pivotal role in binding of DS and in its degradation through the stabilization of the negative charge on the C5 carboxylate moiety. A lysine residue was proposed to serve as the base in the

abstraction of the C5 proton from the iduronate, and an arginine side chain was assigned as the Brønsted acid protonating the O^- leaving group. Figure 1-9 summarizes schematically the active site of *P. heparinus* ChonB.

Figure 1-9. Schematic representation of *P. heparinus* ChonB active site.
The calcium ion (shown in green) binds the carboxylic group of the L-iduronic acid at subsite "+1", lowering the p*Ka* of the H5-proton. Lys250 (shown in red) abstracts the proton, whereas the glycosidic bond is protonated by Arg271 (shown in red), followed by the elimination of the glycosidic bond. Relevant hydrogen bonds are shown in blue.



1.3.7 Endolytic vs. Exolytic Glycosaminoglycan Lyase

The mode of action of *P. heparinus* ChonAC has been established as random endolytic, yielding a mixture of disaccharides, tetrasaccharides and longer oligosaccharides (Hiyama & Okada 1976; Jandik *et al.* 1994; Gu *et al.* 1995). Alanine replacement of a specific arginine residue involved in binding of the GalNAc residue in subsite "–1", namely binding of the *N*-acetyl or *O*-sulfo moieties of the galactosamine changed the enzyme's mode of action to exolytic (Dong *et al.* 2001). It was suggested that the oligosaccharides binds initially within the groove and slide through the tunnel, positioning at the "–2" and "–1" subsites of the enzyme active site. These subsites have

been suggested to represent a high affinity substrate recognition domain, while the "+1" and "+2" subsites represent a product release area (Schmidt *et al.* 1999). Furthermore, the researchers postulated that after cleavage, the oligosaccharides can become trapped in the high affinity "–1" subsite and will not overcome the energy barrier necessary to slide through the tunnel to the ''+'' subsites, preventing subsequent cleavage within the same oligosaccharide chain (Huang *et al.* 2001). Thus, removing the arginine from this site effectively reduced the high affinity binding at subsite "–1", making it possible for the oligosaccharide chain bound at the "–1" subsite after cleavage to slide through the tunnel to the ''+'' subsites, leading to a subsequent, processive, step-wise, exolytic depolymerization within the same oligosaccharide chain.

A. aurescens ChonAC on the other hand, presents an exolytic mode of action, releasing disaccharide products from its GAG substrate (Jandik *et al.* 1994). Structural comparison between *P. heparinus* and *A. aurescens* ChonAC provided a rationale for the observed differences (Lunin *et al.* 2004). *A. aurescens* ChonAC contains two insertions in the N-terminal domain relative to *P. heparinus* ChonAC, one 15 residues long (Arg23–Ser38) and the other 25 residues long (Thr343–Gly366). In structural terms, the two segments form α-helices and connecting loops that fold close to the N- and C-termini of the domain. Together, these segments close off a large part of the surface-exposed cleft harbouring the active site. These segments form a wall that converts the open cleft into a deep cavity and precludes binding of an extended oligosaccharide. The size of the cavity restricts the binding of the carbohydrate to the non-reducing end and was estimated to accommodate two to three monosaccharides, correlating well with the exolytic mode of action presented by *A. aurescens* ChonAC.

Interestingly, in ChonB, an alanine replacement of a specific arginine, Arg364 (Pojasek *et al.* 2002) resulted in an altered product profile after exhaustive digestion of DS, suggesting a role for this residue in defining the mode of action of this enzyme. This arginine interacts with both the 4-*O*-sulfate of the GalNAc4S and the carboxyl group of the IdoA at the non-reducing end. Since 4-*O*-sulfation of GalNAc and IdoA both represent hallmark modifications of DS, this arginine was argued to be a main determinant involved in the substrate specificity of the enzyme. Furthermore, the R234A mutant yielded a complete loss of activity in the real-time kinetic assay and an exhaustive

digestion of DS revealed tetrasaccharides, hexasaccharides, and octasaccharide as end products of the reaction as apposed to the WT enzyme yielding only disaccharides. The oligosaccharide products suggest that R234A-ChonB acts in an endolytic manner and lost its ability to recognize and cleave DS regions containing 4,6 di-sulfated GalNAc. Complete characterization of the mode of action of the WT ChonB and the R234A mutant enzymes is necessary to clarify these issues.

Hamai *et al.* showed that *Proteus vulgaris* produces two isozymes of ChonABC, an endoeliminase depolymerizing CS and DS affording disaccharides and tetrasaccharides as end products, and an exoeliminase preferentially acting on CS tetrasaccharides and hexasaccharides to yield the respective disaccharides (Hamai *et al.* 1997). These enzymes were designated ChonABCI and II, respectively.

Using Mass-Spectroscopy in tandem with capillary electrophoresis, researchers from Sasisekharan's laboratory have shown that *P. heparinus* HepI displays a predominantly processive exolytic mode of action, while *P. heparinus* HepII acts in a non-random endolytic manner (Ernst *et al.* 1998; Rhomberg *et al.* 1998a; Rhomberg *et al.* 1998b).

1.4 Thesis Rationale and Objectives

Since they were first described in the literature, bacterial GAG lyases have found uses in the laboratory and in the clinic (Linhardt *et al.* 1986; Pervin *et al.* 1994; Venkataraman *et al.* 1999; Keiser *et al.* 2001). Nevertheless, the identity of the catalytic residues of several of these enzymes and their roles during catalysis are still not understood in details. This study aimed to achieve several goals. The first goal was to obtain a structural description of HepII as the first structure of a heparin/HS-degrading enzyme. Since HepII has a unique sequence without any close homologs, it was likely that its structure may represent a new fold in the GAG lyase family.

The second goal was to investigate HepII substrate specificity with respect to the sulfate modifications and to understand how the enzyme is able to process both epimers of the uronic acid. As well, I sought a structural explanation for the enzyme's endolytic mode of action especially in comparison it to the recently determined structure of the endolytic HepI (Garon M.L, Shaya D., and Cygler M. unpublished results).

The third goal was to propose a chemical scheme describing the depolymerization of HS and heparin by HepII, and to establish roles of the various essential residues during catalysis.

Upon obtaining the structures of native HepII (nHepII) and recombinant HepII (rHepII) complexed with its substrates, I initiated a structure-function study of ChonABC. This study aimed to shed new light on the catalytic mechanism of ChonABC by crystallizing and solving the structure of a homologous enzyme from the gut mutualist *B. thetaiotaomicron*. In particular, I wanted to obtain an explanation for its degradation mechanism with regards to the IdoA-containing DS, an endeavor that has been ongoing in our lab since the first structure of *P. vulgaris* ChonABCI was solved (Huang *et al.* 2003). To extend the structural study, I undertook a site-directed mutagenesis studies aiming to identify the residues involved in DS degradation.

Due to its new clinical importance, an ongoing collaboration between the laboratories of Prof. Linhardt R. J. and ours was established to improve the methods of clinical usage and delivery of ChonABC. Since the previous biochemical characterization of BactnABCII was very preliminary and incomplete (Linn *et al.* 1983) we have undertaken a more through *in-vitro* biochemical characterization of this enzyme.

1.5 Experimental Procedures

1.5.1 Protein X-ray Crystallography

Determination of atomic the structure of molecules goes back to the experiments by Max von Laue, who, in 1912, only ~17 years after the discovery of the X-ray phenomena by Roentgen, showed that crystals diffract X-rays and thus demonstrated that the distances between atoms in the crystal are of the same order of magnitude as the wavelength of the X-rays (von Laue 1912). The theory of diffraction showed that the scattered amplitudes of the X-rays measured from the diffraction pattern are related by a Fourier transform to the electron density distribution within the crystal (Blundell & Johnson 1976), and suggested that the analysis of the diffraction effect can provide information about the distribution of atoms within the crystal, revealing their spatial configuration in the molecules scattering the X-rays. Experimentally, however, only the intensities of the scattered rays (related to the square-root of the scattered waves) could

be measured but not their phases. This analysis was initially applied to crystals of inorganic molecules and led to the development of methods for extracting the experimental phases (Patterson function, direct methods (Stout & Jensen 1989)) that allowed extending the application of X-ray crystallography to the determination of structures of small organic compounds. The first breakthrough in the application of X-ray diffraction to proteins was pioneered by Max Perutz and coworkers, who demonstrated that binding of a few heavy atoms with large number of electrons to the protein (within the crystal) could result in measurable changes in the diffraction pattern. These changes allowed the derivation of the phases of the diffracting rays, which in turn led to the determination of the first protein structures, myoglobin and haemoglobin in the early 1960's (KENDREW *et al.* 1958; Perutz 1960). Over the last 50 years there has been a tremendous progress in development of instrumentation for recording X-ray diffraction data and in the methodologies for extracting the missing phases from the experimental data, which has led to over 50,000 protein structures determined to date, including the impressive ~2.5 MDa RNA-protein complex, the ribosome (http://www.rcsb.org/pdb/home/home.do, (Ban *et al.* 2000; Schluenzen *et al.* 2000).

The process of protein structure determination by X-ray diffraction contains many intermediate steps. The experimental steps involve preparation of a highly-pure protein samples, their crystallization, diffraction data collection, and deciphering of the atomic structure from the diffraction pattern by computational methods (Drenth 1994). In this chapter I will provide a brief overview of the technical steps involved in protein X-ray crystallography, starting from obtaining an expressing clone, to the final protein model deposited in the Protein Data Bank (PDB, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ, http://www.rcsb.org/). Special attention will be dedicated to the experimental steps that proved to be a bottleneck during my thesis research.

1.5.1.1 Protein Sample Preparation

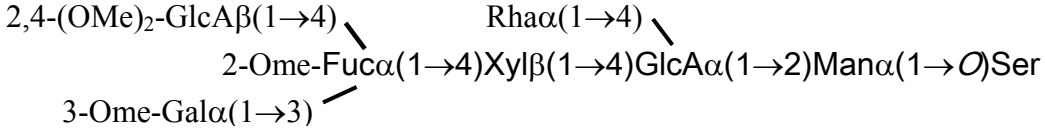
Upon obtaining a soluble protein, the main effort toward structure determination is to produce well-ordered protein crystals that yield a diffraction pattern extending to high resolution when irradiated with X-rays. The high quality of a protein sample,

meaning purity and homogeneity, are judged by a combination of electrophoresis, chromatography, and dynamic light scattering (DLS). Chemical and conformational homogeneity are essential for obtaining well diffracting protein crystals. Protein purity issues are dealt with by standard chromatographic techniques (Deutscher 1990) until purity of >95% is reached. The sample homogeneity is a crucial determinant for the crystallization process, and can be evaluated by DLS or chromatography techniques. A range of biophysical factors, including temperature, the type of buffer, pH, ionic environment, as well as factors related to the protein itself, such as protein stability and, in case of glycosylated proteins, the size and uniformity of glycosylation, influence the protein sample's homogeneity. Buffer, pH and ionic environment issues can be optimized by monitoring the DLS properties of the protein samples in a range of buffers, pHs and salts, known to promote protein homogeneity (Carter, Jr. & Sweet 1997). Protein degradation can be prevented or decreased by the addition of protease inhibitors during purification or changing the expression strain. A more drastic measure that can be employed for proteins that prove very sensitive to proteolysis is to mutate the protease-sensitive site on the protein to eliminate its degradation. Protein glycosylation presents a difficulty for crystallization, due to the usual heterogeneity in glycosylation and the significant flexibility often displayed by the glycan chains (Wormald *et al.* 2002). Several techniques have been used to resolve those issues; among them are mutating the glycosylation sites to prevent glycosylation altogether, enzymatic deglycosylation of the protein following purification, and using expression systems that do not glycosylate proteins, the most common example being *Escherichia coli*.

In the context of this thesis, I encountered protein homogeneity issues, with HepII expressed in *P. heparinus*, which led to fragile protein crystals and an anisotropic X-ray diffraction pattern. This sample heterogeneity was related to heterogeneous protein glycosylation, as judged by mass spectroscopy (Shaya, D., and Cygler, M., unpublished results). Studies conducted with HepI and other GAG lyases have shown that *P. heparinus* glycosylates several GAG lyases with a distinct 1156 Da, branched, heptasaccharide glycan consisting of GlcA, Man, Rha, Xyl, 2-OMe-Fuc, 3-OMe-Gal and 2,4-(OMe)2-GlcA (Fig. 1-10 (Huang L. *et al.* 1995)). I have resolved the problem by limiting the growth of the *P. heparinus* cells using shorter culture times and minimal

medium components. The total culture times and temperatures for both HepII and ChonABCII displayed a strong influence on the resulting protein sample and therefore crystal quality and so have been optimized to yield better-diffracting crystals.

Figure 1-10. Glycosylation of *P. heparinus* GAG lyases (Huang L. *et al.* 1995)



1.5.1.2 Protein Crystallization

Upon obtaining adequate protein sample one needs to determine the appropriate conditions for protein crystal nucleation and growth. A large number of factors, including protein concentration, buffer type, pH, precipitant identity, salts, and temperature influence the crystallization process. Since these variables cover a multidimensional space too large for an exhaustive search, sparse matrix sampling methods have been developed to explore this space in order to rapidly and efficiently identify chemical conditions that promote protein crystal nucleation and growth (Jancarik & Kim 1991; Huang L. *et al.* 1995). The experiments are set-up using several parallel techniques, each one utilizing different geometric set-ups. The most common techniques employed in my research are the vapor diffusion in the hanging drop or sitting drop format, and batch crystallization (McPherson 2004).

The initially established crystallization-promoting conditions frequently have to be optimized varying the pH of the solution, concentrations of the crystallization solution components, and by testing the effect of various additives (metal ions and other compounds used at very low concentrations) in order to obtain bigger and better diffracting crystals. Often manual manipulation techniques, such as micro-seeding or macro-seeding are employed (McPherson 2004). In the context of this thesis, I used micro-seeding by crushing small needle crystals of HepII to prepare nucleation seeds,

diluting them, and seeding them into a fresh crystallization drop in order to obtain larger and better-ordered crystals.

1.5.1.3 Data Collection

Measurement of diffraction data from one or more single crystals constitutes the last experimental step. This involves exposing the crystal to X-ray radiation and recording the diffraction images using an appropriate detector. In order to minimize the extent of radiation damage to protein molecules within the crystals, the crystals are frozen in a cold N_2 stream (100 °K). Prior to freezing, the crystals are mixed with a cryoprotectant solution that permeates the crystal, preventing the formation of ice inside the crystal. In many instances, such as the case of HepII, the cryoprotectant serves also as a crystal-stabilizing agent and not merely as a preventive measure against ice formation. In this context, the major breakthrough with nHepII crystals was the identification of the stabilizing effect of high concentrations of sodium formate on HepII crystals when incorporated in the cryoprotectant (see chapter 2). Once an adequate cryoprotection solution is identified, the cryoprotected crystals are mounted in nylon loops of appropriate size and placed in the direct path of the X-ray beam. The crystals are than slowly rotated and exposed to the X-ray radiation for short periods of time, allowing systematic recording of the diffraction data for different crystal orientations (data frames).

1.5.1.4 Data Processing

Recorded diffraction data frames contain information about the directions and intensities of the scattered X-rays. This information is extracted from the data frames (images) during the data processing step, which is performed *in silico*. This process involves assigning Miller (*hkl*) indices to each diffraction spot (reflection), integration of the pixels corresponding to each reflection and converting the intensities to structure factor amplitudes by applying corrections depending on the type of instrumental setup and design (Otwinowski & Minor 1997). Further data processing involves the determination of the crystal parameters and its orientation with respect to the X-ray beam

and the detector, scaling the data recorded on different frames and merging multiple measurements and symmetry related reflections into a unique dataset (Jones *et al.* 1996).

1.5.1.5 The Phase Problem

Once the experimental measurements are converted into a dataset, the next step involves reconstructing the real space lattice, that is, the electron density of the protein in the crystal, from the reciprocal lattice, the X-ray diffraction pattern. This reconstruction is done through the Fourier transform, which requires the structure factor amplitudes and phases for all reflections. As explained above, the X-ray experiment records the intensity of the reflections, converted to structure factor amplitudes, but does not measure directly the phases. This information is essential for calculating the electron density distribution. Over the years, various methods and algorithms have been devised to estimate the experimental phases. These include Multiple Isomorphous Replacement (MIR), Molecular Replacement (MR), Multiple-wavelength Anomalous Dispersion (MAD) and their various combinations (Blundell & Johnson 1976). Improvement in computational algorithms such as solvent flattening and non-crystallographic symmetry averaging makes Single Anomalous Dispersion (SAD) the most common method to solve new structures and MR to solve structures where the structures of sequence homologs (>30% sequence identity) already exist (Carter, Jr. & Sweet 1997).

The MIR method involves using heavy-atom (electron dense) derivatization of the protein crystal such that the derivatized crystal is isomorphous (contains minimal changes in the unit cell and the orientation of the molecules in the crystal) to the native, underivatized crystal. The presence of heavy atoms introduces measurable intensity changes, which are used to deduce the positions of the heavy atoms. The native protein phases are than estimated from the comparison of the intensities of the native and derivative crystal taking into account the known contribution of the heavy atoms (Blundell & Johnson 1976).

The interaction between radiation and electrons far from the absorption edge of the atom has a predominantly elastic component, resulting in scattering. However, when the wavelength (energy) of the X-rays approaches the absorption edge of a specific atom type, the contribution from inelastic scattering becomes measurable. In simple terms, this

can be interpreted as an excitation of an electron from a low-energy, inner shell to a higher orbital followed by a return to the ground state, which results in photon emission. This event, called anomalous scattering, yields a 90^{o} phase change of the emitted photon. Anomalous scattering is composed of a dispersive term and an absorption term. The differences recorded in the diffraction pattern due to the anomalous scattering are used to estimate the location and phases of the anomalous scatterers in MAD and SAD methods (Drenth 1994). These methods rely on the introduction of anomalous scatterers into the protein crystals. The most commonly used anomalous scatterer is selenium, which is introduced into the recombinant protein in the form of L-selenomethionine during protein expression. Tunable wavelength (energy) synchrotron beamlines are used to irradiate protein crystals at a wavelength that maximizes the absorption and dispersive effects. The SAD method utilizes one wavelength at the absorption peak, maximizing the anomalous signal, whereas the MAD technique uses three wavelengths, the anomalous absorption peak where the anomalous differences are maximum, the inflection point, where the dispersive effect is maximum, and at a remote energy, where anomalous signal is minimal. In the context of this research I used selenomethionine-labeled rHepII crystals irradiated at the Se K(III) absorption maximum edge (SAD regime), and extracted the phase information from this single dataset to determine the structure of rHepII and BactnABCII.

1.5.1.6 Model Building and Refinement

The phases derived as described above are combined with the structure amplitudes and used to calculate the electron density distribution by applying an inverse Fourier transform. This electron density is displayed on a graphics workstation utilizing specialized graphical application software that allows an initial protein model to be built into this density. This initial model is then improved (refined) by computational approaches, such as the least-squares (Drenth 1994), maximum likelihood (Murshudov *et al.* 1997), or simulated annealing (Brunger *et al.* 1998) methods to minimize the differences between the experimental structure factors amplitudes and the structure factors amplitudes calculated from the atomic model (Drenth 1994). The progress of refinement is monitored by the agreement between the observed (F_{obs}) and calculated

(F_{calc}) structure factors amplitudes, evaluated by the *R*-factor calculated according to the formula:

$$R_{work} = (\Sigma||F_{obs}(h\ k\ l)|- |F_{calculated}(h\ k\ l)||)/\ \Sigma|F_{obs}(h\ k\ l)|$$

where the *h k l* summation is over all of the reciprocal lattice points that were experimentally measured. To monitor model bias, a small percentage of reflections are set aside and they are not used for the refinement. These reflections are used to calculate a parameter, designated as $R_{free,}$ in the same manner as the R-factor is calculated. The value of R_{free} provides a measure of the agreement of measured and calculated structure factor amplitudes for an unbiased set of reflections (Brunger *et al.* 1987). Ideally during refinement, the diference between R_{work} and R_{free} will attain a minimum.

1.5.1.7 Resolution in X-ray Crystallography

The level of detail of a graphical image is defined by its resolution. The term 'resolution' in X-ray crystallography is defined by Bragg's law:

$$d_{min}=\lambda/(2 \cdot \sin\theta_{max})$$

where θ_{max} is the maximum angle between the incident beam and the direction of reflections furthest away from the direct beam that still have measurable intensities, λ is the radiation wavelength (Bragg 1913). d_{min}, or resolution, corresponds to the minimal spacing between sets of parallel planes passing through the crystal lattice points that scatter X-rays that yield measurable signal. The higher is the resolution, the is smaller the d_{min} value and the more details that can be discerned in the electron density. For example, at a resolution of 3 Å protein secondary structural elements are well defined but the orientation of sidechains may not be reliably resolved. At 2 Å resolution the orientations of the sidechains are well resolved but individual atoms locations are determined based on chemical bond-length constrains. At a resolution of 1 Å and better, the individual atoms are well resolved in the electron density map, and their positions known with great accuracy. Many of the hydrogen atoms can also be discerned in the electron density map at this resolution.

1.5.2 Kinetic Analysis

Enzyme kinetics denotes the study of chemical reactions that are catalyzed by enzymes, focusing on the measurement and analysis of the reaction rates. The following text will review generally the techniques used through this desertation.

Depolymerization of heparin and HS by HepII and CS-A/C and DS by ChonABC follows the Michaelis-Menten kinetic equations describing enzyme rates of reactions. The general scheme is as follows:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_2} E + P$$

the Michaelis constant $K_m = (k_{-1} + k_2)/k_1$

where: E – enzyme, S – substrate, ES – enzyme-substrate complex formed, P – product. The reaction rate, V, is defined, assuming the reaction is at steady state, square brackets denote concentration, as:

$$V = d[P]/dt = V_{max} ([S]/K_m+[S])$$

where V_{max} is the maximal reaction rate at saturating substrate concentrations. Very often, the equation is represented by its reciprocal Hanes-Woolf plot which follows the linear equation:

$$[S]/V = [S]/ V_{max} + K_m/ V_{max}$$

The Hanes-Woolf plot yields a straight line of slope $1/V_{max}$, a y-intercept of K_m/V_{max} and an x-intercept of $-K_m$. In the context of this thesis we used linear regression analysis with the Hanes-Woolf formalism to calculate the kinetic parameters K_M and K_{cat} for the various mutants of ChonABC analyzed in Chapters 5 and 6.

Chapter 2 Crystallization and Preliminary X-ray Analysis of HeparinaseII from *Pedobacter Heparinus*

Preface

The contents presented in the following chapter have been published as presented:

Shaya, D., Li, Y., and Cygler, M. (2004) Crystallization and preliminary X-ray analysis of heparinase II from *Pedobacter heparinus*. *Acta Crystallogr. D.* **60:** 1644-1646.

Contribution of Authors:

I preformed all the experiments, analyzed the data and wrote the draft to the manuscript. Y.L. instructed me with the crystallization steps. M.C. supervised the project and provided comments and suggestions thought the course of the work. The final manuscript for publication was prepared by M.C. and myself.

2.1 Rationale

The manuscript represents the initial steps toward the structure elucidation of HepII. In the context of this dissertation, the subsequent chapters 3 and 4, describing the structural studies of HepII, relate to this manuscript. This chapter underscores the importance of protein purification in the context of X-ray crystallography. Well diffracting crystals and high quality X-ray diffraction data depend on the high purity of protein preparation and its homogeneity. In the case of HepII, the initial crystals diffracted to low resolution. I started my research by modifying the growth media, purification scheme, and crystallization protocols for native HepII (nHepII) overexpressed and purified from its original host, the soil bacterium *Pedobacter heparinus*, with the aim of obtaining well diffracting protein crystals. This chapter describes the results of these efforts.

2.2 Synopsis

We report here the crystallization and preliminary X-ray analysis of heparinase II from *Pedobacter heparinus.* Heparinases from *Pedobacter heparinus* are widely used in medical and analytical applications, however their three-dimensional structures and detailed catalytic mechanisms have yet to be determined.

2.3 Abstract

Heparinase II from *Pedobacter heparinus* (formerly *Flavobacterium heparinum*) which acts on both heparin and heparan sulfate is one of several glycosaminoglycan-degrading enzymes produced by this organism. This enzyme, with a molecular weight of 84 kDa, utilizes a lytic mechanism to cleave the α(1-4) glycosidic bond between hexosamine (D-glucosamine) and L-iduronic or D-glucuronic acid, resulting in a product with an unsaturated sugar ring on the non-reducing end. The enzyme was crystallized by hanging drop vapour diffusion method. The crystals belong to orthorhombic space group $P2_12_12_1$ and diffract to 2 Å resolutions. There are two molecules in the asymmetric unit consistent with our finding that recombinant heparinase II functions as a dimer in solution.

2.4 Introduction

Glycosaminoglycans (GAGs) are negatively charged linear polysaccharides usually composed of repeating units of uronic acid and sulfated *N*-acetylglucosamine (GlcNAc) or *N*-acetyl-galactosamine (GalNAc) (Ernst *et al.* 1995). The GAG chains are generally attached at its reducing end to a protein core forming proteoglycans (PGs). PGs are important components of the vertebrate extracellular matrix and play critical roles in the formation of the skeleton (Iozzo 1998). The presence of GAGs is usually associated with higher animals but they are not unique to eukaryotes. Several specialized microorganisms also produce unsulfated forms of these polymers (DeAngelis 2002). In addition to a structural role, GAGs act as critical modulators of a number of biochemical signalling events, and play a role in cell growth and differentiation, adhesion, and tissue morphogenesis (Lander & Selleck 2000; Kawashima *et al.* 2002; Esko & Selleck 2002). Heparan sulfate (HS) glycosaminoglycan (Stringer & Gallagher 1997) emerged as key player in many different biological processes ranging from cancer (Linhardt *et al.* 1986; Sugahara & Kitagawa 2002), angiogenesis (Linhardt *et al.* 1992), anticoagulation (Petitou *et al.* 1999), viral and microbial pathogenesis (Shukla *et al.* 1999; Menozzi *et al.* 2002), and multiple aspects of development (Perrimon & Bernfield 2000).

Two different mechanisms are utilized by enzymes that degrade HS. Heparinases are polysaccharide lyases that employ an eliminative mechanism (Linhardt *et al.* 1986), while glucuronidases act through a hydrolytic mechanism (Myette *et al.* 2002). The soil bacterium *Pedobacter heparinus* (formerly *Flavobacterium heparinum*, (Steyn *et al.* 1998)) produces three different heparinases. They differ in size, charge properties, and substrate specificities (Nader *et al.* 1990; Lohse & Linhardt 1992) and display sequence similarities only to proteins from several other microorganisms (http://afmb.cnrs-mrs.fr/CAZY/PL.html). In this database heparinase I has been assigned to PL13 family, heparinase III to PL12 family while heparinase II is among not yet classified enzymes.

Heparinase II is unique in its ability to cleave both heparin and heparan sulfate–like regions of HS (Nader *et al.* 1990). It is a basic protein consisting of 772 residues. An N-terminal signal sequence of 25 residues is cleaved upon export to the periplasmic space to yield the mature protein with MW of 84,545 Da (Su *et al.* 1996). This enzyme

degrades heparin and heparan sulfate through β-elimination of the α(1-4) glycosidic bond between glucosamine and uronic acid (either glucoronic or iduronic acid) sugars, in an endolytic non-random manner (Rhomberg *et al.* 1998b). Although heparinase II exhibits stronger affinity towards heparin, its turnover rate of heparan sulfate is higher (Lohse & Linhardt 1992) and it prefers long polymer substrates to shorter oligosaccharide substrates. Sequence analysis showed that heparinase II posses a Cardin-Weintraub heparin binding consensus which reads ^{444}FFKRTIAH^{451} (Godavarti & Sasisekharan 1996). Extensive biochemical and site-directed mutagenesis revealed that His451 as well as Cys348 are essential for heparin degradation (Shriver *et al.* 1998a; Shriver *et al.* 1998b). It was also proposed that within the single substrate-binding pocket of heparinase II there are two proximate "active sites": one site contains residues positioned for the catalytic cleavage of heparin, the other contains residues that are positioned for the cleavage of heparan sulfate (Rhomberg *et al.* 1998b). We report here the crystallization of this enzyme.

2.5 Methods and Results

2.5.1 Protein Expression and Purification

Recombinant heparinase II (residues 26-772) was purified from *P. heparinus* cells (ATCC No. 13125) overexpressing the cloned gene product under the control of a heparin sensitive promoter (Blain *et al.* 2002). The bacteria were cultivated at 296 K with 225 RPM shaking for aeration, in defined medium FH (Su *et al.* 2001) supplemented with 10% crude grade heparin (Dongying Hi-tech Chemical Industry Co., Dongying City, Haochun West Road, Shandong province, China) and 0.1 mg/ml trimethoprim (Sigma–Aldrich chemie, Steinheim, Germany). The cells were harvested when the optical density of the culture reached A_{600} of 2.5 and lysed using a French-Press. 30 ml of cleared lysate (centrifugation for 30 minutes at 45,000 RPM, 4°C) was loaded on a water cooled (4°C) 25X800 mm SP Sepharose® Fast Flow column (Pharmacia Biotech, Uppsala, Sweden) pre-equilibrated with buffer A (10 mM Na-Phosphate pH 7.5). The column was washed with 50 volumes of buffer A and the protein was eluted in a linear gradient from 0 – 1 M NaCl in buffer A over 10 column volumes. The heparinase II containing fractions were loaded on a hydroxyapatite column (Bio-Gel® HTP, BioRad Laboratories, Richmond,

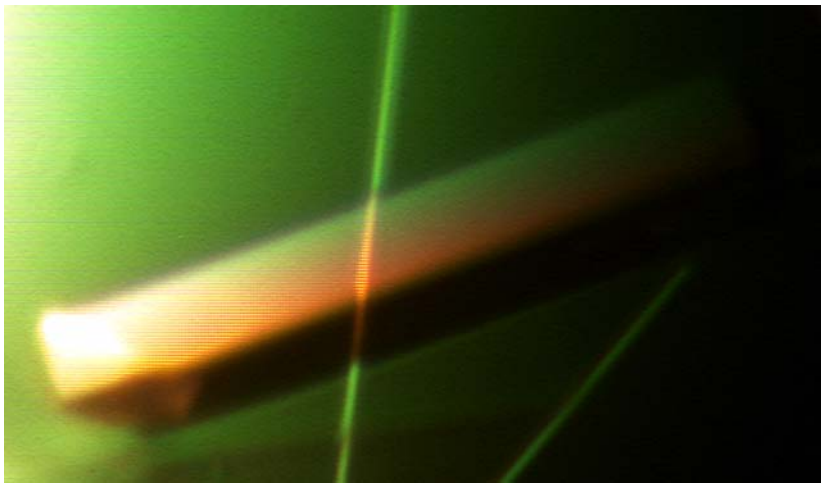
Canada) pre-equilibrated with buffer B (10 mM Na-Phosphate pH 6.0). The column was washed with 10 column volumes of buffer B with addition of 100 mM NaCl. The protein was eluted in 3 column volumes with a salt gradient from 400 mM to 700 mM NaCl in 15 mM Na-Phosphate pH 6.0. Purified heparinase II was concentrated by ultrafiltration using a Centriprep-50 and the buffer was exchanged to 10 mM Na-Phosphate pH 7.5, 100 mM NaCl, 20 mM Na-Formate, 5 mM dithiothreitol (DTT) using a 5 ml HiTrap desalting column (Amersham Biosciences AB, Uppsala, Sweden) to yield a protein with a concentration of 6.5 mg/ml. The protein showed a single band on a 12.5% (w/v) SDS-PAGE gel. Analysis of purified heparinase II by dynamic light scattering (DynaPro MSPRII, Protein Corporation, Piscataway, NJ, USA) showed that the protein was monodisperse with a molecular weight of 150 kDa indicating a dimer in solution. This was confirmed by size exclusion chromatography (Superdex 200 10/300Gl, Amersham Biosciences AB, Uppsala, Sweden) that yielded an apparent molecular weight of 162 kDa on a calibrated column. Enzyme activity was measured spectrophotometrically by monitoring at 232 nm the formation of unsaturated oligosaccharide products released from heparin (Fluka Chemie Neu-Ulm, Switzerland). Protein concentrations were determined by the method of Bradford (Bradford 1976) using bovine serum albumin as a standard.

2.5.2 Crystallization

Crystallization conditions for heparinase II were initially found by screening PEG/Ion screen™ (Hampton Research Laguna Niguel, CA, USA) using the hanging drop vapour diffusion method at 292 K. Small cubic-shaped crystals appeared from drops containing 0.5 µl of heparinase II (6.5 mg/ml, 10 mM Na-phosphate pH 7.5, 100 mM NaCl, 20 mM Na-formate) and 0.8 µl of reservoir solution (20% (w/v) polyethylene glycol (PEG) 3350, 200 mM sodium dihydrogen phosphate monohydrate, pH 4.5) and equilibrated over 60 µl of reservoir solution. Crystals appeared within 3 days to a week. They showed anisotropic diffraction extending to 3Å resolution in the best direction. Screening around the initial conditions led to crystals (obtained from 17% (w/v) PEG 3350, 200 mM Na-phosphate pH 5) that had a different morphology (rectangular plates) but the same unit cell dimensions. Well diffracting crystals were obtained by micro

seeding into drops consisting of 1 µl protein and 1 µl reservoir solution consisting of 16-17.5% (w/v) PEG 3350 and 200 mM Na-phosphate pH 5. Rectangular prism–shaped crystals measuring approximately 0.05 x 0.05 x 0.75 mm^3 grew overnight at 292 K (Fig. 2-1). These crystals belong to the orthorhombic space group $P2_12_12_1$ with cell dimensions a = 70.0, b = 119.3, c =200.7 Å and contain two molecules in the asymmetric unit. The Matthews coefficient Vm is 2.49 $Å^3$ Da^{-1} (Matthews 1968) and the estimated solvent content is 51%. Occasionally, we found in the same drops crystals that could be indexed with three times longer a axis, 210 x 119 x 200 Å. We have not investigated these crystals in detail at this time.

Figure 2-1. Crystal of hepII grown by the hanging-drop vapor-diffusion method.



2.5.3 Data Collection and Processing

Diffraction data were collected at the beamline X25, NSLS, Brookhaven National Laboratory, on an ADSC Q315 area detector. Crystals were transferred briefly from mother liquor to a cryoprotectant solution (22.5% (w/v) PEG 3350, 0.07 M Na-phosphate pH 5, 2.17 M Na-formate), mounted in a nylon loop (Hampton Research, Laguna Niguel, CA, USA) and placed directly in a cold nitrogen stream at 100 °K. The crystal-to-detector distance was set to 280 mm. Oscillation images of 1° were recorded with an exposure time of 7 seconds (Fig. 2-2). A native dataset was collected at a wavelength of 0.97940 Å to 2.0 Å resolution (Table 2-1). Images were processed with the *HKL2000* program

package (Otwinowski & Minor 1997) with statistics as shown in Table 2-1. The structure determination is in progress.

2.6 Acknowledgements

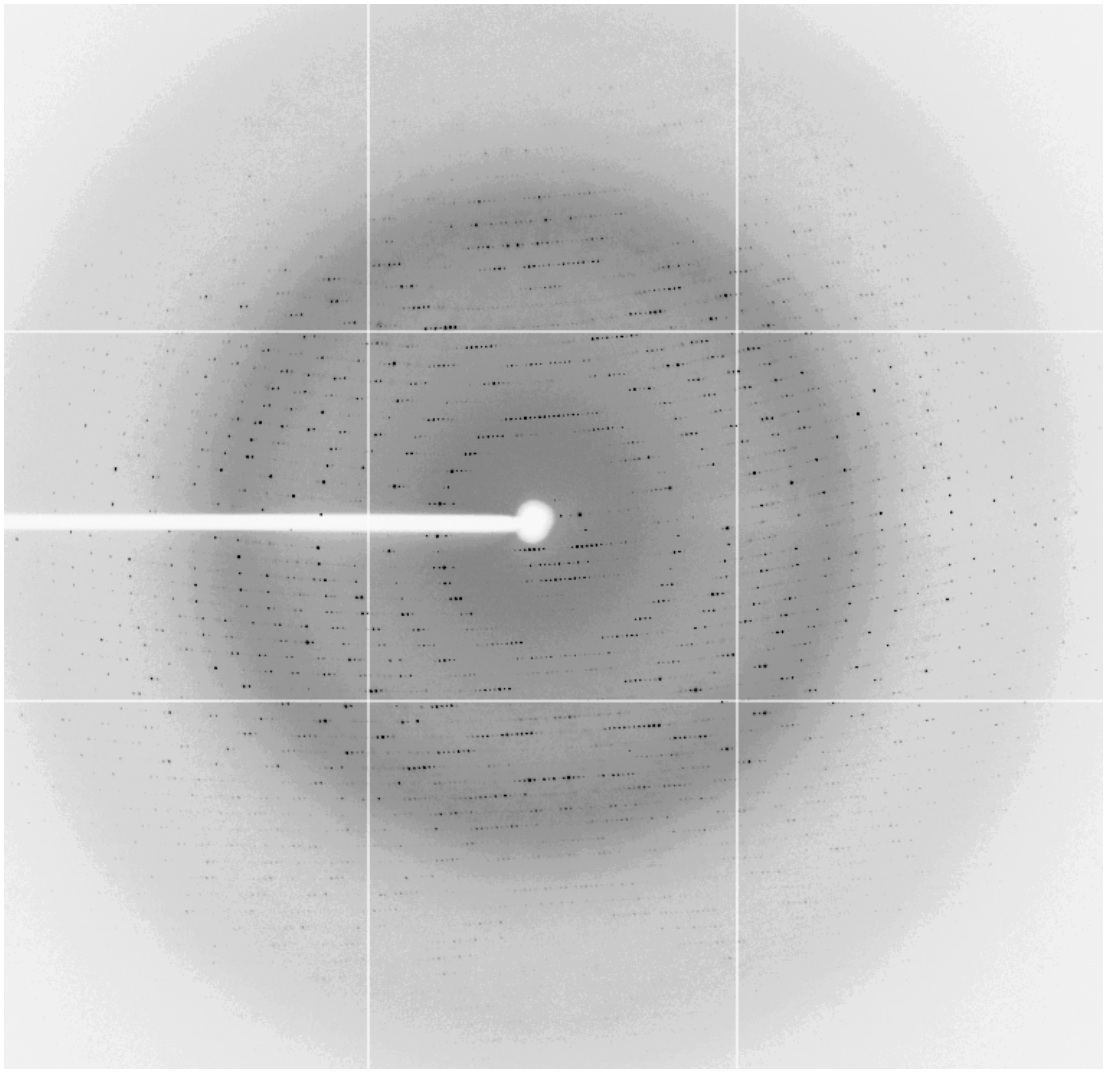
We would like to thank Dr. Joseph D. Schrag for help in data collection and Dr. Michael Becker and the X25 beamline staff for technical support, and Dr. Allan Matte for comments on the manuscript. The clone for heparinase II expression was provided by Dr. Bernard Eggimann from Biomerlin Pharmaceuticals (Montreal). This research was supported by CIHR grant No. 200009MOP-84373-M-CFAA-26164 to MC.

Table 2-1. Statistics of diffraction data collected from a single heparinase II crystal.

Values in parentheses refer to the last resolution shell	
Space group	$P2_12_12_1$
Unit-cell parameters (Å, º)	$a = 70.0$, $b = 119.3$, $c = 200.7$; $\alpha = 90.0$ $\beta = 90.0$, $\gamma = 90.0$
Crystal mosaicity (º)	0.25
Resolution range (Å)	50.00-2.00 (2.07 – 2.00)
Total No. of reflections measured	817,975
No. of unique reflections	112,668
Average redundancy	7.3 (7.1)
Average $I/\sigma(I)$	10.6 (3.3)
R_{sym}[1]	0.075 (0.393)
Completeness (%)	98.6 (98.0)

[1] $R_{sym} = \Sigma|I_{obs} - Iavg|\Sigma I_{avg}$

Figure 2-2. Oscillation frame obtained at beamline X25, NSLS, BNL.
The he extent of the diffraction pattern is shown by this oscillation frame. The oscillation range is 1º and the exposure time was 7 s.



Chapter 3 Crystal Structure of Heparinase II from *Pedobacter heparinus* and its Complex with a Disaccharide Product

Preface

The contents presented in the following chapter have been published as presented:

Shaya, D., Tocilj, A., Li., Y., Myette, J., Venkataraman, G., Sasisekharan, R., and Cygler, M. (2006) Crystal structure of heparinase II from Pedobacter heparinus and its complex with a disaccharide product. *J Biol Chem.* **22**:15525-35.

Contribution of Authors:

I preformed all the experiments and analyzed the data for native heparinase II. I performed all the crystallographic experiments and the structure determination with the selenomethionine labeled heparinase II and wrote the draft of the manuscript. A.T. and Y.L. instructed me in crystallization steps and computational structure determination. Cloning, expression, purification and selenomethionine labeling of heparinase II in E. coli were preformed by J.M., and supervised by G.V. and R.S. M.C., supervised the project, orchestrated the collaboration, and prepared the final manuscript for publication with my aid.

3.1 Rationale

Obtaining the experimental phases is an impeding step in structure elucidation. In particular, the phases are necessary for the calculation of the Fourier transform deriving electron density from the measured diffraction data. In this chapter, as the next step after obtaining the dataset for nHepII in Chapter 2. I determined the experimental phases utilizing the anomalous signal component in the diffraction data of HepII by employing the Single-wavelength Anomalous Dispersion (SAD) phasing technique. In order to obtain the anomalous signal, HepII was cloned into *E.coli* expression system (rHepII) bearing a histidine tag (to ease the purification) and labeled with selenomethionine. The labeled protein was crystallized, and the crystallographic data were collected. Using the experimental phases I have calculated the electron density map and built and refined the structural model of nHepII. The rHepII was co-crystallized with a tetrasaccharide substrate, which allowed me to determine the structure of HepII complexed with a bound disaccharide product, which highlighted the active site. The co-crystal structure laid the ground for the initial description of the catalytic mechanism, the topic of Chapter 4.

3.2 Abstract

Heparinase II depolymerizes heparin and heparan sulfate glycosaminoglycans yielding unsaturated oligosaccharide products through an elimination degradation mechanism. This enzyme cleaves the oligosaccharide chain on the non-reducing end of either glucuronic or iduronic acid, sharing this characteristic with a chondroitin ABC lyase. We have determined the first structure of a heparin degrading lyase, that of heparinase II from *Pedobacter heparinus* (formerly *Flavobacterium heparinum*), in a ligand-free state at 2.15 Å resolution and in complex with a disaccharide product of heparin degradation at 2.30 Å resolution. The protein is composed of three domains: an N-terminal α-helical domain, a central two-layered β-sheet domain and a C-terminal domain forming a two-layered β-sheet. Heparinase II shows overall structural similarities to the polysaccharide lyase family 8 (PL8) enzymes chondroitin AC lyase and hyaluronate lyase. In contrast to PL8 enzymes, however, heparinase II forms stable dimers, with the two active sites formed independently within each monomer. The structure of the N-terminal domain of heparinase II is also similar to that of alginate

lyases from the PL5 family. A Zn^{2+} ion is bound within the central domain and plays an essential structural role in the stabilization of a loop forming one wall of the substrate-binding site. The disaccharide binds in a long, deep canyon formed at the top of the N-terminal domain and by loops extending from the central domain. Based on structural comparison to the lyases from the PL5 and PL8 families having bound substrates or products, the disaccharide found in heparinase II occupies the "+1" and "+2" subsites. The structure of the enzyme-product complex, combined with data from previously characterized mutations, allows us to propose a putative chemical mechanism of heparin and heparan-sulfate degradation.

3.3 Introduction

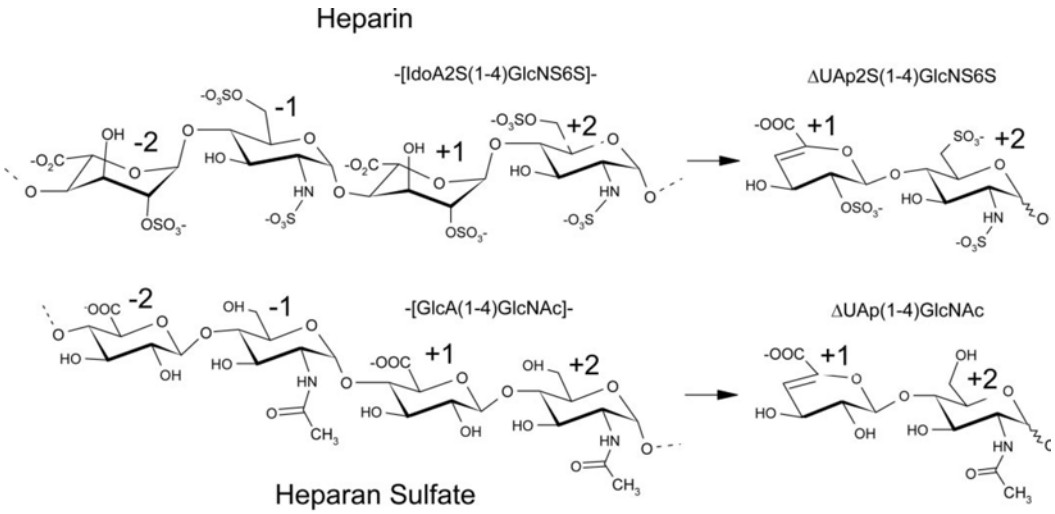
Heparin and heparan sulfate glycosaminoglycans (HSGAGs) are major components of the extracellular matrix and are also found at the cell surface as part of proteoglycan-cell surface receptors (Jackson *et al.* 1991; Kjellen & Lindahl 1991). HSGAGs are negatively-charged linear polysaccharides composed of repeating disaccharide units of hexopyranosyluronic acid (uronic acid, GlcA) and 2-amino-2-deoxy-α-D-glucopyranose (glucosamine, GlcNAc) (Esko & Selleck 2002). Heparan sulfate proteoglycans (HSPG) are abundant on the cell surface as receptors, where two to five polysaccharide chains are attached *via* a serine residue to a core protein (Iozzo 1998). Heparin PGs are primarily synthesized and stored in mast cells. Consequently short oligosaccharide sequences are released as bioactive heparin GAGs by the action of heparin hydrolase. These complex polysaccharides provide docking sites for numerous protein ligands and receptors involved in diverse biological processes, ranging from cancer (Linhardt *et al.* 1986; Sugahara & Kitagawa 2002), angiogenesis (Linhardt *et al.* 1992), anticoagulation (Petitou *et al.* 1999), inflammatory processes, viral and microbial pathogenesis (Shukla *et al.* 1999; Menozzi *et al.* 2002), and multiple aspects of development (Perrimon & Bernfield 2000).

HSGAGs are not homogeneous in their chemical structure but are subject to extensive modification of the linear chain during biosynthesis. These modifications are catalyzed through the sequential action of several enzymes and include epimerization of β-D-glucopyranosyluronic acid to α-L-idopyranosyluronic acid (L-iduronic acid, IdoA),

N-deacetylation coupled with *N*-sulfation of the glucosamine moiety, and *O*-sulfation of hydroxyl groups (2-*O*, 3-*O* or 6-*O*). Heparin typically contains ~90% of iduronic acid and 10% of glucuronic acid, with predominant sulfation of iduronic acid at the 2-OH position. The hexosamine usually contains 3-*O* sulfation and is predominantly *N*-sulfated (GlcNS). Occasionally, it is also present in the *N*-acetylated form (GlcNAc). Occurrence of a glucosamine unsubstituted at the 2-N position is rare. Heparan sulfate is a related polysaccharide with a higher percentage of GlcA and a lower percentage of 2-*O* and *N*-sulfation within the polymer. Modification of HS is not uniform but rather is concentrated within specific regions of the polymer, giving rise to short sequence motifs responsible for the interactions between HS and a diverse repertoire of proteins leading to its multiple biological roles.

Two distinct chemical mechanisms are employed by proteins in order to break down HSGAGs: hydrolysis and β-elimination. Mammalian enzymes employ solely hydrolytic mechanisms to cleave the glycosyl-oxygen (C1-On) bond. The bacterial enzymes are either hydrolases or lyases. Lyases act exclusively on the oxygen-aglycone (On-C4) bond adjacent to the uronic acid moiety at position C5 and result in a product with a C-4–C-5 double bond (Fig. 3-1). The mammalian enzyme heparanase, an endo-β-glucuronidase, cleaves HS utilizing a hydrolytic mechanism. It involves the addition of a water molecule to afford saturated oligosaccharide products (Kussie *et al.* 1999; Toyoshima & Nakajima 1999). Bacterial heparinases on the other hand cleave their polysaccharide substrates through an elimination mechanism (Linhardt *et al.* 1986). The proposed elimination mechanism involves neutralization of the negative charge of the uronic acid acidic group, presumably by interaction with a positively charged amino acid, thereby reducing the acidity of the C-5 proton (Gacesa 1987). Next, a general base abstracts the proton from the C-5 position of the uronic acid, which leads to a β-elimination of the 4-*O*-glycosidic bond concomitant with formation of a double bond between C-4 and C-5 of the uronic acid (Gacesa 1987). In the last step, a general acid donates a proton to the O^- leaving group of a glucosamine, reconstituting the OH function at the reducing end of the cleaved bond.

Figure 3-1. Degradation of heparin and heparan sulfate by heparinase II.
The naming of sugar sites follows the nomenclature introduced by (Davies *et al.* 1997).
The *arrows* indicate the cleavage site.



Heparinases have been identified in several bacterial species, with the enzymes from the soil bacterium *Pedobacter heparinus* (formerly *Flavobacterium heparinum*, (Steyn *et al.* 1998)) being best characterized. This organism produces three heparinases differing in molecular weight, charge properties, and substrate specificities (Nader *et al.* 1990; Lohse & Linhardt 1992). Heparinase I (HepI; heparin lyase I; heparinase; EC 4.2.2.7), with a molecular weight of 43.8 kDa and a pI of 9.3, has a relatively stringent substrate specificity, acting on the heparin-like linkages of 2SIdoA (1-4) 6SGlcNS (Desai *et al.* 1993a; Desai *et al.* 1993b). Heparinase III (HepIII; heparitinase I; EC 4.2.2.8), with a molecular weight of 73 kDa and a pI of 10, prefers an undersulfated GlcNS / GlcNAc (1-4) GlcA heparan-sulfate linkage (Desai *et al.* 1993a; Desai *et al.* 1993b). Heparinase II (heparin lyase II; heparitinase II; no assigned EC number), the largest in the group with a molecular weight of 84.5 kDa and a pI of 8.9, is unique in its ability to cleave both heparin and heparan sulfate–like regions of the polysaccharide, regardless of the sulfation pattern (Nader *et al.* 1990) . Based on its primary sequence, HepI has been assigned to the polysaccharide lyase family PL13 and HepIII to the family PL12, whereas HepII has a unique sequence and has not yet been classified (CAZy database, http://afmb.cnrs-

mrs.fr/CAZY/PL.html). No structural information is available for any of the HSGAG lyases.

Heparinase II (HepII) is a basic protein consisting of 772 residues. An N-terminal signal sequence of 25 residues is cleaved upon export to the periplasmic space to yield the mature protein with a MW of 84,545 Da (Su *et al.* 1996). It acts in an endolytic manner (Rhomberg *et al.* 1998b) and displays broad selectivity, catalyzing the cleavage of linkages adjacent not only to IdoA and GlcA but also to the rare α-L-galacturonic acid residues present in the polymer (Moffat *et al.* 1991a; Moffat *et al.* 1991b). The glucosamine in the substrate could be either GlcNAc or GlcNS. HepII will also degrade the rare disaccharide containing $GlcNH_3$ provided that the uronic acid is 2-*O*-sulfated. Although HepII has greater affinity for heparin, its turnover rate for heparan sulfate is higher (Lohse & Linhardt 1992). Moreover, the enzyme displays preference towards degradation of glycosidic bonds containing GlcA over ones containing IdoA. Heparinase II shows greater catalytic efficiency for longer rather than shorter oligosaccharide substrates (Rhomberg *et al.* 1998b). Previous chemical modification and site-directed mutagenesis studies suggested that cysteine and histidine residues are involved in catalysis (Shriver *et al.* 1998a; Shriver *et al.* 1998b), and that the cleavage of the two substrates, heparin and heparan sulfate, occurs at separate but proximal active sites within a single HepII polypeptide chain (Rhomberg *et al.* 1998b).

We report here the crystal structures of HepII as well as its complex with a heparin disaccharide reaction product. We have adopted the nomenclature of Davies *et al.* (Davies *et al.* 1997) for the sugar binding subsites, where "+" subsites refer to those toward the reducing end of the substrate and "–" subsites to those towards the non-reducing end of the polysaccharide chain. Cleavage occurs between sugars occupying the "–1" and "+1" subsites.

3.4 Experimental Procedures

3.4.1 Protein Cloning Expression and Purification

3.4.1.1 Native Heparinase II

Purification of HepII from *P. heparinus* was carried out as described previously (Shaya *et al.* 2004). Briefly, *P. heparinus* cells overexpressing HepII (Blain et al., 2002)

were cultivated in FH medium (Su et al., 2001) supplemented with 10% (w/v) heparin (Dongying Hi-tech Chemical Industry Co., Dongying City, Shandong Province, People's Republic of China) and 0.1 mg ml^{-1} trimethoprim antibiotic (Sigma-Aldrich, Germany). The cells were harvested, disrupted using a French Press, the lysate clarified by ultracentrifugation (100,000x*g*, 30 min, 4°C) and the protein supernatant containing HepII collected. HepII was purified using two chromatographic steps, SP-Sepharose (Pharmacia Biotech, Uppsala, Sweden) and hydroxylapatite (Bio-Gel HTP, BioRad Laboratories, Richmond, Canada). HepII-containing fractions were concentrated to 6.5 mg ml^{-1} by ultrafiltration using a Centriprep-50 concentrator (Millipore, MA, USA) and the buffer exchanged to 10 mM sodium phosphate pH 7.5, 100 mM NaCl, 20 mM sodium formate, 5 mM DTT.

3.4.1.2 Cloning and Expression of HepII in *E. coli*

P. heparinus was obtained as a lyophilized stock from the American Type Culture Collection (ATCC, Manassas, VA), stock no. 13125. Re-hydrated stock was used to seed a 100 ml culture which was grown aerobically at 30°C with moderate shaking to an optical density (A_{600}) between 1.5 and 2 in defined media containing 0.1% (w/v) heparin as described (Myette *et al.* 2002). Genomic DNA was isolated from the bacterial culture using the Qiagen DNeasy DNA purification kit according to the manufacturer's instructions for gram-negative bacteria based on approximately 2 x 10^9 cells per column. Purified genomic DNA was concentrated by ethanol precipitation and resuspended in Tris-EDTA, pH 7.5 at a concentration of 0.5 mg ml^{-1}. The quality of genomic DNA was assessed by UV absorbance at 260/280 nm, by electrophoresis on a 0.5% agarose gel and confirmed by PCR using *P. heparinus* specific oligonucleotides primers.

The heparinase II gene lacking the signal sequence (residues 1-23) was directly amplified by PCR from *P. heparinus* genomic DNA using the forward primer 5' **CAT ATG** TAT TCC CAA ACC AAG GCC GAT GTG GTT 3' and reverse primer 5' **CTC GAG** GAT CCT TAT CTC AAA AAA CGG TAG GTT CCT 3' (*Nde*1 and *Xho*1 restriction sites are denoted in bold). Oligonucleotides primer sequences were based on the published *hepB* sequence annotated within the NCBI (Genbank) database (accession no. U27585). PCR conditions included 100 ng genomic DNA and Pfx DNA Polymerase

(Invitrogen, Carlsbad CA) for 35 cycles using an annealing temperature of 61°C and 3 minute extensions at 68°C. PCR products were purified by agarose gel electrophoresis and 3' dA overhangs were generated using 200 μM dATP and Taq DNA polymerase (New England Biolabs, Beverly MA) for 10 minutes at 72°C. Modified PCR products were ligated into the TOPO/TA PCR cloning vector 4.0 (Invitrogen) and used to transform One-shot TOP10 chemically competent cells (Invitrogen, Carlsbad, CA). Positive clones were verified by restriction enzyme analysis. The 2.2 kb *hep* II gene was subcloned into the T7-based expression plasmid pET28a (Novagen) as an *Nde* 1-*Xho* 1 cassette for recombinant expression in *E. coli* as an amino terminal His_6 fusion protein to facilitate purification. Plasmid DNA sequencing, using both vector and gene-specific primers, was used to confirm the final expression clones.

Sequence-verified plasmid pET28a/hepIIΔNterm was subsequently transformed into BL21(DE3) for expression. One-liter cultures were grown at room temperature (~20°C) in LB media supplemented with 40 μg ml^{-1} kanamycin, and protein expression induced with 500 μM IPTG added at an A_{600} of 1.0. Induced cultures were allowed to grow for 15-18 hours at room temperature prior to harvesting.

3.4.1.3 Purification of Recombinant Heparinase II

Bacterial cells were harvested by centrifugation (6000 x *g*, 15 min, 4°C) and resuspended in 30 mL of binding buffer (50 mM Na_2HPO_4, pH 7.9, 0.5 M NaCl, 5 mM imidazole). Lysis was performed by the addition of 0.1 mg ml^{-1} lysozyme and incubation for 20 minutes at room temperature followed by intermittent sonication in an ice-water bath using a Misonex XL sonicator at 40-50% output. This lysate was fractionated by centrifugation (20,000x*g*; 30 min, 4°C) and the supernatant filtered through a 0.45 μm filter. Hep II was purified by immobilized metal-chelate chromatography using a 5 mL Hi-Trap column (GE Healthcare) pre-charged with 200 mM $NiSO_4$ and subsequently equilibrated with binding buffer. The column was run at a flow rate of ~3 ml min^{-1} that included a wash step with 50 mM imidazole. HepII was eluted in 5 ml fractions using imidazole elution buffer (50 mM Na_2HPO_4, pH 7.9, 0.5 M NaCl, 250 mM imidazole).

HepII from the IMAC step was buffer exchanged on a Sephadex G-25 column equilibrated with 20mM Na_2HPO_4, pH 6.8, 150mM NaCl. The His_6 tag was cleaved by

addition of biotinylated thrombin at 0.5 units per mg of protein and incubating overnight at 4°C. Thrombin was removed by binding to streptavidin agarose at 4°C for two hours using the Thrombin Capture Kit (Novagen).

The enzyme was further purified on a Source 15S cation exchange column (GE Healthcare). The enzyme was diluted 1:4 (v/v) with 20mM Na_2HPO_4, pH 6.8, loaded onto the column, and a gradient from 0 to 1 M NaCl applied over 10 column volumes. Protein concentration was determined by the Bio-Rad protein assay. Protein purity was assessed by SDS-PAGE followed by staining with Coomassie Brilliant Blue or Sypro Ruby Protein (Molecular Probes).

3.4.1.4 Expression and Purification of Selenomethionine-Labeled HepII

Selenomethionine (SeMet) labelling of HepII by methionine inhibition was performed using a modified M9 medium (Doublie 1997). An overnight culture was used to inoculate 100 ml of LB supplemented with 400 mg l^{-1} kanamycin and grown overnight at 37°C. The culture was pelleted and resuspended in 8 ml of modified M9 medium supplemented with 20% (w/v) glucose, 100 µM $CaCl_2$ and 1X BME vitamin solution (Sigma Chemical Co.). A 1 ml volume of resuspended culture was used to inoculate each of 4 X 500 ml of M9 media and cells grown at 37°C in 2 l baffled flasks until an OD_{600} of 0.5-0.6. Solid amino acid supplement was added as a pre-weighed mixture, and SeMet (Sigma) pre-dissolved in sterile water was added at a final concentration of 50 mg l^{-1}. Cell growth was allowed to continue for ~30 minutes, IPTG added to 0.5 mM and the cells cultured for 15 hours at 30°C. HepII was subsequently purified as described above.

The efficiency of SeMet incorporation was assessed by MALDI-MS using saturated sinapinic acid matrix solution in 50% (v/v) acetonitrile:water as a matrix. Molecular masses were calibrated using both bovine serum albumin (BSA) and β-galactosidase as protein standards. The molecular mass difference between unlabeled and SeMet-labelled enzyme indicated greater than 90% incorporation.

Purified SeMet-labelled protein was concentrated to 6.2 mg ml^{-1} by ultrafiltration using a Centricon YM-100 concentrator (Millipore, MA, USA) and the buffer exchanged to 25 mM Na-Phosphate pH 6, 150 mM NaCl, 10 mM DTT, 0.2% (w/v) NaN_3.

3.4.3 Crystallization and Data Collection

3.4.3.1 Native Heparinase II

Native HepII crystals were obtained from the protein overexpressed in *P. heparinus* and diffraction data collected as described previously (Shaya *et al.* 2004). Briefly, initial crystals were obtained from 6.5 mg ml^{-1} protein in buffer and a reservoir solution containing 17% (w/v) PEG 3350, 200 mM sodium phosphate pH 5. The crystals were optimized by micro seeding. These crystals belong to space group $P2_12_12_1$, with unit cell parameters a = 70.0, b = 119.3, c = 200.7 Å and two molecules in the asymmetric unit. For data collection, crystals were transferred briefly to a cryoprotectant solution (22.5% (w/v), PEG 3350, 0.07 M sodium phosphate pH 5, 2.17 M sodium formate) and flash cooled in a nitrogen stream at 100 K (Oxford Cryosystems, Oxford, U.K.). Diffraction data were collected at beamline X25 (NSLS, Brookhaven National Laboratory) on a Q-315 area detector (ADSC, San Diego, CA) and processed to 2.15 Å resolution using HKL2000 (Otwinowski & Minor 1997). Data collection statistics are summarized in Table 3-1.

3.4.3.2 Selenomethionine-Labeled Heparinase II

HepII overexpressed in *E. coli* did not yield crystals under the same conditions as for the protein produced in *P. heparinus*. Somewhat different conditions were found using the PEG Suite screen (Nextal Biotechnologies, Montreal, Canada) with the hanging drop vapour diffusion method at 20°C. The protein crystallized only when it was supplemented with 3 mM low molecular weight heparin (Sigma-Aldrich Oakville, Canada). Cubic-shaped crystals of SeMet-labelled HepII appeared in drops containing 20% (w/v) PEG 3350, 200 mM magnesium chloride, within 3-5 days. These crystals belong to space group $P2_1$, with unit-cell parameters a = 52.0, b = 163.0, c = 95.2 Å, β = 105.4°, and contain two molecules in the asymmetric unit. The Matthews coefficient V_M is 2.28 $Å^3$ Da^{-1} (Matthews 1968) and the solvent content is estimated to be 46%. Crystals were transferred briefly to a cryoprotectant solution (25% (w/v) PEG 3350, 200 mM magnesium chloride, 12% (v/v) glycerol), and flash cooled in the nitrogen stream at 100 K (Oxford Cryosystems, Oxford, U.K.). Diffraction data extending to 2.3 Å resolution

were collected at beamline X8C (NSLS, Brookhaven National Laboratory), using a Quantum-4 CCD area detector (ADSC, San Diego, CA) at a wavelength of 0.97910 Å corresponding to the measured absorption inflection point, in order to optimize the *f''* component of the Se anomalous scattering. Images were processed with the HKL2000 program package (Otwinowski & Minor 1997) (Table 3-1).

3.4.4 Structure Determination and Refinement

3.4.4.1 E. coli Overexpressed Heparinase II

Single-wavelength anomalous dispersion phasing was performed using the data collected for SeMet-labelled HepII at 0.97910 Å wavelengths. A total of 31 of 42 Se sites were located using the program SOLVE (Terwilliger 2002). The heavy atom sites were refined using the program SHARP (Bricogne *et al.* 2003) and 11 additional Se sites located in the anomalous difference map. Inclusion of all of the identified Se sites gave phases with an overall FOM of 0.35 to 2.5 Å resolution. Solvent flattening with RESOLVE (Terwilliger 2003) increased the FOM to 0.65 and automatically build a partial model consisting of 60% of main chain atoms and 45% of side chain atoms. The model was completed manually using the program O (Jones 1991) and refined using the program REFMAC5 (Murshudov 2003). Since all crystals showed some level of anisotropic diffraction, the refinement protocol included the TLS option with each of the three domains treated separately. The inclusion of TLS reduced the R-free and R-work by ~2% but the average B-factor for the protein was very little affected. The analysis of TLS indicates largest liberation along the direction approximately parallel to the long axis of the molecule (Supplementary material). Solvent molecules were added using ARP/wARP (Perrakis *et al.* 1999). Electron density maps showed the presence of a disaccharide within the substrate-binding site of each molecule which was subsequently modeled as a tri-sulfated disaccharide corresponding to the heparin-degradation product, 2SΔUap (1-4) 6SGlcNS A strong peak in the electron density surrounded by two histidines, aspartate and water molecules was deemed to be a metal ion. This was assumed to be a Zn^{2+} ion based on the ligand geometry and on the fact that its refined B-factors were comparable to those of neighbouring atoms. For other metals tested the B-factors differed significantly from the neighbours. Refinement at 2.3 Å resolution converged with an

R_{factor} of 0.176 and R_{free} of 0.221, with the model consisting of residues Asp30-Arg772 of each monomer, each with a bound metal ion (Zn^{2+}) and a total of 783 water molecules. The final refinement statistics are shown in Table 3-1.

3.4.4.2 *P. Heparinus* Overexpressed Heparinase II

The HepII model was positioned within this unit cell using MolRep (Vagin & Teplyakov 1997) and subsequently refined at 2.15 Å resolution using REFMAC5, to an R_{factor} of 0.193 and R_{free} of 0.232 (Table 3-1). The N-terminal residue was identified as a pyroglutamate, formed following cleavage of the signal sequence (Su *et al.* 1996). The final model includes two molecules, each containing residues 26–772. A metal ion with octahedral coordination was identified in the electron density map. Based on the nature of coordinating groups, and their distances to the metal ion (2.2 – 2.3 Å), it was inferred to be a Zn^{2+} ion. The electron density map showed that Thr134 was *O*-glycosylated and a tetrasaccharide corresponding to the known sequence of a typical *O*-glycosylation site found in *P. heparinus* (Féthière *et al.* 1999) fitted well to the electron density. The model also contains 939 water molecules, two phosphates and three formate molecules.

For all molecules within the different crystal forms, Asn405 is in the disallowed region of the Ramachandran plot. This residue has well-defined electron density and is involved in product binding. Each molecule contains four *cis*-peptides modeled adjacent to proline residues.

Table 3-1. X-ray crystallographic data

Parameters	Values	
	Native HepII	SeMet HepII
Data collection:		
Space group	$P2_12_12_1$	$P2_1$
Cell dimensions:		
a (Å)	70.0	52.1
b (Å)	119.3	163.3
c (Å)	200.7	95.3
α (degrees)	90.0	90.0
β (degrees)	90.0	105.5
λ (degrees)	90.0	90.0
Wavelength (Å)	0.97940	0.97910
Resolution range (Å)	102.6-2.2 (2.21-2.15)	91.7-2.3 (2.36-2.30)
Observed / unique reflections[a]	817,975 / 86,490 (6290)	429,705 / 126,846 (4120)[b]
Average redundancy	7.4 (7.3)	3.4 (3.1)[b]
Completeness (%)	98.7 (98.2)	96.0 (86.4)
R_{sym}[c]	0.075 (0.226)	0.11 (0.487)
Average I/σ (I)	30.8 (10.1)	10.00 (2.0)
Refinement statistics:		
R-work[d]	0.193	0.176
R-free[d]	0.232	0.221
No. of atoms/average B-factor:		
Protein (main chain + side chain)	11,938/27.8	5936/14.1 (molecule A) and 5936/31.2 (molecule B)[e]
Disaccharide product	----------	35/21.9 (molecule A) and 35/53.2 (molecule B)[e]
O-Linked glycan	82/49.1	
Ligand (formate and phosphate)	19/57.2	
Solvent	928/31.1	783/21.5
Ions (Zn^{2+})	2/26.7	
Ramachandran plot:		
Allowed (%)	99.0	98.8
Generously allowed (%)	0.7	1.0
Root mean square deviation:		
Bonds (Å)	0.008	0.008
Angles (degrees)	1.10	1.12

[a] All measured reflections with I > 1σ(I).

[b] The Friedel pairs were not merged.
[c] $R_{sym} = (\Sigma|I_{obs} - I_{avg}|)/\Sigma I_{avg}$.
[d] R-work/free = $(\Sigma|F_o - F_c|)/\Sigma F_o$.
[e] The two molecules in this crystal form had substantially different B-factors.

Coordinates of native HepII and the HepII–2SΔUap (1–4) 6SGlcNS complex have been deposited in the Protein Data Bank (Berman *et al.* 2000) with codes 2FUQ and 2FUT, respectively.

3.5 Results and Discussion

3.5.1 Overall Monomer Architecture

The HepII crystal structure revealed a homodimer in the asymmetric unit. Each monomer has an α + β architecture consisting of three domains with an overall S-like shape with dimensions of 95 x 55 x 50 Å (Fig. 3-2e). The N-terminal domain (residues 26-356) contains fourteen α-helices arranged in a double-layered $(\alpha/\alpha)_6$ toroid. The central domain (residues 357-676) contains sixteen β-strands arranged in a two-layered stack of β-sheets. Several of the loops connecting the strands of this domain are long, contain short secondary structure elements and extend away from the β-sheet. The C-terminal domain (residues 677-772) contains nine β-strands packed together in a manner resembling a β-barrel.

The N-terminal domain begins with two short β-strands in an anti-parallel arrangement and continues with 13 α-helices forming a double-layered $(\alpha/\alpha)_6$ toroid (Fig. 3-2b). Helices α4-α13 (Fig. 3-2a) are arranged in six α-helical pairs into a toroid and are inclined by ~30-45^o to the toroid axis. Five of them are hairpins formed by helices consecutive in sequence (α4-5, α6-7, α8-9, α10-11 and α12-13) while helices α2 and α14, distant in linear sequence, form the sixth pair. The inner layer of the toroid is formed by helices α4, α6, α8, α10, α12 and α14, with their N-termini at the top of the toroid, proximal to the central domain. The outer layer consists of helices α5, α7, α9, α11, α13 and α2, all oriented in the opposite direction to the inner helices.

The bottom of the toroid, opposite to the central domain (Fig. 3-2a), is formed by short loops ranging from 3 to 9 residues connecting the helices within each hairpin. The first 60 N-terminal residues cover this face of the toroid, with Val32-Ile43 forming a α-

hairpin and Leu51 and Tyr52 protruding deep toward the center of the domain and contacting the inner helices α6, α8, α10 and α14 of the toroid. Loops connecting the consecutive hairpins are located at the top of the toroid, proximal to the C-terminal domain. They are substantially longer than the hairpin loops and extend away from the toroid. These loops create a deep canyon along the top of the toroid with two high rims at opposite sides (Fig. 3-2a). The long loop Arg296-Tyr318, connecting the fourth and fifth helical hairpins, together with the side chains belonging to the N-terminal part of helix α10 forms the bottom of the canyon. This loop and the exterior of helices α9 and α11 provide the interface for packing against the C-terminal domain. Residues Gln359-Pro369 comprise a surface-exposed loop connecting the N- and central domains.

The central domain (residues Leu370-Asn676) is composed of twenty anti-parallel β-strands arranged in two β-sheets, a short distorted α helix and two short β-hairpins (Fig. 3-2c). The sheets are roughly parallel to each other forming a two-layered β-sandwich (Fig. 3-2a). Both sheets are extensive, each comprised of eight β-strands. These sheets are twisted, with a ~60° angle between the first and last strands. The first β-sheet starts with four long β-strands (β1-β4) that form the surface contacting the N-terminal domain. The fifth strand, β5, is composed of two segments (β5a and β5b, respectively, Fig.3-2c) interrupted by a long Ω loop formed by Ile423-Asn452. Two other long loops extend from this face of the first β-sheet extending the second rim of the canyon described above. These long loops contain the short secondary-structure elements mentioned above. The tips of these loops (Leu404-His406, Tyr429-Asn437, Gly467-Ser471 and Ile491-Ala492) form a wall that extends the canyon located at the top of the N-terminal domain and contribute to the product-binding site. The second β-sheet is composed of 8 long β-strands, five of them over twelve residues long. There are four crossovers between the two β-sheets, all of them on the same side of the sheets (Fig. 3-2c).

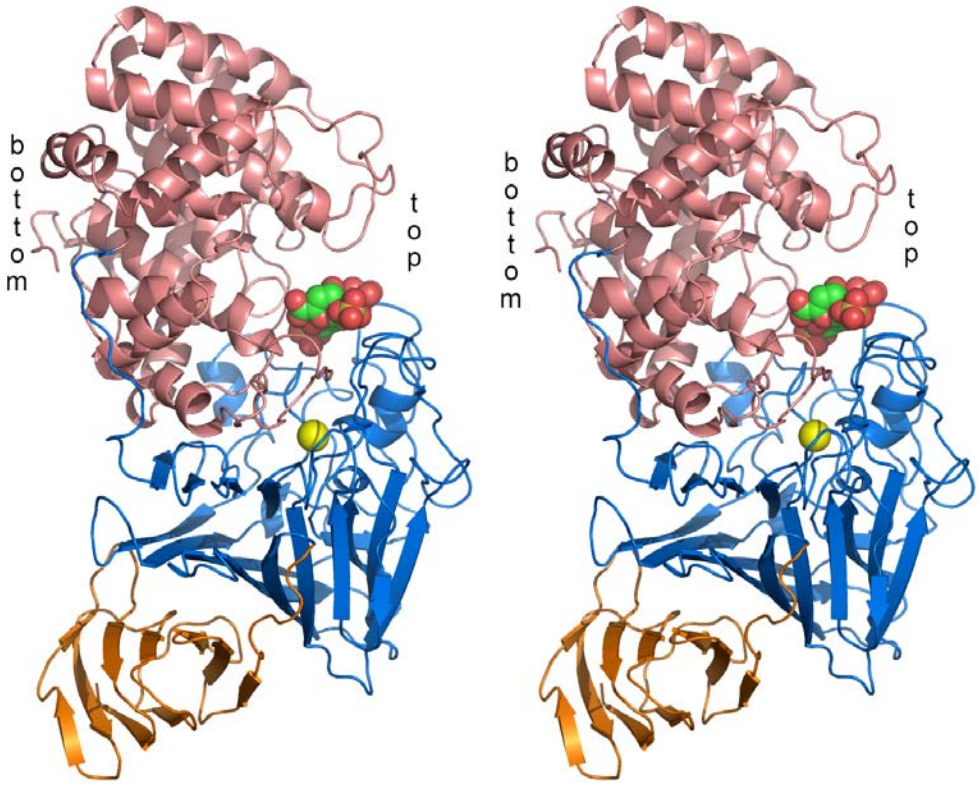
The C-terminal domain, composed of residues His682-Arg772, is connected to the central domain by a short linker. This domain is composed of two β-sheets with three strand crossovers between them (Fig. 3-2d). These sheets are smaller than those of the central domain and are composed of shorter strands. The first β-sheet is composed of five

strands, while the second has four β-strands. This domain packs against the central domain together forming a four-layered β-sheet stack.

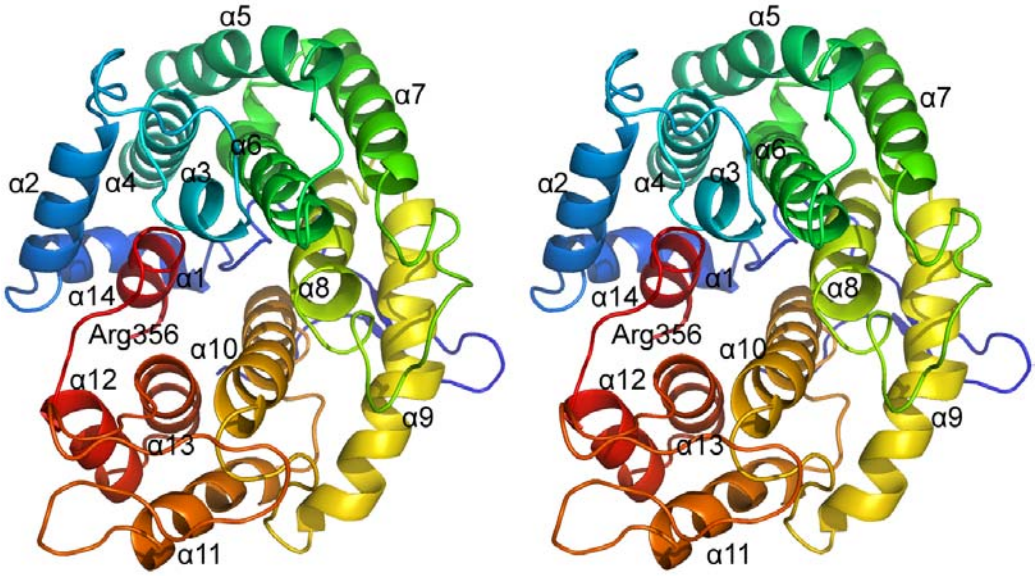
3.5.2 Zinc Binding Site

The initial SAD electron density map showed a strong peak located near the interface of the N- and C-terminal domains. Asp425, His408, His451 and three water molecules surround this peak with nearly perfect octahedral coordination by the nitrogen and oxygen atoms with distances of 2.1-2.4 Å (Fig. 3-2). Based on the type of coordinating atoms, distances and coordination geometry, a Zn^{2+} atom was modeled into the electron density. The refined B-factor of this atom is similar to the B-factors of surrounding residues. Other plausible metal ions tested during the refinement resulted in unreasonable B-factors. The residues Asp425 and His451 coordinating the Zn^{2+} ion are located at the ends of the Ile424-His451 Ω loop extending from the middle of strand β5. The zinc ion assists in holding these ends together. His408 comes from the neighbouring loop connecting strands β3 and β4. Both loops participate in interactions with the N-terminal domain and in particular with the tip of the Gln301-Arg312 loop. Two of the water molecules coordinating the Zn^{2+} ion form hydrogen bonds to Gly306 from this loop and participate in a hydrogen bond network bridging the two domains. The Zn^{2+} clearly plays a structural role and is not directly involved in catalysis being located >15 Å from the substrate-binding site.

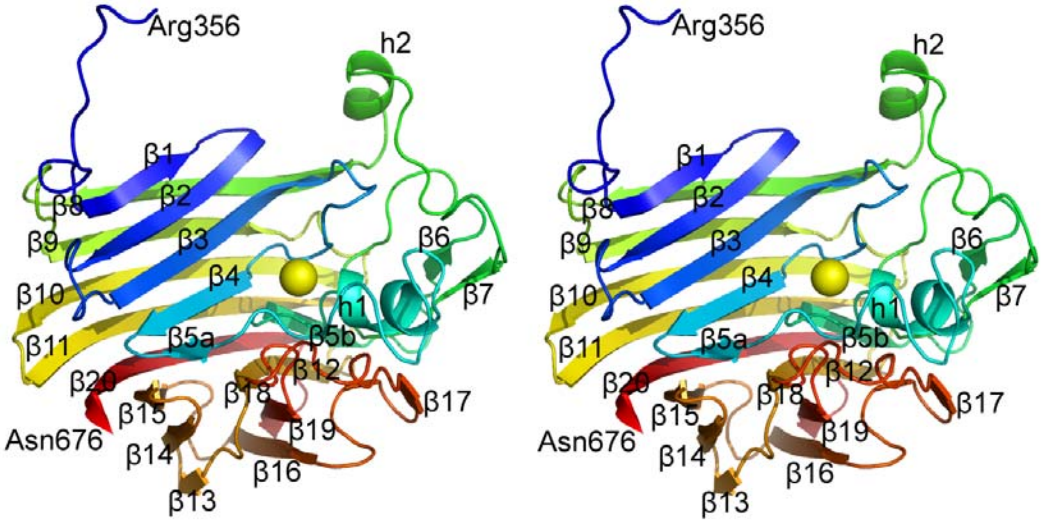
Figure 3-2. Stereo view of the schematic representation of heparinase II.
A. Monomer, with N-terminal domain (salmon), central domain (blue), and C-terminal domain (orange). The product is shown in a CPK representation, and the Zn^{2+} ion is shown as a yellow sphere.



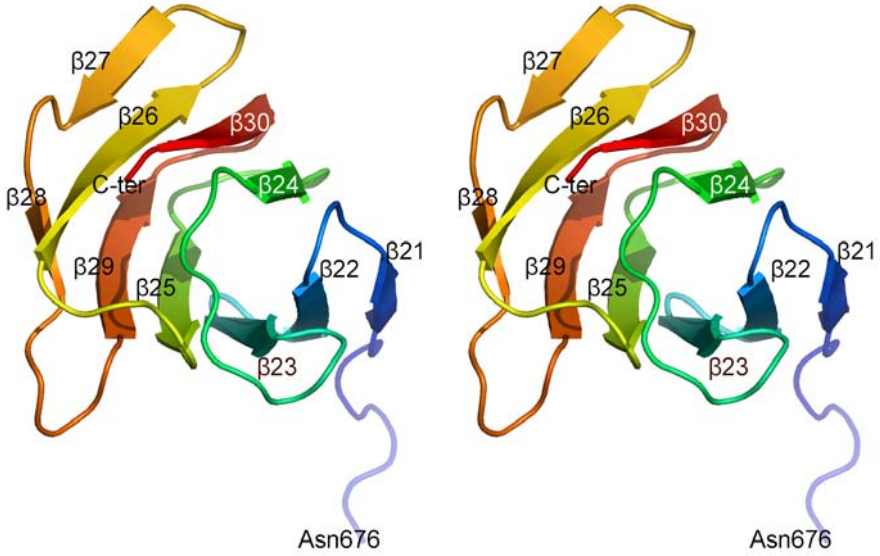
B. The N-terminal domain colored blue to red from the *N*- to *C*- terminus. Secondary structure elements are marked.



C. The middle domain, colored as in B.



D. The C-terminal domain, colored as in B.



E. Dimer, the first molecule colored as in A and the second molecule colored in shades of green. On the left is the schematic representation, and on the right is the surface representation. The C-terminal domain of one molecule packs into a depression between the central and the *C*-terminal domains of the other molecule. This figure was prepared using PyMol (available on the World Wide Web at www.pymol.org).

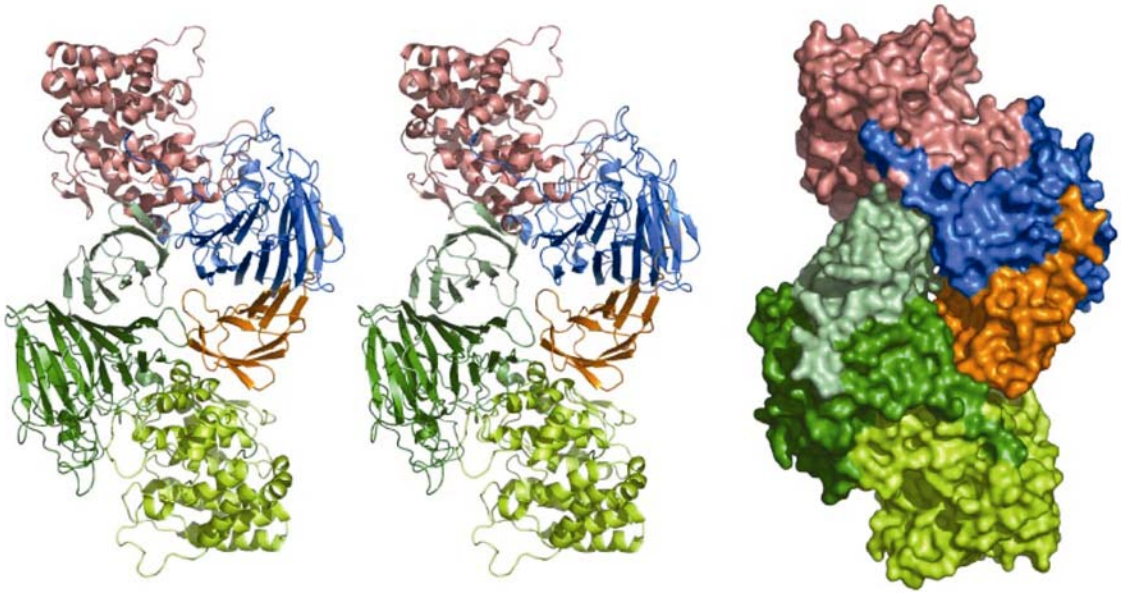
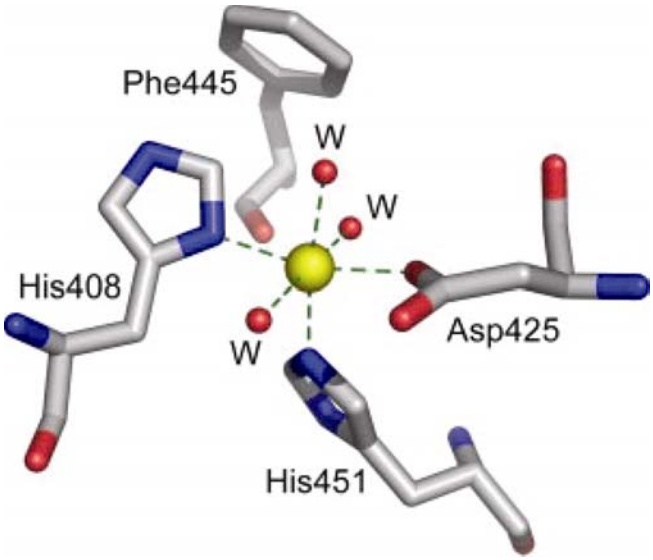


Figure 3-3. Zn^{2+} tetrahedral coordination sphere with the ligands.
A 3.0-Å distance cut-off was applied to choose neighbouring residues. The distances between Zn^{2+} ion and the liganding atoms (in molecule A nHepII) were as follows: $ND1^{His-408}$, 2.23 Å; $NE2^{His-451}$, 2.14 Å; $OD2^{Asp-425}$, 2.08 Å; Waters, 2.20, 2.23, and 2.31 Å.



3.5.3 Glycosylation of Native HepII

P. heparinus has the enzymatic machinery for the glycosylation of proteins that are directed to the periplasmic space (Reinhold *et al.* 1995). Of the five GAG lyases produced by this bacterium (Heparinase I-III, chondroitinase AC & chondroitinase ABC), all except HepIII are glycoproteins. The sequence of a branched heptasaccharide, attached to Ser39, in HepI has been established using a combination of enzymatic digestion, NMR and mass spectrometry (B. Eggimann, personal communication). This sequence is galactose-ɑ(1→4) [galactose-α(1→3)] (2-O-Me) fucose-ɑ(1→4) xylose-ɑ(1→4) glucuronic acid-α (1→2) [rhamnose-α(1→4)] mannose-α (1→O) Ser. The same oligosaccharide was identified previously in the structures of chondroitinase AC (Féthière *et al.* 1999c) and chondroitinase B (Huang *et al.* 1999). The structure of native HepII expressed and purified from *P. heparinus* showed electron density extending from the side-chain of Thr134 indicating that this is a site of *O*-glycosylation. Four sugars linked to this threonine side chain, Man-[Rha]-GlcUA-Xyl, were fitted to the electron density. Weak density indicating the presence of two more sugars was visible but was too weak to reliably model. The branched rhamnose sugar fills a depression in the protein surface

formed between the loop Asp83-Gly103 and helix α6 and forms direct as well as water-mediated hydrogen bonds to the protein. The remaining sugars, in the main oligosaccharide branch, extend away from the surface of the protein.

3.5.4 Dimer Formation

Gel filtration and dynamic light scattering studies both indicate that HepII forms homodimers in solution. Both HepII crystal forms showed the presence of homodimers (Fig. 3-2e). The dimers are formed from the two independent molecules in the asymmetric unit and are related by 2-fold non-crystallographic symmetry. The association of molecules into dimers is the same in all crystal forms; the different dimers superimpose with a root-means-square deviation of 0.4 Å for all Cα atoms. The dimer has an elongated shape with approximate size of 130 x 85 x 50 Å.

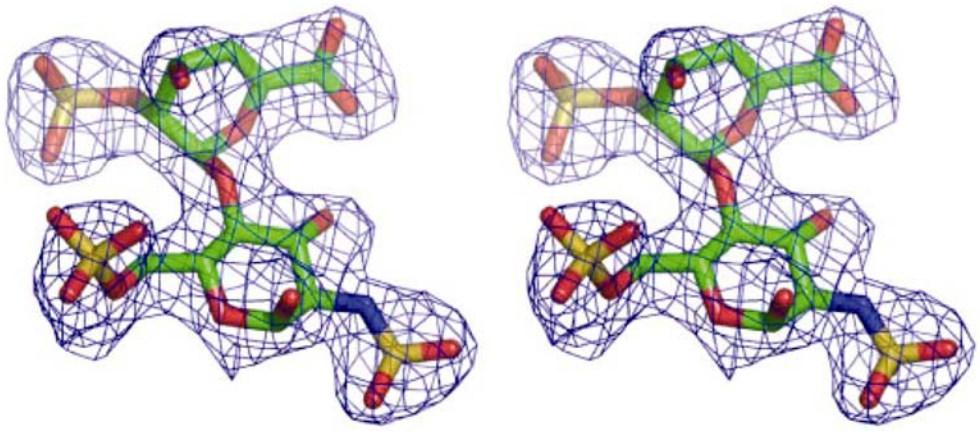
The surface area buried upon dimer formation calculated using the method of Lee and Richards (Lee & Richards 1971) with a 1.4 Å probe radius is ~1300 $Å^2$ per molecule, corresponding to ~5% of the total surface area of each monomer. The dimer is formed by the outside-facing β-sheet of the C-terminal domain packing against the concave surface made of the helices α9, α11, α13 of the N-terminal domain and the edge β-strands of the two β-sheets of the central domain of the other monomer (Fig. 3-2e). The intermolecular interactions include van der Waals contacts and hydrogen bonds, many of them bridged by over 70 water molecules and two phosphate ions within the interface. The dimer interface is located opposite to the substrate-binding canyon, suggesting that the active sites of each Hep II monomer could function independently.

3.5.5 Substrate Binding-Site

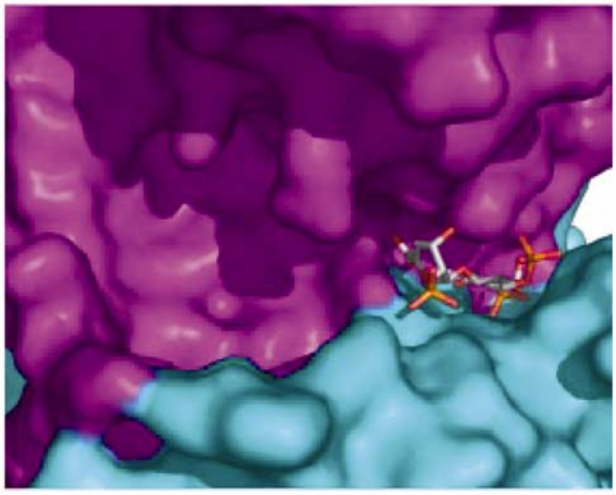
The structure of HepII crystallized in the presence of heparin included electron density that we have identified as a disaccharide, corresponding to the elimination degradation product of heparin, with an unsaturated uronic acid at the non-reducing end (Fig. 3-4a). The disaccharide product, 2SΔUap (1–4) 6SGlcNS binds within a deep, elongated canyon that spans ~30-35 Å in length. This canyon is formed at the top of the N-terminal domain and is extended by loops from the central domain (Fig. 3-4b). It is clearly divided into two parts separated by a ridge, (1) the narrow less deep end where the

"+" sites are located and (2) the wide and deeper part in which the oligosaccharide on the non-reducing side of the cleaved bond resides. The disaccharide occupies the narrow section formed at the interface between the N-terminal and central domains, with its non-reducing end (2SΔUAp) pointing toward the center of the canyon (Fig. 3-4b). This section is ~15 Å long, ~11 Å deep and 7-8 Å wide.

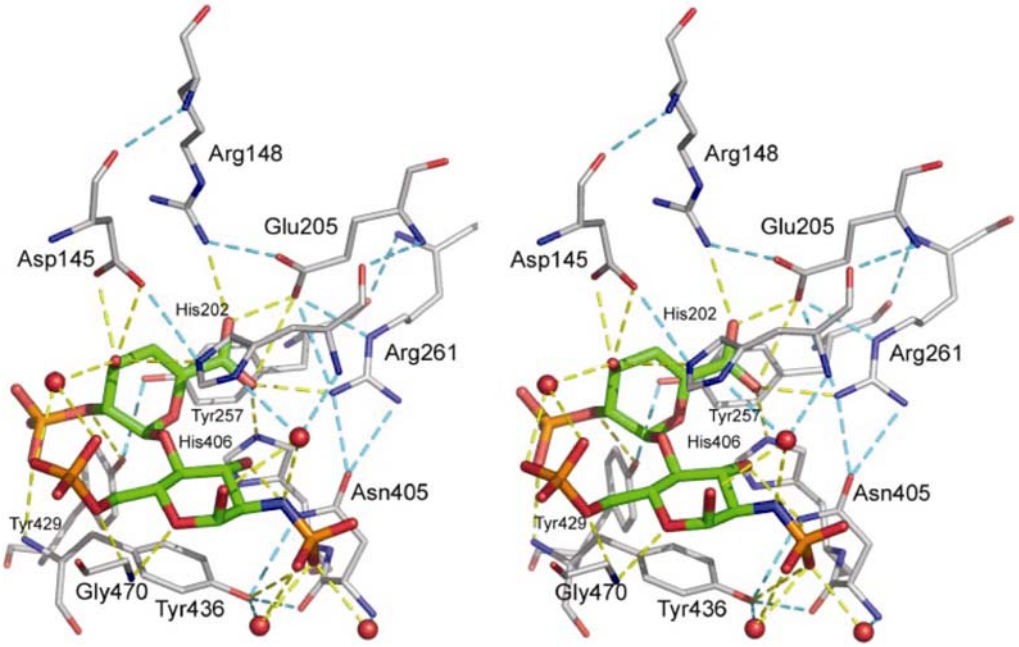
Figure 3-4. Substrate binding site.
A. Disaccharide product ΔUAp2S (1–4) GlcNS6S with electron density ("omit map" calculated with phases derived from the model without the disaccharide) contoured at the 3σ level.



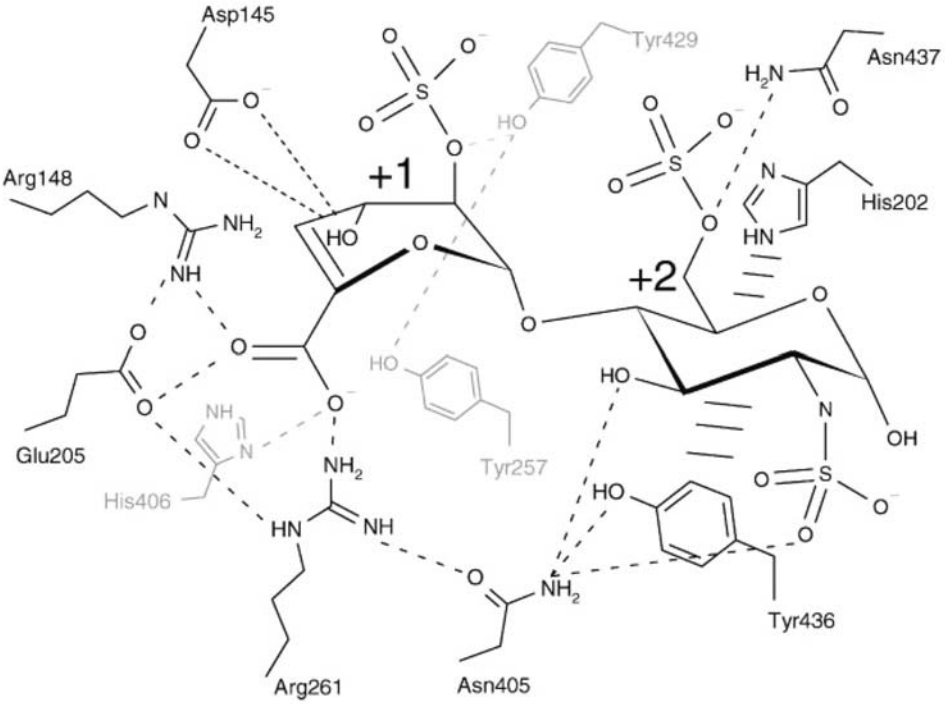
B. Surface representation of the binding site with a disaccharide product shown in a *ball-and-stick representation*. The N-terminal domain is shown in *magenta*, and the central domain is *cyan*. The product occupies the plus sites, and the minus sites are empty.



C. Ball-and-stick representation of the disaccharide product bound to HepII. The disaccharide is shown in thicker lines, its carbon atoms are green, the surrounding HepII residues are shown in thin lines, and their carbon atoms are grey. The hydrogen bonds are shown by dashed lines and are coloured yellow when between the disaccharide and protein residues and cyan between the protein residues. The red spheres represent water molecules participating in the hydrogen-bonding network.



D. Schematic representation of interactions between the disaccharide product and HepII. The residues shown in grey approach the sugars from below. Three parallel lines indicate stacking of a side chain with the sugar ring. Substituents in positions 1, 2, and 3 of the uronic acid are in axial orientations in our structure.



Residues located in loops belonging to the N-terminal domain (in particularly loops connecting helices α7-α8, α9-α10, α11-α12) form one rim of the canyon. The opposite rim of the canyon is composed of residues belonging to the loops of the C-terminal domain (the Ω loop connecting strand α5a to α5b, loops connecting strands α5b to α6 and α8 to α9). Elaborate hydrogen-bonding networks maintain the position and correct orientation of the side-chains creating the binding site (Fig. 3-4c,d). One such network involves the sidechains and some mainchain atoms of Tyr436, Asn405, Arg261, Glu205 and Arg148. The last three residues form salt bridges and make contacts to the carboxylic group of the uronic acid substrate. Another hydrogen-bonded network includes the sidechains of His406, Asp307, Lys446 and Tyr468.

The disaccharide sugar rings are approximately parallel to the walls of the canyon. The carboxylic group of the uronic acid and the *N*-sulfate of the glucosamine point towards the bottom of the cleft (Fig. 3-4a) while the 6-O and 2-O sulfate groups of the heparin disaccharide 2SΔUap (1→4) 6SGlcNS point upward and are exposed to the

solvent. The disaccharide forms intimate contacts with the protein through hydrogen bonds, stacking and van-der-Waals interactions. All hydroxyl groups of the disaccharide are engaged in hydrogen bonds to protein side-chains, one bridged by a water molecule (Fig. 3-4c, Table 3-3). The sugar ring of glucosamine is sandwiched by two aromatic rings, Tyr436 on one side and His202 on the other side. Such stacking interactions are typical for carbohydrate-protein interactions in general (Quiocho *et al.* 1997; Muraki 2002; Sharon & Lis 2002) and have been observed in other GAG lyases (Huang *et al.* 2001). The carboxylic group of the uronic acid is firmly held by interactions with four sidechains: His406, Arg261, Glu205 and Arg148 (Fig. 3a,c). Apart from His406, the atoms hydrogen bonded to the carboxylic group are nearly coplanar. The sidechain of His406 forms a second hydrogen bond to the carboxylic groups of Asp307 and therefore has to be protonated. Furthermore, the very close distance of 2.4 Å between the carboxylic oxygen O6a of the uronic acid and the OE2 atom of Glu205 clearly indicates that one of these two acidic groups is protonated. Based on the short distance between these two oxygens it is plausible to suggest that this proton participates in a low-barrier hydrogen bond (Cleland *et al.* 1998).

As expected from the ability of the enzyme to cleave sulfated as well as non-sulfated polymers, the sulfate groups contribute little to substrate binding due to their limited contacts with the protein. Only three hydrogen bonds are formed, a direct H-bond between 6-S and backbone NH of Tyr436, and two H-bonds mediated through water molecules (Fig. 3-4b, Table 3-2). Nevertheless, even this small binding contribution by the sulfate groups may lead to the observed higher affinity of HepII toward heparin substrates (Lohse & Linhardt 1992).

Based on the proximity of the sidechain of Tyr257 to the unsaturated uronic acid (OH^{Tyr257}-C-5^{2SUAp} distance 3.8 Å) and its superposition on the catalytic tyrosine of chondroitin lyase and alginate lyase (see below) we propose that this residue plays a key catalytic role. Another potential candidate for a catalytic residue is His202, which is on the opposite side of the uronic acid with the $NE2^{His202}$–C-$5^{2S\Delta UAp}$ distance of 4.2 Å.

The disaccharide product occupies the "+1" and "+2" subsites. This differentiates HepII from other lyases, where the disaccharide product binds more tightly to the "–"

subsites, as observed in the crystal structures of enzyme-product complexes (Huang *et al.* 1999; Yoon *et al.* 2001).

Previous investigations showed that HepII cleaves long oligosaccharides preferably near their reducing ends, releasing di- or tetrasaccharides (Rhomberg *et al.* 1998b). This is consistent with the HepII structure presented here, which shows that the "+" side (narrow end) of the canyon can accommodate most likely 3-4 subsites, while the "–" half-site in the wider and longer part of the canyon could accommodate more subsites.

3.5.6 Structural Similarities

Although there is no detectable sequence similarity between HepII and other lyases, its structure shows clearly recognizable similarity to polysaccharide lyases from two unrelated sequence-families, PL5 and PL8.

The PL5 family contains predominantly alginate lyases, ~370 amino acid long single domain enzymes, though some of them have an additional ~250 amino acid long C-terminal domain of presently unknown three-dimensional structure (CAZy, http://afmb.cnrs-mrs.fr/CAZY/PL.html). They have similar topology to the N-terminal domain of HepII and are represented by *Sphingomonas sp.* alginate lyase A1-III (PDB codes 1QAZ and 1HV6 (Yoon *et al.* 1999)). Like HepII, this alginate lyase has a toroidal shape made of five α-helical hairpins and a sixth pair formed by joining the N- and C-terminal helices. Its structure was also determined in complex with a trisaccharide reaction product located in subsites "–1", "–2" and "–3". In the superposition of these two enzyme structures, the trisaccharide present in AI-III forms a continuation of the disaccharide located in HepII, supporting the notion that this disaccharide occupies the "+1" and "+2" subsites. Furthermore, the catalytic Tyr246 of alginate lyase occupies the same position as Tyr257 of HepII. There are differences in the loops surrounding the substrate-binding cleft, most likely related to the fact that alginate lyase does not posses another domain to aid its catalytic activity.

Topological similarity to lyases from family PL8 extends not only to the N-terminal domain but also to the other domains. This family includes chondroitinases AC, chondroitinases ABC, hyaluronate lyases and xanthan lyases (CAZy, http://afmb.cnrs-

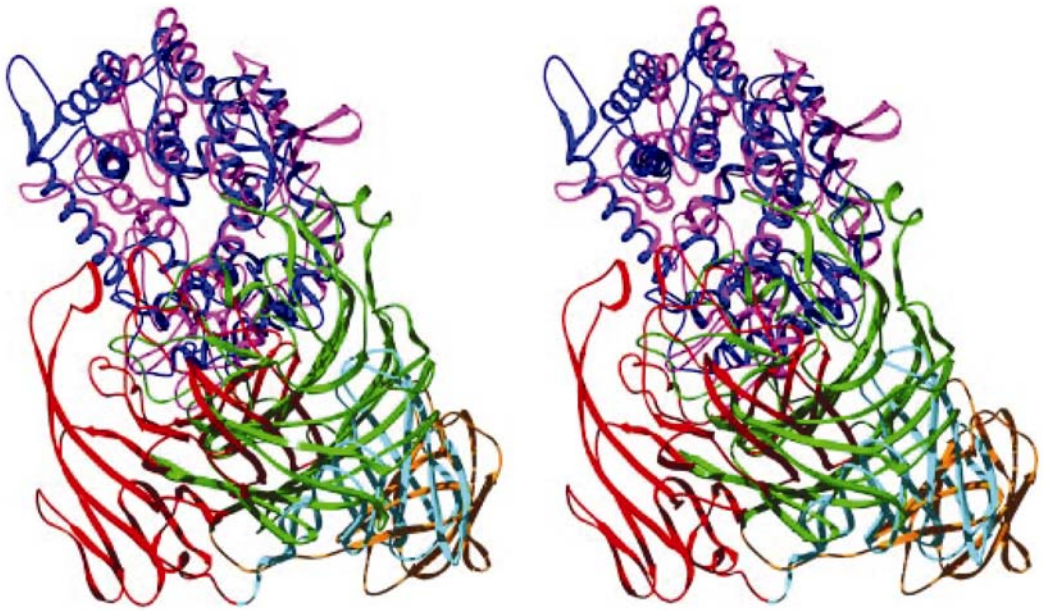
mrs.fr/CAZY/PL.html). These enzymes are ~700-1000 amino acid long and contain several domains (Féthière *et al.* 1999c; Li *et al.* 2000; Hashimoto *et al.* 2003; Huang *et al.* 2003). Like HepII, they have an N-terminal toroidal α-helical domain and their superposition on the N-terminal domain of HepII shows that the toroid architecture (positions and orientations of helices) is well preserved, especially the five inner helices. The loops connecting the hairpins differ in length and conformation but the common features are long canyons along the top of their N-terminal domains where the substrate-binding site is located. The shapes of these canyons are influenced by the loop conformations and differ between the aligned protein structures. The C-terminal helix α14 in HepII, which together with helix α2 forms the sixth hairpin, has its counterpart in PL8 family lyases displaced by ~10 Å and oriented nearly perpendicularly to the toroid axis, leaving the toroid partially open (Lunin *et al.* 2004). This structural alignment also resulted in a superposition of the HepII reaction product onto the "+1" and "+2" sugar-binding subsites of chondroitinases AC (Huang *et al.* 2001; Lunin *et al.* 2004). Moreover, this superposition also juxtaposes the catalytically essential tyrosines sidechains of chondroitinase AC (Tyr234 in *P. heparinum* and Tyr239 in *A. aurescens*) and hyaluronan lyase (Tyr408 in *Streptococcus pneumoniae*) with Tyr257 of HepII.

The central domain and C-terminal domain of HepII form together a four-layered β-sheet stack and its topology is similar to the C-terminal domain of PL8 family lyases. In the description of chondroitinase AC and hyaluronan lyase topology, the C-terminal half of the protein, was described as a single domain (Féthière *et al.* 1999c; Li *et al.* 2000). Comparison with HepII indicates that a two-domain description of this region, like that adopted for HepII, better represents the actual topology. Superposition of the central and C-terminal domains of HepII on the corresponding domain(s) of the other two proteins shows that their β-sheets overlay very well with an r.m.s deviation in the range of 1.5-1.7 Å. However, the relative orientation of these two domains in HepII is different from that in chondroitinase AC and hyaluronan lyase (Fig. 3-5), justifying their assignment as separate domains.

The overall structure of HepII differs from that of other family PL8 enzymes not only in the relative orientation of the central and C-terminal domains but also in the relative orientation of the N-terminal and central domains (Fig. 3-5). This change in

orientation of the central domain in HepII relative to that in other PL8 enzymes brings it more into contact with the substrate. Another difference between HepII and family PL8 proteins is their oligomeric state; while HepII form dimers, other PL8 enzymes exist in solution as monomers.

Figure 3-5. Comparison of overall structures of HepII and Chon AC.
Both structures are shown in a simplified schematic representation. The Hep2 domains are as follows: N-terminal (magenta), central (green), and C-terminal (orange). The ArthroAC domains are as follows: N-terminal (blue), central (red), and C-terminal (cyan). The superposition is based on N-terminal domains. The central domains (green and red) and the C-terminal domains (orange and cyan) in these two proteins are oriented differently relative to the N-terminal domain.



3.5.7 Mutagenesis and Chemical Modification Studies

Previous structure-function studies of *P. heparinus* HepII have been carried out through the complementary use of chemical modification and site-directed mutagenesis (Shriver *et al.* 1998a; Rhomberg *et al.* 1998b; Shriver *et al.* 1998b). In particular, these studies focused on the potential catalytic role of cysteine and histidine residues. The results from these mapping studies indicated that of the three cysteines, present in HepII, Cys348 was uniquely susceptible to chemical modification by sulfhydryl-reactive reagents. Mutation of this cysteine (Cys348Ala) resulted in a selective loss of activity

towards heparin but not toward heparan-sulfate (Shriver *et al.* 1998a). Indeed, the HepII structure indicates that Cys348 is the only cysteine with a surface accessible sulfhydryl group. It is located in the substrate-binding canyon at either the "–2" or "–3" subsite and could interact with the substrate. Indeed, pre-incubation of HepII with either heparin or heparan sulfate protected Cys348 against this modification. The effect of the Cys348Ala mutations only on heparin substrate suggests differences in interactions at the "–" subsites of an oligosaccharide containing iduronic acid vs. glucuronic acid.

Using a similar chemical modification and site-directed mutagenesis approach, the influence of His238, His406, His408, His451 and His579 on catalysis was investigated (Shriver *et al.* 1998b). Only His238, His451 and His579 were accessible to chemical-modifying reagents. The results of these experiments are somewhat difficult to reconcile with the structure of HepII since these three histidine sidechains are not readily surface accessible but rather buried within the protein. Modification could result from either flexibility of the molecule in solution allowing access of diethylpyrocarbonate or partial unfolding of HepII in the presence of this molecule. On the other hand the effect on activity of mutating these histidines to alanines agrees well with the structure reported here. His238 is located within a bent helix, α9, where the disruption of hydrogen bonding pattern occurs. This sidechain provides support to the loop Glu188-Ala203, which forms part of the substrate-binding site, specifically making a hydrogen bond to the backbone carbonyl of Ala203. Removing the His238 sidechain most likely affects the conformation of this loop and compromises substrate binding. Mutating His408 and His451 to alanines abolishes the enzyme's activity towards both substrates. As these two sidechains are involved in coordination of the Zn^{2+} ion, their mutation would destroy the zinc-binding site and affect local structure of the Ω loop, which contains residues interacting with the substrate (Tyr429, Tyr436 and Asn437). Additionally, the neighbouring His406 and Asn405 are part of the substrate-binding site and form hydrogen bonds with sugars in "+1" and "+2" subsites. Finally, His579 located at the end of the sixth strand of the first β-sheet in the central domain is distant from the active site but close to the Zn^{2+} binding site and forms hydrogen bond to Ile449. Mutation of His579 most likely affects the position of His451 and leads to disruption of the Zn^{2+} binding site, which in turn alters the Ω loop conformation. Based on these data, the role of Zn^{2+} appears to be preservation

of the structural integrity of one wall of the substrate binding site, while the rest of the residues pointed in the studies mentioned are either in direct contact with the substrate or have an effect in positioning important residues for substrate binding.

3.5.8 Putative Catalytic Mechanism

HepII has the ability to cleave the oxygen-aglycone bond between the uronic acid and hexosamine irrespectively of the configuration at the C-5 atom (iduronic or glucuronic acid). It is in this respect similar to chondroitinase ABC but differs from the stereo specific hyaluronan lyases, chondroitinase AC and chondroitinase B. The HepII-disaccharide reaction product complex provides the first insight into the active site and allows formulation of a possible mechanism. The β-elimination reaction mechanism, as proposed by Gacesa (Gacesa 1987) requires neutralization of the negative charge of the carboxylic group of uronic acid in order to reduce the pKa of the C-5 hydrogen. In other polysaccharide lyases this is accomplished either by a Ca^{2+} ion ligand interacting with this group (chondroitinase B (Michel *et al.* 2004), pectate lyase (Scavetta *et al.* 1999), pectin lyase (Mayans *et al.* 1997)) or by the formation of a hydrogen bond between asparagine oxygen atom and the acidic group of the uronic acid that requires its protonation, thus neutralizing the negative charge (chondroitinase AC (Lunin *et al.* 2004)). HepII appears to utilize a strategy similar to that of chondroitinase AC. The carboxylic group of uronic acid makes a short contact (2.4 Å) with the sidechain of Glu205, which is flanked by Arg261 and Arg148 (Fig. 3a,c). Such a short contact indicates formation of a hydrogen bond, requiring protonation of one of the acidic groups, and is characteristic of a low energy barrier hydrogen bond, where the proton is shared by the two oxygen atoms (Cleland *et al.* 1998). The sidechain of Tyr257 is positioned in a similar manner to the catalytic tyrosine in chondroitin AC lyases, hyaluronate lyases and alginate lyases and most likely functions as a general base abstracting the C-5 proton. In the context of the present structure, the location of Tyr257 on one side of the uronic acid ring is consistent with proton abstraction from the glucuronic acid. For abstraction of the proton from the C-5 of iduronic acid this ring would have to be bound to the enzyme in an orientation that exposes the opposite side of the ring to Tyr257. Although this may be possible, we cannot assert this from our present

structure. However, the present structure indicates another possibility for proton abstraction from iduronic acid. The sidechain of His202 is located on the opposite side of the uronic acid ring to Tyr257 and its NE2 atom is ~4 Å away from C-5. Therefore this histidine could serve as a general base in case of iduronic acid containing substrate, without the need for a different binding mode. The previously shown preference of HepII for the cleavage of GlcA over IdoA (Rhomberg *et al.* 1998b) is in agreement with the structural view.

The best candidate for the role of general acid is Tyr257. This residue is appropriately positioned to donate a proton to the glycosidic oxygen bridging hexosamine and uronic acid, concomitant with breaking the C-4–O-4 bond and formation of C-4–C-5 double bond.

Thus, based on the structure of the HepII with a disaccharide product we propose that the elimination mechanism employed by HepII involves the following steps: (1) neutralization of the negative charge of the carboxylic group of the uronic acid by formation of a strong hydrogen bond with Glu205 that reduces the pKa of hydrogen bound to C-5; (2) proton abstraction by either Tyr257 (GlcA) or His202 (IdoA) acting as a general base; (3) proton donation by Tyr257 acting as a general acid to restore the OH functional group of the hexosamine product concomitant with breaking the C-4–O-4 bond and formation of C-4–C-5 double bond.

We have mutated Tyr257 to either Phe or Ala. The preliminary activity data show that both mutants are kinetically inactive relative to the wild type enzyme, but Y257F shows some minimal, but discernible activity in an exhaustive digestion assay. Further study of these and other mutants to delineate the catalytic mechanism is in progress.

3.6 Conclusions

We describe here the structure of HepII, the first HSGAG degrading enzyme, and its complex with a bound disaccharide degradation product. The structure shows that HepII has overall architecture similar to the enzymes from the PL8 lyase family but with a somewhat different relative disposition of domains, which leads to a significantly larger participation of the β-sheet domain in substrate binding and, possibly, catalysis. The shape of the substrate-binding site, located in a deep and long canyon formed at the

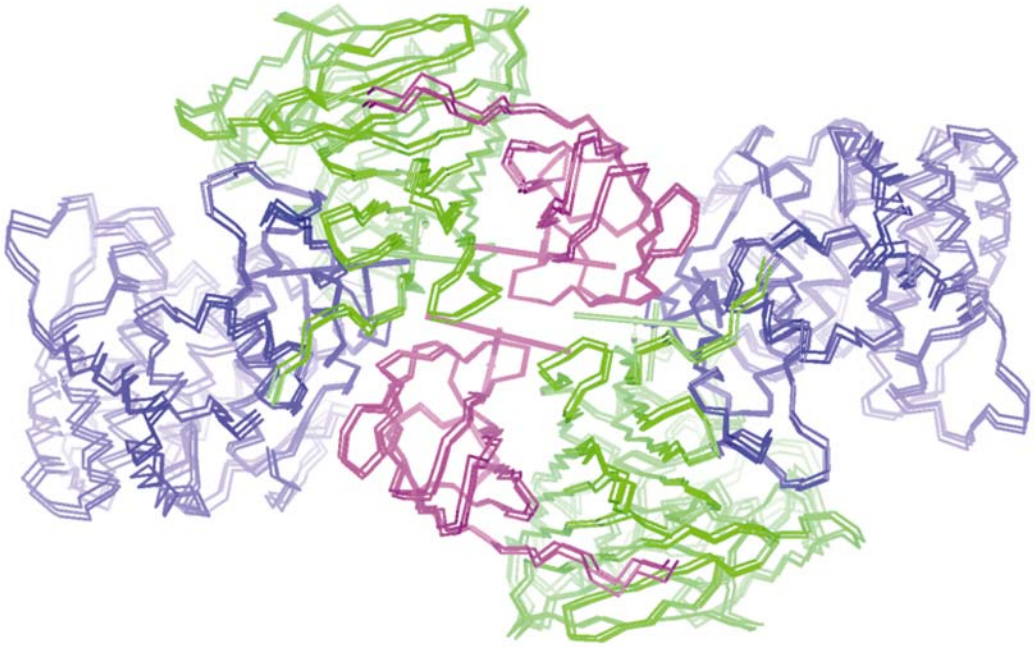
interface between the N-terminal and middle domain, suggests that the protein binds tightly at least 4-6 sugars of the GAG chain. Based on the structure with a bound unsaturated disaccharide reaction product we were able to identify residues forming the putative active site. While the present structure shows a potential role for Tyr257 in catalysis, the mechanism by which this enzyme is able to abstract the C-5 proton from either side of the uronic acid ring remains to be elucidated.

3.7 Acknowledgments

We would like to thank Anil Salgotra and Steve Wudyka for their valuable technical assistance in the purification of SeMet heparinase II. We also thank Drs. Allan Matte and Joseph D. Schrag for critical comments on the manuscript. Financial support from National Institutes of Health USA GM 057073 to RS is acknowledged.

3.8 Appendix A – Supplementary Material

Figure 3-S1. TLS components of thermal motions of the hep2 dimer
The large domain is colored blue, the middle domain is green and the C-terminal domain is magenta. The figure shows the average position and the distribution of the end points simulated by the program TLSviewer (Painter & Merritt 2005). Liberation axes are shown explicitly.



Chapter 4 Insight to the Catalytic Mechanism of Heparinase II by Site-Directed Mutagenesis and the Crystal Structure with its Substrate

Preface

The content presented in the following chapter is based on a manuscript submitted to publication as presented:

Shaya, D., Linhardt R., J., and Cygler, M. (2008) The Catalytic Mechanism of Heparinase II Revealed by Site-Directed Mutagenesis and the Crystal Structure with its substrate

4.2 Rationale

After obtaining the structures of HepII and its complex with the heparin degradation product, I continued my studies with this enzyme aiming at the to explaination of its activity against both HS and heparin. The first objective was to ascertain the hypothesis put forward in the previous chapter regarding the presence of one active site responsible for both HS and heparin binding and degradation. Upon obtaining the structure of the enzyme-HS product complex, I went on to obtain an enzyme-substrate complex to gain a deeper insight into the catalytic mechanism. In the context of this dissertation, this chapter concludes my studies of HepII and establishes a proposed catalytic scheme for its enzymatic activity, including assignments for the catalytic roles assumed by its active site residues.

4.3 Abstract

Heparinase II (HepII) is an 85-kDa dimeric enzyme depolymerizing both heparin and heparan sulfate glycosaminoglycans (GAGs) through a β-elimination mechanism. It possesses the striking ability to degrade the glycosidic bond between the hexauronic acid and hexosamine irrespective of the configuration at the C5 substituent of the uronic acid (either iduronic acid, IdoA, or glucuronic acid, GlcA). The β-elimination mechanism involves the neutralization of the negative charge on the uronic acid, followed by proton removal from carbon 5 (a chiral center) of the uronic acid. The two different hexauronic acid epimers present their C5 bound carboxyl group and hydrogen on opposite sides of the sugar ring, posing a topological challenge for the degradation of both substrates in a single active site.

Previously, we have determined the crystal structure of HepII from *Pedobacter heparinus* (previously known as *Flavobacterium heparinum*) in complex with heparin disaccharide degradation product, and determined the location of the enzyme's active site. Here we present the structure of HepII complexed with a heparan sulfate disaccharide degradation product, proving that a single binding/active site is responsible for the degradation of both uronic acid epimer-containing substrates.

We have identified a tyrosine and two histidine residues located at equal distances form C5, the chiral center of the reaction site, and these amino acids are located on both sides of the uronate product. The HepII double mutant, H202A/Y257A, showed complete lack of activity and we have crystallized this mutant with a heparan sulfate-derived tetrasaccharide and solved its structure. Based on all the structures and the kinetic evaluation of the active site mutants we propose that His202 acts as a general base abstracting the C5 proton during degradation of GlcA containing glycosidic bonds in heparan sulfate while Tyr257 in a deprotonated form abstracts a proton from IdoA containing glycosidic bonds in heparin. Furthermore, we propose that Tyr257 also acts as the general acid, donating proton to the glycosyl-leaving group at the final step of the β-elimination. Since the intermediate for both substrates looses its chirality by the formation of a planar double bond, we propose that Tyr257 can serve as a general acid for both substrates.

4.4 Introduction

Heparin and heparan sulfate (HS) glycosaminoglycans (GAGs) are negatively charged, linear polysaccharides, composed of repeating disaccharide units of uronic acid and glucosamine residues (GlcN, 2-amino-2-deoxy-α-D-glucopyranose) (Esko & Selleck 2002). Heparin typically contains ~90% iduronic acid (IdoA, α-L-idopyranosyluronic acid) and 10% glucuronic acid (GlcA β-D-glucopyranosyluronic acid), with a high content of 2-*O*-sulfo groups on the IdoA residue. The glucosamine residue in heparin is predominantly substituted with *N*-sulfo groups (GlcNS, where S is sulfo) with only a small number of *N*-acetyl groups and 6-*O*-sulfo groups and much less frequently with 3-*O*-sulfo groups. In contrast, HS is somewhat more diverse in its primary structure and characterized by higher percentage of the GlcA epimer, *N*-acetyl substituted GlcN (GlcNAc) and a lower percentage of 2-*O*-sulfo and *N*-sulfo groups. The modifications in HS are not uniform; rather, they are concentrated within specific regions of the polysaccharide, giving rise to short sequence motifs responsible for the interactions between HS and a diverse repertoire of proteins leading to its multiple biological roles. HS-GAGs are abundant at the cell surface as part of proteoglycan cell surface receptors (Kjellen & Lindahl 1991; Jackson *et al.* 1991). These complex polysaccharides provide

docking sites for numerous protein ligands involved in diverse biological processes, ranging from cancer and angiogenesis, anticoagulation, inflammatory processes, viral and microbial pathogenesis, and to multiple aspects of development (Linhardt *et al.* 1986; Jackson *et al.* 1991; Kjellen & Lindahl 1991; Shukla *et al.* 1999; Petitou *et al.* 1999; Perrimon & Bernfield 2000; Sugahara & Kitagawa 2002; Menozzi *et al.* 2002).

Specialized microorganisms express GAG-degrading lyases serving nutritional purposes of both themselves and their vertebrate hosts. The lyases depolymerize GAGs through a *β*-elimination mechanism, characterized by charge neutralization of the carboxylic acid group, and the release of unsaturated products (Gacesa 1987).

Heparinase II (heparin lyase II; heparitinase II, HepII; no assigned EC number) has the ability to cleave both heparin and HS-like regions of GAGs, regardless of their sulfation patterns (Nader *et al.* 1990). In general, the efficiency of enzymatic reactions relies on the stereo-specificity of the active site residues to fit the particular configuration of the substrate at the active center. Thus HepII is a remarkable enzyme in its ability to cleave the bond next to the uronic acid by abstracting the C5 bound proton of either epimer of the uronic acid (GlcA or IdoA).

The mature protein consists of 747 residues, starting with Glu26 (which is converted to pyroglutamate) and finishing in Arg772, has a molecular mass of 84,545 Da and a pI of 8.9 (Su *et al.* 1996). Presently the enzyme has no recognizable sequence homologs and therefore has not been assigned to a specific lyase family (CAZy data base, http://www.cazy.org/fam/acc_PL.html).

HepII acts in an endolytic manner (Rhomberg *et al.* 1998b) and displays broad substrate specificity, catalyzing the cleavage of linkages adjacent not only to IdoA and GlcA but also to the rare α-L-galacturonic acid residues present in the same GAGs (Moffat *et al.* 1991a). HepII will accept either GlcNAc or GlcNS and even the rare GlcN at the hexosamine-binding subsite, provided that the uronic acid contains a 2-*O*-sulfo group. Although HepII has greater affinity for heparin, its turnover rate for HS is higher (Lohse & Linhardt 1992) thus displaying a preference toward degradation of glycosidic bonds next to GlcA rather than IdoA. Moreover, HepII shows greater catalytic efficiency for longer (hexa-, octa, etc.) oligosaccharide substrates than for tetrasaccharides (Rhomberg *et al.* 1998b).

Previously, we have described the structure of the native HepII expressed in its original host *P. heparinus;* PDB accession code 2FUQ (Su *et al.* 1996; Shaya *et al.* 2004; Shaya *et al.* 2006) and of rHepII (expressed in *Eschericia coli*) complexed with a disaccharide product (PDB accession code 2FUT (Shaya *et al.* 2006b)).

HepII is the first heparin/HS degrading lyase for which the structure has been determined. The structure combined with solution studies has shown that the enzyme's biological unit is a dimer, with the two active sites located at opposite sides of the dimer and separated by ~80 Å (Shaya *et al.* 2006). Each monomer of HepII is comprised of an α+β lyase fold similar to the enzymes from the PL8 family (http://afmb.cnrs-mrs.fr/CAZY/). Briefly, each monomer contains three sub-domains: an $(\alpha/\alpha)_6$ incomplete toroid constitutes the N-terminal domain (residues PCA26 - Pro365), two extended β-sheets comprise the central β-sheet domain (residues Gly361 – His682), and 2 concave β-sheets in a β-barrel-like arrangement comprise the C-terminal β-sheet domain (residues Glu683 – Arg772).

There are marked differences in the relative disposition of the domains in HepII compared to the other PL8 family enzymes, which result in a significantly larger participation of the central β-sheet domain in substrate binding. In the PL8 family enzymes the active site is located at the top of the N-terminal α-helical toroidal domain and composed solely of residues belonging to this domain. HepII contains a Zn^{2+} ion bound at the interface between the N-terminal and central domains, which is hexa-coordinated by the sidechains of His409, Asp425, His451 and 3 water molecules. The Zn^{2+} is located 15 Å away from the heparin product. We proposed that this ion serves a structural role in maintaining the orientation between both domains and the proper conformation of the β-sheet loops to create one rim of the active site cleft (Shaya *et al.* 2006).

The HepII-disaccharide complex provided the first insight into the enzyme's active site and allowed us to speculate on the residues performing the *β*-elimination and their specific roles. Here we have continued investigating HepII catalytic mechanism by preparing the active site Y257F, Y257A, H202A and H406A single mutants and H202A/Y257A, H406A/Y257A and H202A/H406A double mutants. The mutants activity was evaluated in real-time by monitoring the formation of a double bond at 232

nm (Yamagata *et al.* 1968) while residual activity was measured after prolonged overnight incubation with excess of substrates and products analyzed by capillary electrophoresis separation. Selected inactive mutants were co-crystallized with heparin/heparan sulfate tetrasaccharides and the structures of these complexes were determined.

4.5 Materials and Methods

4.5.1 Preparation of the HS Tetrasaccharide Substrate

HS sodium salt was obtained from Celsus Laboratories (Cincinnati, OH). The HS was partially digested (90% reaction completion) with nHepII and fractionated by gel-permeation chromatography using a P-10 column (Bio-Rad, Hercules, CA) to obtain uniform sized oligosaccharides. The fraction containing tetrasaccharides was desalted on a Bio-Rad P-2 column and concentrated by freeze-drying. Fractions containing individual tetrasaccharides were collected from semi-preparative SAX-HPLC (Shimadzu, Columbia, MD) using a Spherisorb column (Waters Corp, Milford, MA), desalted on a Bio-Rad P-2 column, and freeze-dried (Pervin *et al.* 1995). The structure of each tetrasaccharides was determined by 1D and 2D proton NMR and electron detachment dissociation mass spectrometry (Wolff *et al.* 2007). The second tetrasaccharide eluting from the SAX-HPLC had the structure ΔUA (1→4) GlcNAc (1→4) GlcA (1→4) GlcNAc (where ΔUA is 4-deoxy-α - L-*threo*-hex-4-enopyranosyluronic acid).

4.5.2 Site-Directed Mutagenesis

The cloning of HepII was described earlier (Shaya *et al.* 2006b). HepII mutants were constructed using the QuickChange Site-Directed Mutagenesis Kit (Stratagene), according to the manufacturers' protocol, with appropriate DNA primers and the mutations were confirmed by DNA sequencing. The mutant proteins were expressed in *E. coli* BL21 (DE3).

4.5.3 Kinetic Evaluation of HepII Mutants

HepII lyase activity was measured in real time (over 5 min) based on the increase of absorbance at 232 nm resulting from the UV-absorbing, double-bond containing

products (Yamagata *et al.* 1968). Residual activity was evaluated using an exhaustive digestion of substrates in excess of rHepII mutants coupled with capillary electrophoresis for detection of reaction products (Venkataraman *et al.* 1999). Protein concentrations were determined using Bradford Protein Assay Reagent kit (Bio-Rad).

4.5.4 HepII Expression, Purification and Crystallization

Native HepII was purified in order to obtain the structure of wtHepaII-HS complex. Purification and crystallization of HepII from *P. heparinus* was carried out as described previously (Shaya *et al.* 2004; Shaya *et al.* 2006).

Expression, purification and crystallization of rHepII mutants Y257F, Y257A, H202A, H406A, H202A/Y257A, H406A/Y257A and H202A/H406A was preformed as previously described (Shaya *et al.* 2006).

4.5.5 Data Collection, Structure Determination and Refinement

Native HepII crystals were soaked briefly in reservoir solution supplemented with 20 % glycerol and 3 mM heparan sulfate tetrasaccharide substrate and flash frozen in N_2 stream at 100 ºK (Oxford Cryosystems, Oxford, UK). Diffraction data were collected on an HTC area detector (Rigaku) mounted on a Micromax 007 generator (Rigaku) with Osmic mirrors. The data were processed to 2.35 Å resolution using the program d*TREK® (Rigaku). Data collection statistics are summarized in Table 4-1.

rHepII mutant crystals were soaked in 25% (w/v) PEG 3350, 200mM ammonium acetate or 200 mM $MgCl_2$ or 200 mM Mg-formate, 12% (v/v) glycerol, supplemented with the 3 mM heparin or HS tetrasaccharide substrate for varying time periods ranging from ~ 30 sec to 20 min and flash-frozen in the nitrogen stream at 100 ºK (Oxford Cryosystems). Diffraction data extending to 2.1 Å resolution were collected at beamline X8C (NSLS, Brookhaven National Laboratory), using a Quantum-4 CCD area detector (ADSC, San Diego, CA) and processed using HKL2000 (Otwinowski & Minor 1997). Data collection statistics are summarized in Table 4-1.

The structures were determined by molecular replacement using the program MolRep (Vagin & Teplyakov 1997) with the PDB deposited coordinates of nHepII and rHepII as search models (PDB codes 2FUQ, 2FUT respectively (Shaya *et al.* 2006)). The

structures were refined using the program REFMAC5 (Murshadov, 1997). The refined structure of nHepII-HS product complex extends from PCA26 to Arg772 and contains 3 protein molecules in the asymmetric unit. The model also contains 3 Zn^{2+} ions, 511 water molecules, 10 phosphate molecules, 3 glycosylation sites containing 4-membered glycans and 3 disaccharide products bound to each monomer of HepII. The refined structure of rHepII H202A/Y257A-HS tetrasaccharide complex contains two molecules in the asymmetric unit, each comprised of Ala29-Arg772, 2 Zn^{2+} ions, 784 water molecules, 4 acetate molecules, and two tetrasaccharide molecules. The refinement statistics for both structures are summarized in Table 4-1.

4.5.6 Protein Data Bank Accession Codes

The coordinates and structure factors have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (**http://www.rcsb.org/**) as entries 3E80, 3E7J.

Table 4-1. X-ray crystallographic data

Data Collection		
Protein	**nHepII HS-ΔUAp**	**H202A/Y257A-tetra**
Space group	$P2_12_12$	$P2_1$
Cell dimensions:		
a (Å)	201.3	51.95
b (Å)	209.4	162.21
c (Å)	59.2	93.84
α (Å)	90.0	90.00
β (Å)	90.0	105.97
γ (Å)	90.0	90.00
Wavelengt (Å)	1.54	1.10
Resolution range (Å)	50.00-2.35 (2.43-2.35)	50.00-2.10 (2.16-2.10)
Observed/unique reflections[a]	477,225/98,168 (6,464)	230,689/72,010 (3,630)
Average redundancy	4.6 (3.8)	3.3 (1.7)
Completeness (%)	98.1 (90.0)	82.78 (36.10)
Rsym[b]	0.134 (0.302)	0.047 (0.353)
Average $I/\sigma(I)$	8.3 (4.4)	15.5 (1.50)
Refinment Statistics		
Rwork[c]	22.92 (36.10)	19.94 (32.10)
Rfree[c]	26.83 (38.00)	23.45 (35.00)
Number of atoms / average B-factor:		
Molecules in the ASU	3	2
Protein	17,934/24.02	11,876/49.80

Product/ Substrate	78/30.76	104/68.68
O-linked glycan	123/35.03	------
Ligand (phosphate/acetate)	45/43.24	20/64.61
Solvent	536/21.15	782/49.78
Ions (Zn^{+2})	3/26.43	2/48.81
Ramachandran plot:		
Allowed (%)	98.9	98.9
Generously allowed (%)	0.8	0.9
Disallowed (%)	0.2	0.2
r.m.s deviation from ideal values:		
Bonds (Å)	0.010	0.007
Angles (°)	1.560	0.966

[a] All measured reflections with $I>1\sigma$

[b] Rsym = $(\Sigma \mid I\text{obs} - I\text{avg}\mid)/ \Sigma I\text{avg}$

[c] Rwork/free = $(\Sigma \mid F\text{obs} - F\text{calc}\mid)/ \Sigma F\text{obs}$

4.6 Results and Discussion

4.6.1 Structure of nHepII-HS Disaccharide Degradation Product

The previously determined structure of HepII was complexed with product of heparin degradation and we postulated that both heparin and HS bind HepII in the same location with one active site performing degradation of both substrates (Shaya *et al.* 2006b). To validate this hypothesis we soaked native crystals of HepII with a representative HS tetrasaccharide (ΔUAp (1→4) GlcNAc (1→4) GlcA (1→4) GlcNAc), and determined the structure of this complex (Table 4-1, PDB accession code 3E80). The electron density map indicated the presence of only the disaccharide degradation product (ΔUA (1→4) GlcNAc) bound in "+1", "+2" subsites, similarly to what was observed for HepII co-crystallized with heparin ((Shaya *et al.* 2006) Fig. 4-2). No change in the protein structure was noticed between the two complexes and a structural overlay based on the superposition of only protein atoms, revealed both disaccharide products at the same position (Fig. 4-1). The disaccharides were found in a deep, elongated, positively charged cleft formed on the surface of HepII between the N-terminal and the central domains, with both sugar rings oriented parallel to the rims of the cleft and the *N*-acetyl/*N*-sulfo group pointing toward the bottom of the cleft. The binding site cleft is clearly divided into two parts, a narrow less deep half where the "+" subsites are located (towards the reducing end of the substrate (Davies *et al.* 1997)) and a wider deeper part in which the "–" subsites reside. The products occupy "+1", "+2" subsites, differentiating

HepII form other PL8 family lyases, where the disaccharide product bound tighter to the "–" subsites (Huang *et al.* 1999; Yoon *et al.* 2001). Elaborate hydrogen-bonding network maintains the position and correct orientation of sidechains lining the binding site. The disaccharide forms intimate contacts with the protein through hydrogen bonds, aromatic stacking and van der Waals interactions (Table 4-2).

4.6.2 HepII Mutant Analysis

Three amino acids were identified in the vicinity of the C5 atom of the uronic acid, the chiral center of the reaction, namely His202, Tyr257, and His406 (Fig. 4-2). Tyr257 and His406 stack parallel to each other on one side of the sugar ring (the direction of the C5 bound proton when subsite +1 is occupied by GlcA-HS substrate). This unique coupling of Tyr-His was found in the active site of ChonAC, ChonABC, hyaluronan lyase, and alginate lyase ((Li *et al.* 2000; Li & Jedrzejas 2001; Huang *et al.* 2003b; Lunin *et al.* 2004a; Yamasaki *et al.* 2005). These enzymes utilize the tyrosine side chain as a general base, in a deprotonated state, actively removing the C5 bound proton triggering the elimination. Hence we produced HepII mutants Y257F and Y257A. They showed no measurable activity in a real-time assay at 37°C over 5 min monitoring the absorbance at 232 nm. In hope of capturing the enzyme-substrate complex we have crystallized both mutants in the presence of tetrasaccharide substrate and solved their structures. Rather unexpectedly, the structures showed only the presence of a disaccharide degradation product bound in the "+" subsites. The mutant proteins were therefore analyzed for residual activity by overnight incubation with excess substrate at 37°C and the reaction mix was analyzed by capillary electrophoresis for detection of oligosaccharide products. This analysis confirmed that the single mutants retained residual activity. Similar analysis revealed that H202A single mutant afforded the same results, i.e. low residual activity on the overnight assay and only the disaccharide degradation product present in the active site of the soaked HepII crystals. To fully inactivate the enzyme we have produced and tested three double mutants H202A/Y257A, Y257A/H406A, and H202A/H406A. The double mutants rendered HepII inactive against heparin and HS. We have thus crystallized the H202A/Y257A mutant and soaked in the HS tetrasaccharide in order to capture the substrate in the HepII catalytic site. This structure showed indeed the

presence of intact tetrasaccharide in the electron density. The crystallographic findings are summarized in table 4-1.

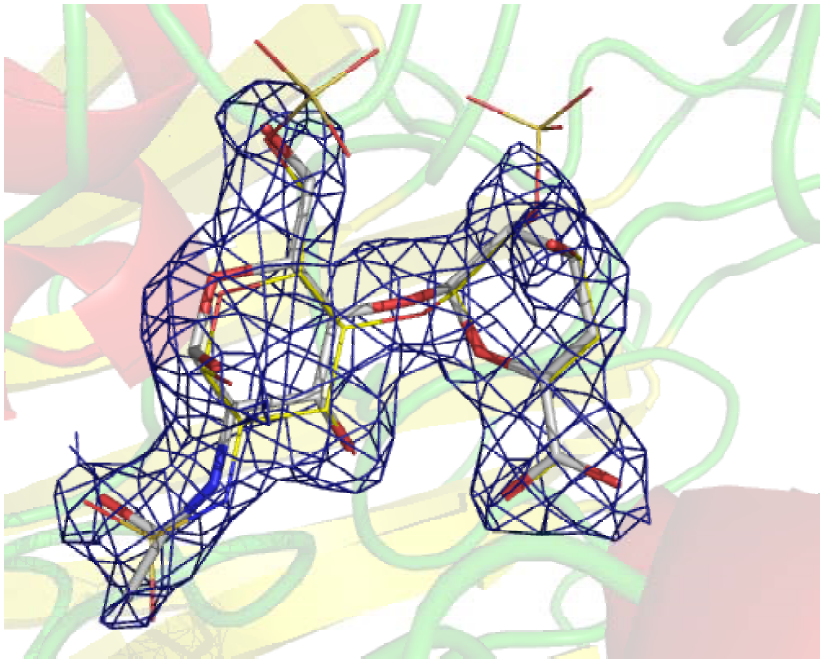
4.6.3 The Structure of rH202A/Y257A-HS Tetrasaccharide Complex

This structure allowed us to map the full substrate-binding site and identify four subsites from "–2" to "+2". The "–" subsites displayed fewer direct contacts between the protein and the oligosaccharide than the "+" subsites, representing weaker and less specific binding (Table 4-2). At the "+1" subsite the tetrasaccharide contains a GlcA rather than a ΔUA (present in the product), which is present in a different conformation and shows somewhat different binding to the protein. Thus, the carboxylic oxygen atoms in the GlcA of the substrate are hydrogen bonded to $His406^{NE}$ and $Tyr429^{OH}$ and through a water molecule to $Arg261^{NH2}$, whereas the carboxylic oxygen atoms of the product bind directly to $Glu205^{OE2}$ $His406^{NE}$ and $Arg261^{NH2}$. The GlcNAc in "–1" subsite binds with its acetyl group pointing toward the opening of the cleft as opposite to its binding in "+2" subsite. The N-acetyl nitrogen is bound through a water molecule to $Asp145^{OD2}$ and the oxygen of this group is hydrogen bonded to both $Arg96^{NH2}$ and $Arg96^{NH1}$, which also binds O3 of the GlcNAc. On the opposite side of the GlcNAc its O5 and O6 both are hydrogen bonded to $Arg148^{NH2}$. The GlcA at position "–2" binds the protein indirectly through water molecules. Thus O2 of GlcA is bridged with the backbone oxygen of Asp307, and O3 with the backbone oxygen of Ser256 and $Lys315^{NZ}$ through water molecules.

The bound tetrasaccharide displays a 90° kink between the sugar rings bound in the "–" and "+" sites (Fig. 4-3). A similar bending of the polysaccharide backbone around the glycosidic scissile bond occurs in the enzyme-substrate complex of chondroitin lyase AC from *A. aurescens* (Lunin *et al.* 2004). Furthermore, the structure of chondroitinase AC-substarte complex revealed that the uronic acid at the catalytic center attains a high-energy twisted boat conformation, contributing to the lowering of the reaction energy barrier through conformational strain. The replacement of Tyr257 by an alanine in the HepII double mutant H202A/Y257A created a void in the substrate-binding site. This void was occupied by the bound substrate, which projected the carboxylate group of GlcA in "+1" subsite into this space. The GlcA moiety could therefore bind in this

modified substrate-binding site in an energetically favourable chair conformation. Modeling this substrate conformation in the wild type HepII structure showed that the distance between Tyr257^{OH} and the carboxylate oxygen would be impossibly short, only 0.95 Å. Therefore, we propose that the substrate binding to the wild type enzyme invokes a conformational change in the GlcA in "+1" subsite from a chair conformation to a higher energy twisted boat conformation in a manner similar to that found in chondroitin AC and in other carbohydrate processing enzymes, e.g. lysozyme, (Strynadka & James 1991; Lunin *et al.* 2004).

Figure 4-1. Structure superposition of HepII-HS and HepII-heparin



4.6.4 Proposed Catalytic Mechanism

In the context of the described structures of HepII-product and substrate complexes we propose Glu205/His406 to be responsible for charge neutralization of the uronic acid. The neutralization of the negative charge on the carboxylic group at position "+1" is achieved by formation of a low energy barrier hydrogen bond with Glu205 that requires protonation of the carboxylate group. Glu205 is assisted by His406, which forms a hydrogen bond to the second oxygen atom of the carboxylate group. This charge neutralization in turn, reduces the p*Ka* of the proton bound to C5, which is abstracted by

a basic sidechain during the second step of the β-elimination. Tyr257 is proposed to serve dual functions: during depolymerization of HS, in a deprotonated form, it triggers the degradation, serving as the catalytic base (OH^{Tyr257}–$C5^{\Delta UA}$ distance 3.3 Å, Fig. 2-4) abstracting the C5 proton from GlcA (HS), following the elimination of the glycosidic bond the Tyr257 serves as a general acid, donating a proton to the non-reducing end of the GlcNAc leaving group, restoring the –OH functional group prior to product release.

Table 4-1. HepII HS product / substrate binding.
Distances between the disaccharide product, the tetrasaccharide substrate, and HepII (monomer A), ordered by binding subsites of HepII, the proposed catalytic side chains are boxed.

HepII molecule A	Product ΔUAp(1→4)GlcNAc	Distance Å	Substrate Tetrasaccharide	Distance Å
		"+2" subsite		
$ND2^{Asp-405}$	O1 sulfate at $N^{GlcNS6S}$	2.9	N-acetyl oxygen	3.3
$ND2^{Asp-405}$	$O3^{GlcNS6S}$	2.8	$O3^{GlcNAc}$	3.0
$N^{Gly-470}$	$O5^{GlcNS6S}$	3.2	$O5^{GlcNAc}$	3.0
$N^{Gly-470}$	O1 sulfate at $C6^{GlcNS6S}$	3.2	$O6^{GlcNAc}$	2.8
$ND2^{Asn-437}$	O1 sulfate at $C6^{GlcNS6S}$	2.9	$O6^{GlcNAc}$	3.0
$NE2^{His-202}$	Glycosidic oxygen (O4)	3.4	Glycosidic oxygen (O4)	2.9[a]
		"+1" subsite		
$OH^{Tyr-429}$	$O2^{\Delta UAp2S}$	3.0	O2-carboxyl^{glcA}	2.6
$N^{Tyr-436}$	O1sulfate at $C2^{\Delta UAp2S}$	3.0	---	---
$OD1^{Asp-145}$	$O3^{\Delta UAp2S}$	3.0	---	---
$OD2^{Asp-145}$	$O3^{\Delta UAp2S}$	2.9	$O3^{glcA}$	2.7
$NE2^{His-202}$	$O3^{\Delta UAp2S}$	3.3	---	---
$NE2^{His-406}$	O1 carboxyl$^{\Delta UAp2S}$	2.6	O1-carboxyl^{glcA}	3.0
$NH1^{Arg-261}$	O1 carboxyl$^{\Delta UAp2S}$	2.6	---	---
$OE2^{Glu-205}$	O2 carboxyl$^{\Delta UAp2S}$	2.4	O1-carboxyl^{glcA}	4.3
$NH2^{Arg-148}$	O2 carboxyl$^{\Delta UAp}$	3.0	O1-carboxyl^{glcA}	4.8
$OH^{Tyr-257}$	$C5^{\Delta UAp2S}$	3.8	$C5^{GlcA}$	3.3
$NE2^{His-406}$	$C5^{\Delta UAp2S}$	3.9	$C5^{GlcA}$	4.6
$NE2^{His-202}$	$C5^{\Delta UAp2S}$	4.2	$C5^{GlcA}$	5.0
		"–1" subsite		
$NH1^{Arg-96}$	---	---	Acetyl oxygen	3.2
$NH2^{Arg-96}$	---	---	Acetyl oxygen	3.0
$NH2^{Arg-96}$	---	---	$O3^{GlcNA}$	3.1
$NH2^{Arg-148}$	---	---	$O5^{GlcNA}$	3.1
$NH2^{Arg-148}$	---	---	$O6^{GlcNA}$	3.0
		"+2" subsite		
$NZ^{Lys-316}$	---	---	$O3^{\Delta UAp}$	3.2

[a] The distance between the glycosidic oxygen and $NE2^{His-202}$ is based on modeling the sidechain according to its position in the native structure.

During degradation of heparin, His202, located on the opposite side of C5 4.3 Å away from the chiral center, is proposed to serve as a general base that removes the proton from the IdoA epimer (Fig. 4-2). Since the intermediate formed during heparin degradation shares the same structure as that of HS (the chirality at C5 is lost after the formation of the planar C4-C5 double bond in the uronic acid), Tyr257 is proposed to serve as the general acid during degradation of both heparin and HS substrates.

Figure 4-2. Active site of HepII-HS degradation product complex

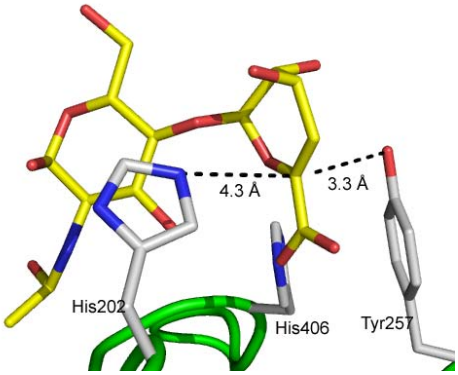
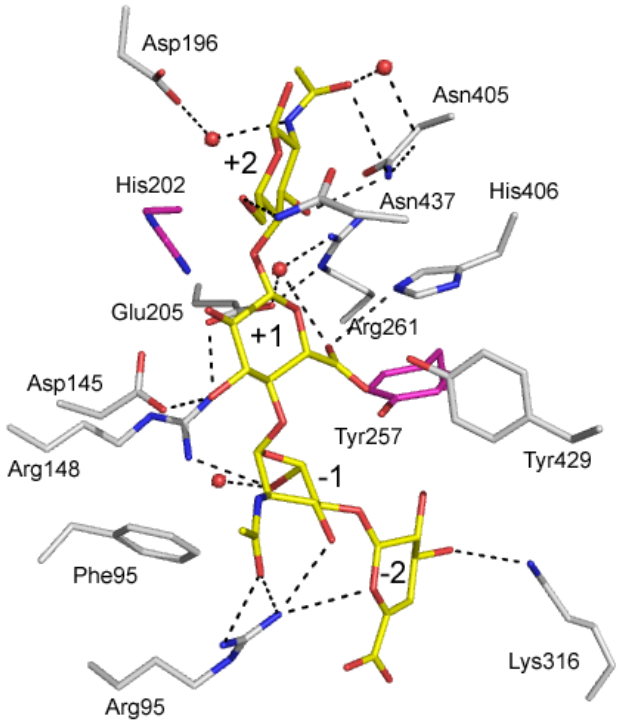


Figure 4-3. Active site of HepIIH202A/Y257A-HS tetrasaccharide
In white/blue/red sticks relevant active site residues, in yellow/blue/red sticks the tetrasaccharide substrate bound, in magenta His202 and Tyr257 modeled based on nHepII structure.



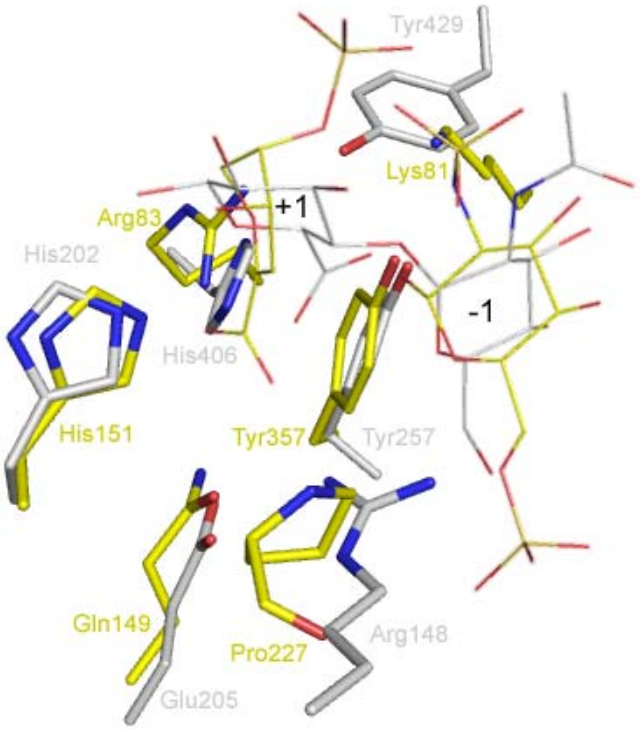
4.7 Conclusions

In this Chapter I have shown that HepII utilizes the same, single active site for the degradation of GlcA and IdoA containing substrates (HS and heparin, respectively). The identity of the sugar binding subsites "-1" and "-2" was revealed by crystallizing HepII H202A/Y257A double mutant with the HS representative tetrasaccharide. Furthermore, the structure of HepII H202A/Y257A complexed to HS tetrasaccharide revealed that the substrate is bound in a constrained manner, introducing a ~90 ° bend in the sugar backbone. We propose that in a native enzyme this bend is accompanied by a high-energy twisted boat conformation of the uronic acid (as shown in the structure of ChonAC) reducing the energy barrier for the glycosidic bond dissociation. Based on the structures presented in chapter 3 and this chapter we propose the following roles for active-site amino acids during degradation: Glu205 and His406 perform charge neutralization of the uronic acid, followed by proton abstraction from IdoA by H202 (heparin) and Y257 in a deprotonated state from GlcA (HS). Finally, Y257 acts as a general base for both substrates and the products are released.

4.8 Appendix A - Mechanistic Similarities between HepII and HepI

Recently, we solved the structures of nHepI from *B. thetaiotaomicron* and its complex with two disaccharide-degradation products (Garron, M., L., Shaya, D., and Cygler, M., unpublished results). The degradation products occupied the catalytic subsites "+2", "+1", "-1", and "-2" and allowed the comparison of the active sites of HepII with HepI. This comparison further validated the assignment of the catalytic roles for HepII active site residues proposed in this chapter. Structurally, both enzymes have in common solely the amino acid arrangement of their catalytic sites. HepI is a smaller protein (42.5 kDa) operating as a monomer in solution, which absolutely requires Ca^{2+} for its processive exolytic degradation (Sasisekharan *et al.* 1996; Rhomberg *et al.* 1998b). The enzyme folds into a β-jellyroll (resembling PL8 and 15 alginate lyses) as apposed to HepII. Nevertheless, a structural overlay based on the active residues revealed that both enzyme share the Tyr-His diad ($Tyr257^{HepII}$ overlays onto $Tyr357^{HepI}$, and $His202^{HepII}$ overlays on $His151^{HepI}$), supporting the roles assigned to His202 as the catalytic base, and Tyr257 as the catalytic acid during degradation of IdoA containing glycosidic bonds. $Glu205^{HepII}$ is substituted with $Gln149^{HepI}$. Interestingly, the catalytic site of ChonAC, degrading CS-A/C contains an asparagine at the similar location (interacting with the carboxylate of the GlcA at subsite "+1"), while its counterpart ChonABC, degrading both CS and DS contains an aspartate at that location. $Arg83^{HepI}$ substitutes $His406^{HepII}$ and occludes GlcA binding in subsite "–1", thus dictating the specificity of HepI toward IdoA containing glycosidic bonds.

Figure 4-4 Structural overlay of the active site of HepII and HepI
HepII active site residues shown in white/blue/red sticks, HepI active site residues shown in yellow/blue/red sticks, the sugars occupying "+1" / "–1" subsites are represented by semi-transparent thin lines.


Tyr429
Lys81
+1
Arg83
His202
-1
His406
His151
Tyr357
Tyr257
Gln149
Pro227
Arg148
Glu205

Chapter 5 Composite active site of chondroitin lyase ABC accepting both epimers of uronic acid

Preface

The content presented in the following chapter has been published as presented:

Shaya, D., Hahn B.S., Bjerkan T.M., Kim W.S., Park N.Y., Sim J.-S., Kim Y.S., and Cygler M. (2008) Composite active site of chondroitin lyase ABC accepting both epimers of uronic acid. *Glycobiology* **18**:270-277

Contribution of Authors:

I performed all the experiments and analyzed the data leading to the structure of chondroitin lyase ABC. I designed and constructed the active site mutants, proposed the kinetic experiments for characterization of these mutants and for defining the metal dependency of chondroitin lyase ABC activity. I have analyzed the data and wrote the draft of the manuscript. T.M.B. assisted in expression of selenomethionine labeled chondroitin lyase ABC. The kinetic characterization of chondroitin lyase ABC, its mutants and its dependence on metal cations, was proposed by me and carried out by W.S.K. and N.Y.P. and supervised by B.-S.H. Cloning of chondroitin lyase ABC was carried out by J.-S.S. under the supervision of Y.S.K. M.C supervised the project, orchestrated the collaboration, and prepared the final manuscript for publication with my aid.

5.1 Rationale

After establishing the manner by which HepII degrades GAGs containing both epimers of the uronic acid, I commenced my studies of chondroitinase ABC (ChonABC). ChonABC was shown to degrade both CS and DS, i.e. GAGs characterized by galactose as their hexosamine moiety and either GlcA (CS, DS) or IdoA (DS) as their uronic acid moiety. In the context of this dissertation, this chapter describes the structure elucidation of ChonABC and the initial analysis of its catalytic activity, efforts that continued in Chapter 6. Structurally, the dimension of the cleft harboring the active site of ChonABC was much wider than the known lyases belonging to PL8, which posed the question of the manner by which this enzyme maintains its substrate specificity and accomplishes catalysis. I addressed those questions by site-directed mutagenesis based on structural comparisons of this enzyme to *P. vulgaris* ChonABC, and *A. aurescens* ChonAC.

5.2 Abstract

Enzymes have evolved as catalysts with high degrees of stereo specificity. When both enantiomers are biologically important, enzymes with two different folds usually catalyze reactions with the individual enantiomers. In rare cases a single enzyme can process both enantiomers efficiently, but no molecular basis for such catalysis has been established. The family of bacterial chondroitin lyases ABC is comprised of such enzymes. They can degrade both chondroitin sulfate (CS) and dermatan sulfate (DS) glycosaminoglycans at the non-reducing end of either glucuronic acid (CS) or its epimer iduronic acid (DS) by a β-elimination mechanism, which commences with removal of the C-5 proton from the uronic acid. Two other structural folds evolved to perform these reactions in an epimer specific fashion: $(\alpha/\alpha)_5$ for CS (chondroitin lyases AC) and β-helix for DS (chondroitin lyases B); their catalytic mechanisms have been established at the molecular level. The structure of chondroitinase ABC from *Proteus vulgaris* showed surprising similarity to chondroitinase AC, including the presence of a Tyr-His-Glu-Arg catalytic tetrad, which provided possible mechanism for CS degradation but not for DS degradation. We determined the structure of a distantly related *Bacteroides thetaiotaomicron* chondroitinase ABC to identify additional structurally conserved

residues potentially involved in catalysis. We found a conserved cluster located ~12 Å from the catalytic tetrad. We demonstrate that a histidine in this cluster is essential for catalysis of DS but not CS. The enzyme utilizes a single substrate-binding site while having two partially overlapping active sites catalyzing the respective reactions. The spatial separation of the two sets of residues suggests a substrate-induced conformational change that brings all catalytically essential residues close together.

5.2 Introduction

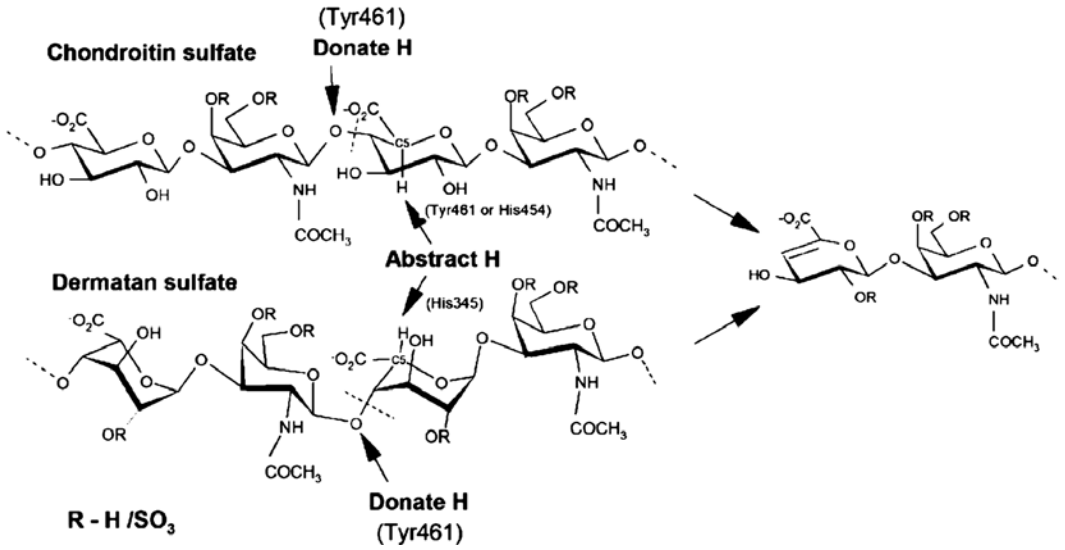
Enzymes acting on chiral substrates catalyze reactions with an extremely high degree of stereo specificity that is associated with preferential stabilization of one stereo isomeric transition state over another (reviewed in (Lamzin *et al.* 1995)). In many instances enzymes with different topologies have evolved to process two enantiomeric substrates. Some examples include L- and D- amino acid oxidases (Pawelek *et al.* 2000), 2-hydroxy acid dehydrogenases (Lamzin *et al.* 1994) and glycosaminoglycan (GAG) lyases (Ernst *et al.* 1995). The active sites of such enzyme pairs can show a near mirror symmetry embedded in dissimilar folds (Lamzin *et al.* 1994; Mattevi *et al.* 1996). A different category of enzymes, racemases and epimerases, evolved to catalyze inversions of stereochemistry (reviewed in (Tanner 2002)), and there is a good understanding of their catalytic mechanism (Allard *et al.* 2001; Tanner 2002). In rare cases an enzyme possesses the ability to accept substrates with either configuration at the active center and convert them into new molecules. For example, acetolacetate decarboxylase possesses the unique ability to process both enantiomers of acetolacetate yielding a single decarboxylation product enantiomer (Najmudin *et al.* 2003). Heparin lyase II (Shaya *et al.* 2006) and chondroitin lyase ABC can depolymerize specific glycosaminoglycans by abstracting C-5 proton from either side of their uronic acid building block (Huang *et al.* 2003). However, the mechanistic and structural details by which these enzymes perform these tasks are presently not known. Here we address these questions for the uronic acid, epimer-nonspecific glycosaminoglycan lyases.

Chondroitin sulfate (CS) and dermatan sulfate (DS), one of the four classes of glycosaminoglycans, are linear, heterogeneous, highly negatively charged polysaccharides composed of $\beta(1\rightarrow4)$ linked disaccharide repeating units containing

uronic acid (1→3) linked to *N*-acetyl-D-galactosamine (GalNAc) (Ernst *et al.* 1995). The uronic acid moiety in CS is exclusively β-D-glucuronic acid (GlcA), whereas DS contains a mixture of α-L-iduronic acid (IdoA) and GlcA epimers (Fig. 5-1). GAGs linked *via* specific serine residues to core proteins form proteoglycans (PGs), which are the major components of the extracellular matrix of vertebrates. PGs modulate fundamental biological processes such as cell adhesion, proliferation, differentiation, signaling, inflammation, and infection (Folkman & Klagsbrun 1987; D'Amore 1990; Vlodavsky *et al.* 1991; Varki 1993; Bernfield *et al.* 1999; Bao *et al.* 2004).

Bacterial GAG-degrading lyases depolymerize their substrates through a β-elimination mechanism characterized by charge neutralization of the carboxylic acid group, removal of a relatively acidic proton from the C-5 carbon (chiral center) of the uronic acid, and the release of the 4-linked hexosamine with the generation of a C-4–C-5 double bond at the uronic acid ring (Gacesa 1987) (Fig. 5-1).

Figure 5-1. Lytic degradation of CS and DS by ChonABC.
Dashed line between the galactose and uronic acid indicates the bond that is being broken.



Enzymes with two distinct structural topologies have evolved to sterospecifically degrade either CS or DS. Chondroitin lyase AC (chondroitinase AC, ChonAC), which specifically depolymerizes the glucuronic acid-containing CS (Fig. 5-1, top), has a catalytic domain characterized by an incomplete $(\alpha/\alpha)_5$ toroid fold (Féthière *et al.* 1999). The crystal

structures of chondroitin lyase AC from *Pedobacter heparinus* (PedAC) and from *Arthrobacter aurescens* (ArthroAC) complexed with inhibitors and substrates (Huang *et al.* 2001; Lunin *et al.* 2004), combined with mutagenesis (Capila *et al.* 2002; Rye *et al.* 2006) showed that the enzyme utilizes an asparagine to neutralize the acidic group of GlcA (low-barrier hydrogen bond) and a Tyr-His-Glu-Arg hydrogen-bonded tetrad as a catalytic center. The role of the general base removing the glucuronic acid C-5 proton was assigned to a tyrosine in a deprotonated state (Lunin *et al.* 2004). The structural and biochemical data (Rye *et al.* 2006) argued for the same tyrosine to protonate also the leaving group, acting as a general acid. In contrast, chondroitin lyase B (chondroitinase B, ChonB), which degrades exclusively iduronic acid containing regions of DS (Fig. 5-1, bottom), has a right-handed β-helix fold (Huang *et al.* 1999) similar to that of pectate and pectin lyases (Herron *et al.* 2003). This enzyme utilizes a Ca^{2+} ion coordinated by two glutamates and one asparagine to neutralize the acidic group of IdoA, whereas its catalytic center contains a lysine acting as a base abstracting the C-5 proton and an arginine performing the role of a general acid during catalysis (Michel *et al.* 2004).

In addition to these C-5 epimer-specific enzymes, another class of enzymes designated chondroitin lyases ABC (chondroitinase ABC, ChonABC) possesses the ability to degrade both CS and DS. Two isomers, PvulABCI and PvulABCII were identified in the bacterium *Proteus vulgaris*, and biochemically characterized (Hamai *et al.* 1997). Their sequences show no recognizable similarity to either ChonAC or ChonB. Nevertheless, the structure of PvulABCI revealed a very similar fold to ChonAC, supplemented with an additional N-terminal domain (Huang *et al.* 2003). The similarity extends to the active site: PvulABCI contains an identical Tyr-His-Glu-Arg tetrad arrangement as that found in ChonAC; their involvement in catalysis was confirmed by mutagenesis (Prabhakar *et al.* 2005). Thus, the unusual ability of ChonABC to abstract the proton from either side of the uronic acid sugar ring presented us with a conundrum: the degradation of CS substrates could occur in the same manner as in ChonAC, whereas the structure provided no clear indications how the same tetrad could be involved in abstracting the C-5 proton from the other side of the uronic acid ring during DS degradation.

The most straightforward way to identify catalytically essential residues is to determine the structure of the enzyme-substrate/inhibitor complex. Despite substantial efforts using multifaceted crystallization approaches we were not able to obtain crystals of such complexes for ChonABC. Therefore, we decided to use a different structural approach to identify other catalytically important residues in addition to the tetrad. We searched for a distantly related member of the ChonABC family aiming to determine its structure and, by comparison with PvulABCI, to identify structurally conserved constellations of residues in the proximity to the catalytic tetrad that could potentially be involved in catalysis of DS substrates. One such enzyme (BactnABC, 1014 amino acid protein) with similar specificity as PvulABCI, was identified in *Bacteroides thetaiotaomicron*, a prominent member of the human gut microbiota, containing a large contingent of proteins responsible for breaking down of the otherwise indigestible dietary oligosaccharides (Backhed *et al.* 2005), and has been partially characterized (Linn *et al.* 1983). BactnABC is only distantly related to PvulABCI (23% sequence identity) hence we decided to investigate its crystal structure.

Comparing the structures of PvulABCI and BactnABC we identified two conserved histidines and an arginine in the substrate binding cleft ~12 Å from the putative catalytic tetrad. Using site-directed mutagenesis, we demonstrate that these histidines are essential only in DS degradation, whereas the arginine is required for catalysis of all substrates. Moreover, we demonstrate the involvement of the divalent metal cations Ca^{2+}/Mg^{2+} in the catalysis of DS substrates. Based on these observations we hypothesize that members of the ChonABC family contain a composite active site consisting of two sets of partially overlapping residues that are involved in the cleavage of CS and DS, respectively.

5.3 Results and Discussion

5.3.1 The Three-Dimensional Structure of BactnABC

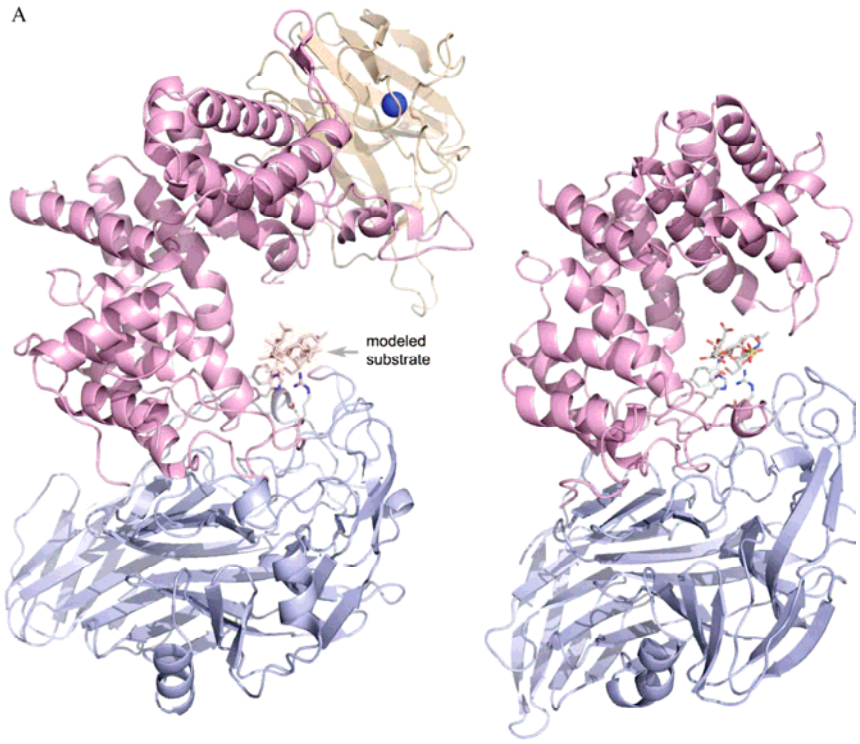
Remarkably, despite low sequence identity, the structure of BactnABC is highly similar to that of PvulABCI (Fig. 5-S1, supplementary data). A structure-based sequence alignment shows a sequence identity of only 23% for the two proteins. The root-means-squares deviation (rmsd) of Cα atoms for individual domains is in the range 1.1-1.45 Å

and the overall superposition shows a small difference in orientation of the N-terminal domains relative the rest of the molecules, suggesting some inter-domain flexibility.

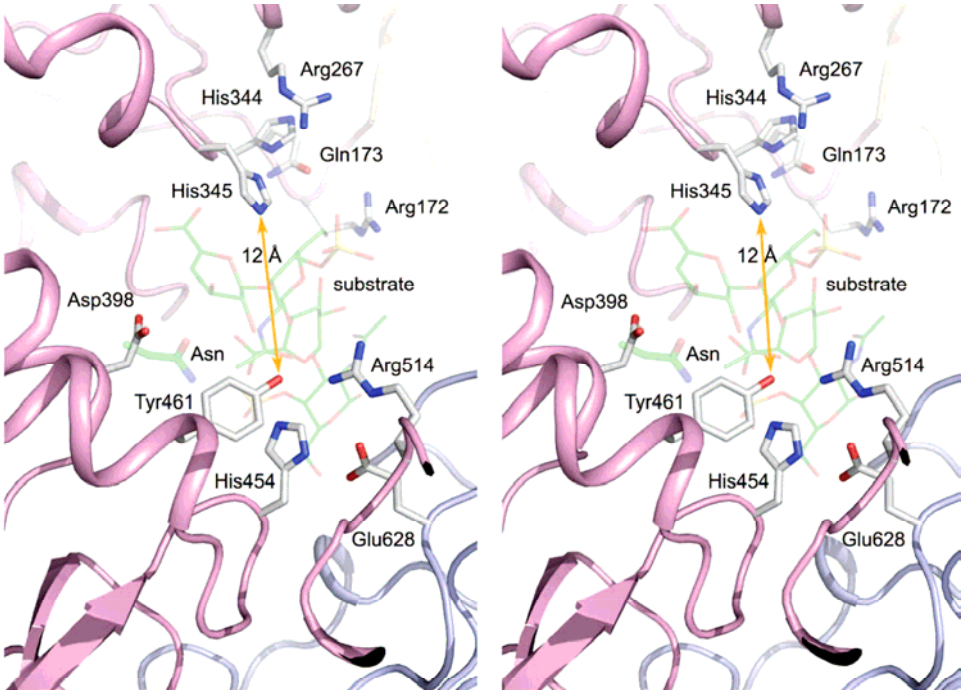
Briefly, the N-terminal domain (residues Ala14-Asp170) contains 10 antiparallel β-strands arranged in a two-layered jellyroll β-sheet fold (Fig. 2a) common to carbohydrate binding modules found in such enzymes like endo 1,4-α-D-xylanases (Charnock *et al.* 2000) and cellulase (Carvalho *et al.* 2004). Such carbohydrate-binding proteins contain a structural Ca^{2+} ion located between the second and third β-strand. We have identified a Ca^{2+} ion in the equivalent location in BactnABC, hepta-coordinated by Ser 24, Glu26, Asp50, Ser56 and Asp161, (2.3-2.7 Å distances). The central domain (residues Ala171–Gln593) has a double-layered $(\alpha/\alpha)_5$ incomplete toroid fold. Despite no recognizable sequence similarity, this domain is structurally similar to the catalytic domains of ChonAC from *P. heparinus* and *A. aurescens* (Lunin *et al.* 2004), hyaluronate lyases(Li *et al.* 2000), xanthan lyase (Hashimoto *et al.* 2003), alginate lyases (Yoon *et al.* 1999), and *P. heparinus* heparinase II (Shaya *et al.* 2006). The noticeable difference between BactnABC and the other proteins is a much more open substrate binding cleft in the former (Fig. 5-2a). The C-terminal domain (residues Gly594–Lys1014) contains a stack of four antiparallel β-sheets and contributes Glu628 to the catalytic tetrad.

Structural alignment of PedAC, ArthroAC PvulABCI and BactnABC shows that the Tyr-His-Arg-Glu catalytic tetrads of ChonAC enzymes (His225, Tyr234 Arg288 and Glu653 in PedAC (Huang *et al.* 2001)) have counterparts in BactnABC (His454, Tyr461 Arg514 and Glu628, supplementary Fig. 5-S2) and in PvulABCI. Furthermore, both BactnABC and PvulABCI have wide substrate binding clefts. On the rim of the substrate binding cleft opposite to the tetrad we have identified three residues conserved between both ChonABC enzymes, His 344, His 345 and Arg172, capable of participating in catalysis or substrate binding. In this region we also found two other conservatively replaced residues that could be involved in substrate binding, namely Gln173 and Arg267 in BactnABC (Tyr, Lys, respectively, in PvulABCI). Contribution of all these residues to the catalyzed reaction was investigated by site-directed mutagenesis.

Figure 5-2. The structure of BactnABC.
A. Left: a cartoon representation of a BactnABC monomer. Colors: N-terminal domain – wheat, middle domain – magenta, C-terminal domain – light blue, Ca2+ ion – blue sphere; the tetrad sidechains are shown in stick mode. The oligosaccharide substrate is modeled based on PedAC. Right: cartoon representation of PedAC with a bound tetrasaccharide IdoA (1→3) GalNAc4S (1→4) IdoA (1→3) GalNAc4S (where GalNAc4S stands for N-acetyl-D-galactosamine sulfated at carbon 4 and IdoA for α-L-iduronic acid). The substrate binding cleft in BactnABC is much more open than that in PedAC.



B. Stereo view of the substrate-binding site of BactnABC with the sidechains of the tetrad and residues identified in this work shown as sticks. The modeled position of the substrate tetrasaccharide IdoA(1→3)GalNAc4S(1→4)IdoA(1→3) GalNAc4S is shown by thin lines. The location of the neutralizing asparagine in the structure of PedAC relative to the tetrad and substrate is marked with Asn. The distance between the residues of the two clusters is greater than 12 A°. Figures were created with the program PyMol (http://www.pymol.org).



5.3.2 Ca^{2+} and Mg^{2+} Dependence of BactnABC Activity

Of the two lyases specific for either CS or DS, the activity of chondroitinase AC is independent of metal ions (Huang et al. 2001), whereas that of chondroitinase B requires Ca^{2+} utilized for charge neutralization of the acidic group of iduronic acid (Michel *et al.* 2004). The effect of divalent cations on the activity of the recombinant C-5 epimer-non-specific BactnABC is therefore of importance in understanding the mechanisms involved in catalysis and substrate recognition. We tested the effects of Ca^{2+} and Mg^{2+} in Tris-HCl buffer at 37°C, the temperature of maximal activity, and pH 7.6 optimum for this enzyme. The addition of these ions had a dramatically different effect on activity toward CS (glucuronic acid) and DS (iduronic acid). The specific activity for CS substrates increased only moderately (~2 fold) while an increase of ~25-fold was

observed for DS (Table 5-2). This suggests that the metal ions play a key role in charge neutralization of iduronic acid (as found for ChonB, (Michel et al. 2004) but are less important for the glucuronic acid epimer (as seen in ChonAC, (Pojasek *et al.* 2001; Lunin *et al.* 2004). No increase in activity was observed for Mn^{2+} and Zn^{2+} (data not shown). Addition of 10 mM EDTA abolished the activity toward either substrate indicating that the presence of divalent cations is essential for BactnABC activity. No indication of a large structural change was found in the CD spectra upon the addition of EDTA (data not shown) making it unlikely that the lack of activity was caused by the removal of the structural Ca^{2+} from the N-terminal domain affecting indirectly the substrate-binding site.

Table 5-1. X-ray crystallographic data

Data collection	
Space group	P63
Cell dimensions (Å)	a = 223.6, c = 112.6
Wavelength (Å)	0.97233
Resolution range (Å)	44.62-2.85 (2.92-2.85)
Observed/unique reflections[a]	419,435/139,287
Average redundancy	3.1 (1.7)
Completeness (%)	94.8 (59.4)
*R*sym[b]	0.063 (0.400)
Averaged $I/\sigma(I)$	20.6 (1.6)
Refinement statistics	
No. of reflections[c]	72,509 (3,102)
R-work[d]	0.220 (0.413)
R-free[d]	0.262 (0.464)
No. of atoms/average *B*-factor	
Protein	15,824/59.3
Solvent	67/40.7
Ions (Ca^{2+})	2/62.1
PO_4^{2-}	16*5 = 80/73.3
Ramachandran plot	

Allowed (%)	99.2
Generously allowed (%)	0.6
Disallowed (%)	0.2
Root-mean-square deviation	
Bond (Å	0.02
Angles (°)	1.9

[a] The Friedel pairs were not merged

[b] *R*sym = (Σ|*I obs* − *I avg*|)/Σ*Iavg*

[c] Reflections used for refinement with merged Friedel pairs

[d] R-work/free = (Σ|Fo − Fc|)/ΣFo

Table 5-2. Ca^{2+}/Mg^{2+} influence upon the specific activity of BactnABC.
All experiments were done in triplicate. The substrate was dissolved in 50 mM Tris–HCl, pH 7.6, and the reaction was monitored at 232 nm over 300 s at 37°C. One microgram of BactnABC was used in the experiment. The CS-A used is from BioCorp while DS and CS-C are from Sigma.

	Specific activity (U mg –1 protein)		
Metal ion	**CS-A**	**CS-C**	**DS**
50 mM tris-HCl	12.1± 0.4	6.8 ± 0.4	0.2 ± 0.1
+10 mM EDTA	n.d[a]	n.d[a]	n.d[a]
+10 mM $CaCl_2$	25.3 ± 0.4	13.6 ± 0.4	4.3 ± 0.1
+20 mM $CaCl_2$	29.6 ± 0.4	12.6 ± 0.4	2.5 ± 0.1
+10 mM $MgCl_2$	38.2 ± 0.4	19.1 ± 0.4	5.2 ± 1.8

[a] n.d – not detected

5.3.3 Two Overlapping Catalytic Sites Reside Within a Single Substrate-Binding Site

The β-elimination reaction mechanism requires neutralization of the negative charge of the uronic acid carboxylic group to reduce the pKa of the proton bound to C-5 (Gacesa 1987). Abstraction of this proton by a general-base-acting residue triggers the formation of a C-4–C-5 unsaturated bond in the uronic acid, accompanied by the

elimination of the 4-linked hexosamine assisted by a general acid donating a proton to reconstitute the –OH of the leaving group.

The C-5 epimer-specific enzymes ChonAC and ChonB adopt completely different structures for stereo selective degradation of their respective substrates. ChonAC, with the $(\alpha/\alpha)_5$ helical toroid catalytic domain, neutralizes the acidic group of the glucoronate by two hydrogen bonds to an asparagine sidechain requiring its protonation, and an additional hydrogen bond to a protonated histidine (Lunin *et al.* 2004). ChonB, on the other hand has a right-handed β-helical fold and utilizes a Ca^{2+} ion coordinated between the protein sidechains and the acidic group of the iduronate for its neutralization (Michel *et al.* 2004). The residues playing the roles of general base and general acid also differ between these enzymes. Thus, ChonAC uses a tyrosine for both functions while ChonB utilizes a rather unusual Lys/Arg pair. A common feature, found in complexes of both these enzymes with their substrates, is that the sugar ring of the uronic acid assumes a higher energy twisted conformation upon binding, making the degradation of the glycosidic bond an energetically more favorable event (Michel *et al.* 2004; Lunin *et al.* 2004).

BactnABC (and PvulABCI), which combines the activities of both of these enzymes, shares structural similarities with ChonACs. However, it contains an additional domain at its N-terminus and its substrate binding cleft, within the central $(\alpha/\alpha)_5$ domain, is significantly wider than the corresponding cleft in ArthroAC and PedAC (Fig. 5-2). While the Tyr-His-Arg-Glu tetrad is structurally conserved, the tryptophan residues stacking against the galactosamine rings of the substrate in ChonACs have no counterparts in ChonABCs. Moreover, the asparagine interacting with the acidic group of the glucoronate in ChonACs is substituted in ChonABCs by an aspartate (Asp398 in BactnABC) tilted away from the proposed binding site of the substrate.

The conservation of the tetrad residues and the effects of their mutation on PvulABCI activity (Prabhakar *et al.* 2005) strongly indicate their involvement in catalysis. As expected, replacing these residues in BactnABC by alanines led to inactivation of the enzyme against all substrates (Table 5-3). Thus the presence of the tetrad would explain the cleavage of CS as described for ChonAC (Lunin *et al.* 2004) but not DS, which is indeed an inhibitor of ChonAC. As well, the manner by which the

negative charge of the acidic group of uronic acid is neutralized, in ChonABC, remains an open question.

Surprisingly, our results clearly show that while the activity toward CS increases only slightly in the presence of divalent metal cations, DS degradation is substantially enhanced in their presence (Table 5-2). We postulate that a divalent metal cation (either Ca^{2+} or Mg^{2+}) binds together with the substrate coordinating the uronate acidic group, possibly assisted by Asp398, and participates particularly in the charge neutralization of the iduronate during DS degradation, paralleling ChonB (Michel *et al.* 2004).

The residues participating in DS degradation (iduronic acid epimer) remained an open question. Other than the strong effect of Ca^{2+} on the activity toward DS, no other similarities between BactnABC and ChonB are evident. Therefore, we compared the structures of BactnABC and PvulABCI in search for other structurally conserved residues in the vicinity of the substrate-binding site that might be involved in catalysis. As described above, we identified His344, His345 and Arg172 as well as two conservative replacements, Arg267 and Gln173 located within the substrate-binding cleft ~12 Å from the tetrad (Fig. 5-2b). These five residues were mutated individually to alanines and the activities of the mutants toward CS and DS were measured (Table 5-3). The expression levels of these mutants and their behaviour during purification were similar to that of the wild type enzyme.

The H344A and H345A mutants show no detectable activity toward DS in the standard assay but still degrade CS, albeit with 10-30 fold lower k_{cat}/K_M (Table 5-3). Overnight incubating these mutants with DS showed traces of disaccharide degradation products on HPLC analysis for the H344A mutant indicating residual activity, whereas no disaccharides were observed for the H345A mutant showing that it is indeed inactive (data not shown). His345 is thus the best candidate for the role of general base for DS substrates. The R172A showed no activity with any substrate, neither in the standard assay (Table 5-3) nor during the overnight incubation. The two other mutations, Q173A and R267A, reduced the activity toward all substrates by a factor of ~2-10, suggesting that they function in substrate binding or in maintaining the integrity of the active site.

Table 5-3. Kinetic analysis of BactnABC and its mutants.
Mutants of the tetrad residues are boxed. Mutants of the basic amino acids identified in the second site are marked in grey.

	CS-A[b]			CS-C [d]			DS [e]		
Enzyme	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M (μM-1 min-1)	K_M (μM)	kcat (min-1)	k_{cat}/K_M (μM-1 min-1)	K_M (μM)	k_{cat} (min-1)	k_{cat}/K_M (μM-1 min-1)
WT	67±1	15,792±240	**235.7**	33±3	10,404±38	**315.27**	61±8	2,307±190	**37.8**
H454A	n.d.[c]	n.d.[c]	**n.d.[c]**	n.d.[c]	n.d.[c]	**n.d.[c]**	n.d.[c]	n.d.[c]	**n.d.[c]**
Y461F	n.d.[c]	n.d.[c]	**n.d.[c]**	n.d.[c]	n.d.[c]	**n.d.[c]**	n.d.[c]	n.d.[c]	**n.d.[c]**
R514A	n.d.[c]	n.d.[c]	**n.d.[c]**	n.d.[c]	n.d.[c]	**n.d.[c]**	n.d.[c]	n.d.[c]	**n.d.[c]**
E628A	n.d.[c]	n.d.[c]	**n.d.[c]**	n.d.[c]	n.d.[c]	**n.d.[c]**	n.d.[c]	n.d.[c]	**n.d.[c]**
R172A	n.d.[c]	n.d.[c]	**n.d.[c]**	n.d.[c]	n.d.[c]	**n.d.[c]**	n.d.[c]	n.d.[c]	**n.d.[c]**
Q173A	102±5	7,389±164	**72.4**	50±2	1,984±5	**39.7**	71±1	408±7	**5.7**
R267A	109±5	12,320±80	**113.0**	98±6	3,706±61	**37.8**	71±2	250±14	**3.5**
H344A	202±20	5,560±28	**27.5**	134±2	2,143±172	**16.0**	n.d.[c]	n.d.[c]	**n.d.[c]**
H345A	171±13	3,828±40	**22.4**	195±19	1,852±129	**9.5**	n.d.[c]	n.d.[c]	**n.d.[c]**
H453A	38±8	1,057±62	**27.8**	20±0.5	385±14	**19.3**	44±4	352±0.2	**8.0**

[a] Results are the mean ± S.D. for at least three experiments.
[b] 1 mg CS-A from porcine trachea (Biocorp) in 50 mM K-phosphate buffer pH 7.6, measured for 300 sec at 37°Cusing 0.5 μg of enzyme
[c] n.d. Not detected
[d] 1 mg Cs-C from shark cartilage (Sigma) in 50 mM K-phosphate buffer pH 7.6, measured for 300 sec at 37°C using 1.0 μg of enzyme
[e] 1 mg DS from porcine intestinal mucosa (Sigma) in 50 mM K-phosphate buffer pH 7.6, measured for 300 sec at 37°C using 1.5 μg of enzyme

The data recorded for the mutants show that the tetrad amino acids and the distal Arg172 are essential for catalysis of all substrates while the distal histidines are essential only for DS degradation. Thus, we suggest that the general base abstracting the C-5 proton from the glucuronic acid of CS is Tyr461 (as in the case of ChonAC), or the nearby His454 (a tetrad component) (Fig. 5-1), while the role of a general base for DS is performed by His345. The most likely general acid residue donating the proton to the bridging oxygen is Tyr461 (as suggested for ChonAC). This residue would play the same

role for both epimers, since the intermediate state of each epimer substrate after removal of the C-5 proton is the same. We hypothesize that Arg172 is involved in substrate binding in subsite "+2" (GlcNAc), playing analogous role to the tryptophanes in ChonAC, rather than directly participating in catalysis. Furthermore, the fact that both the tetrad and the cluster of newly identified residues affect all substrates argues for the presence of a single substrate binding site with a composite active site made of partially overlapping amino acids specific either for CS or DS degradation. The formation of such a composite active site can only be accomplished by narrowing the substrate-binding cleft in the presence of the substrate. Thus, combining a structural comparison of distantly related enzymes from the same family with mutagenesis of structurally conserved residues and their kinetic characterization led to uncovering of novel residues essential for catalysis and their different roles in processing the two epimers of uronic acid.

5.4 Materials and Methods

5.4.1 Bacterial Growth

The *Bacteroides thetaiotaomicron* strain WAL 2926 was obtained from the Korean Culture Center of Microorganisms (KCCM) and cultivated at 37^{o}C in nutrient broth. *E. coli* BL21(DE3) (Novagen) and DL41(DE3) were maintained at 37^{o}C in Luria broth supplemented with 100 μg ml^{-1} ampicillin (Ap) when appropriate.

5.4.2 Standard Methods

Standard recombinant DNA procedures were performed according to Sambrook and Russel (Sambrook & Russell 2001). Genomic DNA of *B. thetaiotaomicron* and plasmid DNA were isolated using the G-spin Genomic DNA Extraction kit (iNtRON, Korea) and the QIAprep Spin Miniprep Kit (Qiagen) respectively, according to the manufacturers' procedures. Transformation of *E. coli* was performed according to the Novagene protocol. DNA amplifications by PCR were done using Taq polymerase, while BactnABC mutants were constructed using the QuickChange Site-Directed Mutagenesis Kit (Stratagene) and the primers listed in supplementary Table S1. Mutations were verified by DNA sequencing.

5.4.3 Cloning of the chondroitin ABC lyase gene from *Bacteroides* genome

The gene coding for BactnABC was amplified by PCR from *B. thetaiotaomicron* WAL 2926 genomic DNA using primers constructed according to the published sequence BT3324 (Xu *et al.* 2003). The PCR fragment was subcloned into the pET-22b vector. DNA sequencing revealed 64 nucleotide differences with the published sequence (gi:29348733, (Xu et al. 2003)) translating into 16 amino acid differences (Supplementary Table S2). These differences are probably related to strain differences between the WAL 2926 and the VPI-5482 isolates (NCBI deposited sequence).

5.4.4 Expression and Purification of BactnABC

E. coli BL21(DE3) was transformed with the expression vector containing BactnABC. Bacteria were cultivated in 1 l LB-Ap medium and grown at 37°C with aeration. At an OD_{600} of 0.6-0.8 the cultures were transferred to 30°C for 30 minutes, induced with 1 mM IPTG and incubated for 4 h at 30°C with aeration. Selenomethionine (SeMet) labeled protein was expressed in *E. coli* methionine auxotroph DL41(DE3) maintained in LeMaster medium supplemented with 25 μg ml l^{-1} L-SeMet (Hendrickson *et al.* 1990). Cultures were incubated for 12-16 hours at room temperature in the medium after induction with 1 mM IPTG.

The cells were harvested by centrifugation at 6000 x g for 30 min at 4 °C and the pellet was resuspended in 25 ml binding buffer (50 mM K_2PO_4 pH 7.6, 400 mM NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol (β-ME)). Cells were lysed by the addition of 0.25 μg ml^{-1} lysozyme followed by sonication. After ultracentrifugation the lysate was applied to a 1.5 ml DEAE-Sepharose (GE Healthcare) column and the unbound protein was loaded onto a 3 ml Ni-NTA column (Qiagen). The column was washed by 50 column volumes of buffer (50 mM K_2PO_4 pH 7.6, 1.0 M NaCl, 10 mM imidazole, 5 mM β-ME) and bound His_8-BactnABC eluted with 50 mM K_2PO_4 pH 7.6, 200 mM NaCl, 300 mM imidazole, 5 mM β-ME. This eluate was diluted 3X with 10 mM K_2PO_4 pH 6 and loaded onto a pre-equilibrated Mono-S HR 10/10 column (GE Healthcare). The column was washed with 30 ml of 10 mM K_2PO_4 pH 6, 200 mM NaCl and the protein eluted by 200-500 mM NaCl gradient over 5 column volumes. Fractions containing BactnABC were combined, concentrated on a Centricon YM-50 concentrator (Millipore Corporation) and applied to a HiLoad 16/60 Superdex 200 column (GE Healthcare), with

50 mM K_2PO_4 pH 7.2, 150 mM NaCl, 5 mM DTT as running buffer. BactnABC containing fractions were pooled and concentrated to 6-8 mg ml^{-1}. SeMet labeled protein and BactnABC mutants were purified in the same manner.

5.4.5. Kinetic Analysis and Divalent Cation Requirement

Activity of BactnABC and its mutants was determined spectrophotometrically at 37°C using chondroitin-4-sulfate, CS-A (BioCorp), DS (Sigma) and chondroitin-6-sulfate, CS-C (Sigma) as substrates at concentrations of 0.1-5 mg ml^{-1} in a reaction buffer containing 50 mM K_2PO_4, pH 7.6. 1.0 µg of BactnABC was added and initial reaction rates were monitored for 300 s following the absorbance at 232 nm. Kinetic parameters (k_{cat} and K_M) were determined using Hanes plots according to the equation: [S]/v = Km/Vm + [S]/Vm. Data were interpreted using the program Hyper Version 1.0.0 (http://homepage.ntlworld.com). The molar absorption coefficient (ε) of 3800 M^{-1} cm^{-1} was used to calculate product release.

The influence of the divalent cations Ca^{2+}/Mg^{2+} on the activity of BactnABC was evaluated in Tris-HCl buffer. The protein was dialyzed into 50 mM Tris-HCl, pH 7.6 and its specific activity against 1 mg of different substrates was measured as a function of $CaCl_2$ concentration. The experiment was repeated with addition of 10 mM EDTA, and 10 mM $MgCl_2$.

5.4.6. Crystallization and Data Collection

Initial crystallization conditions were identified by hanging drop vapor diffusion at 20°C using the Index Screen (Hampton Research). After optimization, the best crystals grew from 17% (w/v) PEG 3350, 0.2 M ammonium sulfate and 100 mM HEPES pH 6.50-7.00 at 4°C. These conditions were also used to produce crystals from SeMet labeled protein. The crystals belong to the space group $P6_3$, with unit cell dimensions a = 223.4, c = 112.6 Å and contain two molecules in the asymmetric unit with V_M=3.53 $Å^3$ Da^{-1} (solvent contents 65%). For data collection the crystals were soaked at 4°C in mother liquor supplemented with 12% (w/v) glycerol, picked up in a nylon loop and flash cooled in liquid N_2 at 100 °K. Diffraction data extending to 2.85 Å resolution were collected at the Se absorption peak (λ=0.97233 Å) of ID22 beamline (APS) on a MAR

CCD 300 mm area detector (MAR Research, Germany) set at a distance of 400 mm. Images were processed with the HKL2000 program (Otwinowski & Minor 1997).

5.4.7 Structure Determination and Refinement

The structure of BactnABC was solved by single-wavelength anomalous dispersion (SAD). A total of 70 selenium sites were located using the program SHELXD (Sheldrick 1991). The heavy atom sites were refined using SHARP (Bricogne *et al.* 2003). Solvent flattening with RESOLVE (Terwilliger 2002) yielded a figure-of-merit of 0.82. The model was built with COOT (Emsley & Cowtan 2004) and refined with REFMAC5 (Emsley & Cowtan 2004) against the SAD dataset, including the TLS option (the molecule was divided into three domains). A strong peak in the electron density was interpreted as a Ca^{2+} ion based on the ligand geometry and reasonable B-factors. Refinement at 2.85 Å resolution converged with an *R*-factor of 0.220 and *R*-free of 0.262, with two SeMet-labeled monomers in the asymmetric unit consisting of residues Ala-14 to Lys-1014, one bound Ca^{2+} ion for each monomer, 16 phosphate ions, and 67 water molecules (Table 5-1). Residues Ser558–Tyr567 are disordered and were not modeled. The final model has good geometry (PROCHECK (Laskowski *et al.* 1993)). Coordinates have been deposited to the Protein Data Bank (Berman *et al.* 2000) with the code 2Q1F.

5.5 Acknowledgements

We would like to thank Dr. Z. Dauter for help with data collection and to Drs. Joseph D. Schrag and John S. Mort for critical comments. This research was supported by the Canadian Institutes of Health Research (CIHR) grant MOP-74725 to M.C. and by the grant from National Institute of Agricultural Biotechnology to B.S.H.

5.6 Appendix A – Supplementary Material

Table 5-S1. Summary of primer sequences for site-directed mutagenesis.
Mutation triplet codons are indicated in bold; bases modified in order to create the desired point mutations are underlined. Forward primers for each mutant are provided.

Mutant	**Primer Sequence**
Arg^{172}→Ala	5' CGAAGGTAGACGCC**GCT**CAGCAGACAGCCGATC 3'
Gln^{173}→Ala	5' GGTAGACGCCCGT**GCG**CAGACAGCCGATCTG 3'
Arg^{267}→Ala	5' TGCCCATTTTCATGGTG**GCC**GCTTCCGAAGCC 3'
His^{344}→Ala	5' GTAGTTGCTGGGGCAATATT**GCT**CATTACGGATATAGCG 3'
His^{345}→Ala	5' GGGGCAATATTCAT**GCT**TACGGATATAGCGTGCGCG 3'
His^{453}→Ala	5'GGACGGAGGAGCTTTC**GCT**CACCGTAATAACTACCC 3'
His^{454}→Ala	5' GACGGAGGAGCTTTCCAT**GCC**CGTAATAACTACCCTGCC 3'
Tyr^{461}→Phe	5' GTAATAACTACCCTGCC**TTT**GCCGTTGGAGGACTGG 3'
Arg^{514}→Ala	5' CCTTGTCTATGTCCGGG**GCC**CATCCCGATGGACAGG 3'
Glu^{628}→Ala	5' CTACCTTTGGGCAGCC**GCA**CATTATCTGGGACACAAC 3'

Table 5-S2. Sequence of BactnABCI.
Amino acids differing from the NCBI sequence highlighted in grey.

```
  1    MLILSFLCPAFLNAQIVTDERMFSFEEPQLPACITGVQSQLGISGAHYKDGKHSLEWTFEPN
 63    GRLELRKDLKFEKKDPTGKDLYLSAFIVWIYNEQPQDAAIEFEFLKDGRKCASFPFGINFKG
125    WRAAWVCYERDMQGTPEEGMNELRIVAPDAKGRLFIDHLITATKVDARQQTADLQVPFVNAG
187    TTNHWLVLYKHSLLKPDIELTPVSDKQRQEMKLLEKRFRDMIYTKGKVTEKEAETIRKKYDL
249    YQITYKDGQVSGVPVFMVRASEAYERMIPDWDKDMLTKMGIEMRAYFDLMKRIAVAYNNSEA
311    GSPIRKEMRRKFLAMYDHITDQGVAYGSCWGNIHHYGYSVRGLYPAYFLMKDVLREEGKLLE
373    AERTLRWYAITNEVYPKPEGNGIDMDSFNTQTTGRIASILMMEDTPEKLQYLKSFSRWIDYG
453    CRPAPGLAGSFKVDGGAFHHRNNYPAYAVGGLDGATNMIYLFSRTSLAVSELAHRTVKDVLL
496    AMRFYCNKLNFPLSMSGRHPDGKGKLVPMHYAIMAIAGTPDGKGDFDKEMASAYLRLVSSDS
559    SSAEQAPEYMPKVSNAQERKIAKRLVENGFRAEPDPQGNLSLGYGCVSVQRRENWSAVARGH
621    SRYLWAAEHYLGHNLYGRYLAHGSLQILTAPPGQTVTPTTSGWQQEGFDWNRIPGVTSIHLP
683    LDLLKANVLNVDTFSGMEEMLYSDEAFAGGLSQGKMNGNFGMKLHEHDKYNGTHRARKSFHF
745    IDGMIVCLGSDIENTNMDYPTETTIFQLAVTDKAAHDYWKNNAGEGKVWMDHLGTGYYVPVA
807    ARFEKNFPQYSRMQDTGKETKGDWVSLIIDHGKAPKAGSYEYAILPGTDRKTMTAFAKKPAY
869    SVLQQDRNAHILESPSDRITSYVLFETPQSLLPGGLLQRTDTSCLVMVRKESADKVLLTVAQ
931    PDLALYRGPSDEAFDKDGKRMERSIYSRPWIDNESGEIPVTVTLKGRWKVVETPYCKVVSED
993    KKQTVLRFLCKDGASYEVELEK
```

Figure 5-S1. Superposition of BactnABCII vs. PvulABCI.
Stereo view of the superposition of BactnABCI (blue) and PvulABCI (orange).

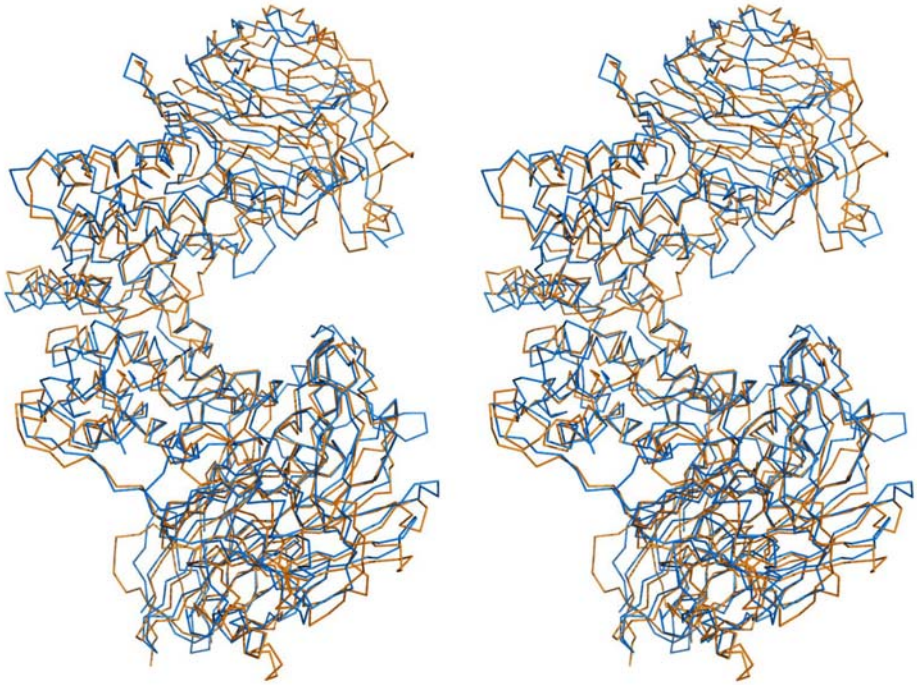
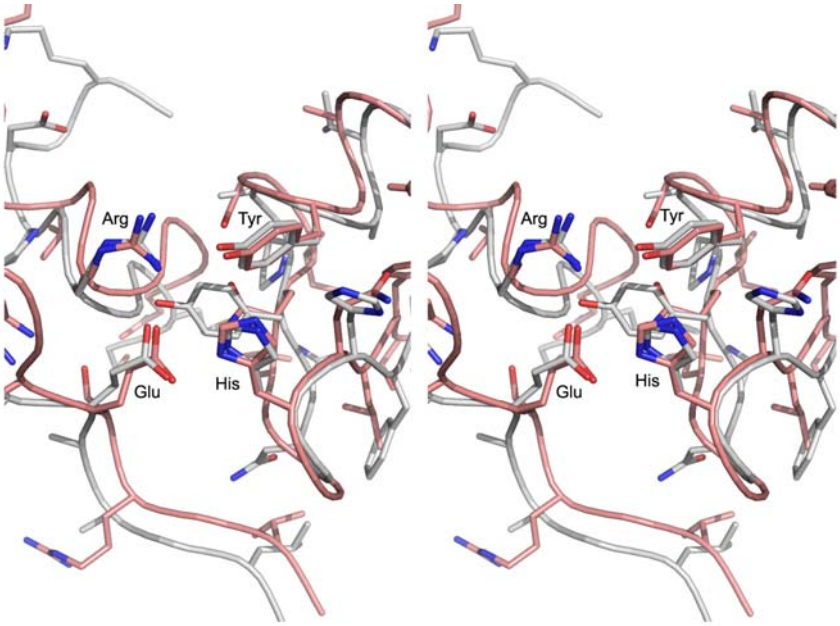


Figure 5-S2. Structural conservation of the tetrad.
Superposition of the active site tetrad residues of ArthroAC (pink) and the equivalent



Chapter 6 Characterization of Chondroitin Sulfate Lyase ABC from *Bacteroides Thetaiotaomicron*

Preface

The contents presented in the following chapter have been published as presented:

Shaya, D., Hahn B.-S., Park N.Y., Sim J.-S., Kim Y.S., and Cygler M. (2008) Characterization of Chondroitin Sulfate Lyase ABC from *Bacteroides thetaiotaomicron* WAL2926. *Biochemisty.* May 31, 2008

Contribution of Authors:

I performed the thermal stability experiments characterizing B.thetaiotaomicron *chondroitin lyase ABC in comparison to the* P.vulgaris *enzyme, designed and constructed the mutant proteins, and proposed their kinetic characterization experiments. I have proposed the experiments characterizing chondroitin lyase ABC, analyzed the data and wrote the draft of the manuscript. The kinetic characterization of chondroitin lyase ABC and its mutants was carried out by N.Y.P. and J.-S.S under the supervision of B.-S.H. and Y.S.K. M.C. supervised the project, orchestrated the collaboration, and prepared the final manuscript for publication with my aid.*

6.1 Rationale

Based on the structure and the preliminary site-directed mutagenesis data, presented in the previous chapter, I commenced the analysis of the active site of BactnABCII. In particular, I have focused on the active site residues involved in DS degradation. In parallel, collaborating with the laboratory of Prof. Yeong Shik Kim we have undertaken a biochemical characterization, establishing the mode of action, conditions for optimal in-vitro activity, temperature dependence and stability of the enzyme. This chapter resumes my studies of BactnABC and proposes catalytic roles for the active site residues involved in DS degradation, as well as establishing its Ca^{2+} dependence for degradation of DS.

6.2 Abstract

Chondroitin sulfate ABC lyase (ChonABC) is an enzyme with broad specificity that depolymerizes via β-elimination, chondroitin sulfate (CS) and dermatan sulfate (DS) glycosaminoglycans (GAGs). ChonABC eliminates the glycosidic bond of its GAG substrates, on the non-reducing end of their uronic acid component. This lyase possesses the unusual ability to act on both epimers of the uronic acid, either glucuronic acid present in CS, or iduronic acid in DS. Recently, we cloned, purified, and determined the 3-D structure of a broad specificity chondroitin sulfate ABC lyase from Bacteroides thetaiotaomicron (BactnABC) and identified two sets of catalytic residues. Here, we report the detailed biochemical characterization of BactnABC together with extensive site directed mutagenesis resulting in characterization of the previously identified active site residues. BactnABC's catalysis is stimulated by Ca^{2+} and Mg^{2+} cations, particularly against DS. It displays extremely low activity toward hyaluronic acid and no activity toward heparin/heparan sulfate. Degradation of CS and DS by BactnABC yields only disaccharide products, pointing to an exolytic mode of action. The kinetic evaluations of the active-site mutants indicate that CS and DS substrates bind in the same active site, which is accompanied by a conformational change bringing the two sets of active site residues together. Conservative replacements of key residues suggest that His345 plays the role of a general base, initiating the degradation by abstracting the C5 bound proton

from DS substrates, whereas either Tyr461 or His454 perform the equivalent role for CS substrates. Tyr461 is proposed, as well, to serve as general acid completing the degradation of both CS and DS by protonating the leaving group.

6.3 Introduction

Glycosaminoglycans are highly sulfated, linear polysaccharides composed of alternating hexosamine (glucosamine or galactosamine) and uronic acid (D-glucuronic or L-iduronic) moieties. Most GAGs are linked through a serine residue to a core protein forming a proteoglycan (PG). PGs are ubiquitous within the extracellular matrix (ECM) and are present on the cell surfaces of vertebrates. Through their interaction with a wide variety of proteins, PGs modulate fundamental biological processes including cell adhesion, proliferation, differentiation, signaling, inflammation, and infection (Varki 1993; Iozzo & Cohen 1993; Iozzo 1998; Lander & Selleck 2000; Sugahara & Kitagawa 2000; Trowbridge & Gallo 2002; Sugahara *et al.* 2003; Bao *et al.* 2004).

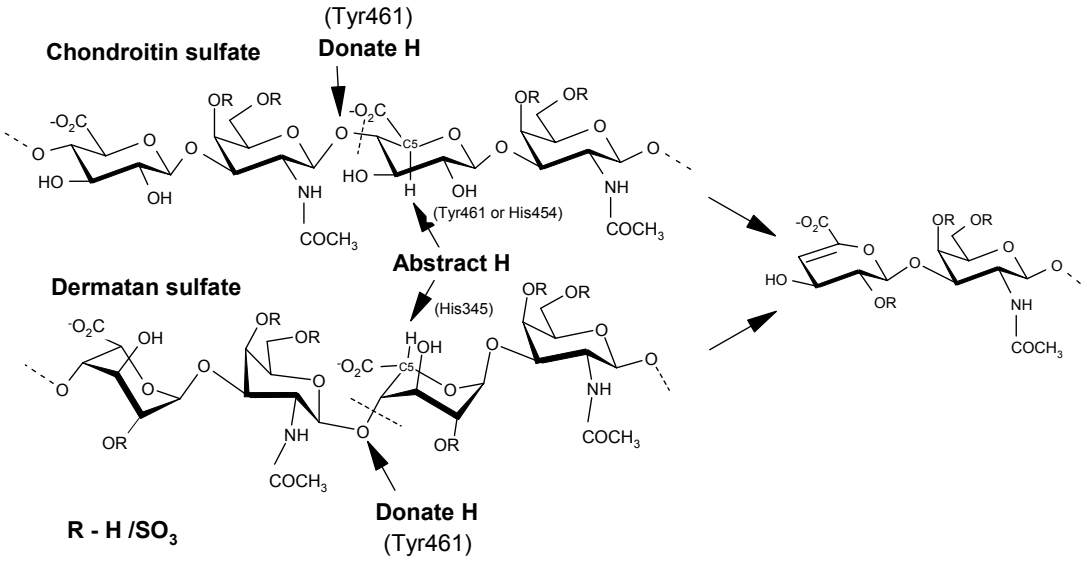
Dermatan sulfate (DS) and chondroitin sulfate (CS) are structurally-related GAGs, composed of alternating 1,4-β-D-N-acetylgalactosamine (GalNAc) and 1,3-*α*-L-iduronic acid (DS) or 1,3-β-D-glucuronic acid (CS) (Ernst *et al.* 1995; Sugahara & Kitagawa 2000). The uronic acid moiety in CS is exclusively β-D-glucuronic acid (GlcA), whereas DS contains a mixture of *α*-L-iduronic acid (IdoA) and GlcA epimers (Fig. 6-1) (Valla *et al.* 2001). During their biosynthesis, CS and DS polysaccharides undergo sulfation at various positions generating their chemical diversity. Chondroitin sulfate is primarily characterized by *O*-sulfation at either the C4 position (chondroitin 4-sulfate, CS-A) or the C6 position (chondroitin 6-sulfate, CS-C) of GalNAc, whereas DS is primarily sulfated at the C4 position of the galactosamine. Additional, less frequent *O*-sulfation is found at the C2 position of IdoA in DS creating di-sulfated disaccharide units that have been implicated in regulating the biological functions of this polymer (Iozzo 1998).

Specialized microorganisms express GAG-degrading lyases using their terminal products as a source of carbon. Unlike the eukaryotic GAG hydrolases, these prokaryotic enzymes depolymerize GAGs through a β-elimination mechanism characterized by removal of a relatively acidic proton from the C5 carbon (the chiral center) of the uronic

acid, and the release of the 4-linked hexosamine with the generation of a C4-C5 double bond at the uronic acid ring (Gacesa 1987) (Fig. 6-1). During the past two decades GAG lyases have found uses as tools to dissect the complex structure of GAG polysaccharides (Aguiar *et al.* 2003). The broad specificity ChonABC from the bacterium *Proteus vulgaris,* depolymerizing both CS and DS, was recently shown to enhance axonal regeneration and improve nerve plasticity following spinal cord injury in model systems (Bradbury *et al.* 2002; Lemons *et al.* 2003). Similar broad specificity ChonABC lyases have been identified in the genome of *Bacteroides thetaiotaomicron* (Linn *et al.* 1983). A thorough biochemical characterization of this enzyme, together with a simple expression-purification protocol, could offer novel therapeutic tools to overcome the debilitating spinal cord injury.

ChonABC was initially identified in the gram-negative bacterium *Proteus vulgaris*. Two orthologs, the endolytic PvulABCI (gi: 1095454, MW of 115.2 kDa and 1,021 amino acids) and the exolytic PvulABCII (gi:1828879, with MW of 114.1 kDa and 1,013 amino acids) were biochemically characterized (Hamai *et al.* 1997; Prabhakar *et al.* 2005b). We have previously determined the crystal structure of PvulABCI (Huang *et al.* 2003b). Continuing our efforts to explain the enzyme's broad substrate specificity and its mechanism of action, we recently cloned, expressed and determined the structure of chondroitinase ABC from the gram-negative Bacteroides thetaiotaomicron WAL2926 (BactnABC, MW of 115.2 kDa and 1,016 amino acids, sequence in supplementary material table 6-S2). Comparison of the structures of these two enzymes led to the identification of two clusters of essential active site residues: in addition to the Tyr461-His454-Glu628-Arg514 tetrad we have identified His344, His345 and Arg172, located on the opposite rim of the substrate binding site, as affecting significantly catalytic activity (Huang *et al.* 2003). We present here the biochemical characterization of BactnABC, providing conditions for optimal activity, and report the enzyme's product profile against a broad spectrum of GAG substrates. We performed an extensive mutational analysis of the residues implicated in activity, using both, short-term kinetic assays and overnight end-point product formation assays to confirm their functional roles in catalysis.

Figure 6-1. Chemical schemas of CS and DS degradation



6.4 Materials and Methods

6.4.1 Site-Directed Mutagenesis of BactnABC

Genomic DNA was isolated from cultures of *Bacteroides thetaiotaomicron* strain WAL2926 (Korean Culture Center of Microorganisms), and the gene BT3324 (gi: 29348733) was subcloned into the expression vector pET-22b (Novagene) with a C-terminal His_6 tag as described previously (Shaya 2008). DNA sequencing revealed 16 amino acid differences (supplementary material table 6-S1) with the published sequence (gi: 166235375) related to strain differences between the WAL2926 and the VPI-5482 isolates (NCBI deposited sequence). BactnABC mutants were constructed using the QuickChange Site-Directed Mutagenesis Kit (Stratagene), according to the manufacturers' protocol, using DNA primers listed in Table 6-S2 and their sequences confirmed by DNA sequencing. The mutants were expressed in *E. coli* BL21 (DE3).

6.4.2 Purification of Recombinant BactnABC

Recombinant BactnABC and derived mutant enzymes were expressed in *E. coli* strain BL21(DE3), which is devoid of GAG lyase activity (Prabhakar *et al.* 2005), and purified as described earlier (Shaya 2008). Briefly, cells expressing BactnABC were

cultured in Luria broth at 37^{o}C supplemented with 100 μg ml^{-1} ampicillin and protein expression induced with 1 mM IPTG. Cells were disrupted by sonication and purified using a combination of DEAE (GE Healthcare), Ni-NTA (Qiagen), Mono-S HR 10/10 (GE healthcare) and HiLoad 16/60 Superdex 200 (GE healthcare) columns. The purity of the eluted fractions was evaluated by SDS-PAGE stained by Coomassie Blue R-250, which showed a single band on an overloaded gel, corresponding to the expected MW of the recombinant enzyme. Fractions containing purified, active BactnABC were pooled, concentrated to 8 mg/ml by ultra-filtration (Centricon YM-50 concentrator, Millipore Corporation), and used in subsequent characterization experiments.

6.4.3 Biochemical Characterization of Recombinant BactnABC

6.4.3.1 BactnABC Activity

BactnABC lyase activity was measured based on the increase in absorbance at 232 nm (Yamagata *et al.* 1968). One unit of enzyme is the amount required for the eliminative cleavage of substrate, yielding 1 μmol of UV-absorbing, double bond-containing products per min as calculated using a molar absorption coefficient (ε) of 3800 M^{-1} cm^{-1}. All experiments were carried out in a 1 cm path length quartz cuvette with a total volume of 1 ml, and included at least three repetitions. Protein concentrations were determined using Bradford Protein Assay Reagent kit (Bio-Rad).

6.4.3.2 Buffer Selection

We first established the appropriate buffer for subsequent biochemical characterizations by comparing enzyme specific activity in Tris versus phosphate buffer system. Accordingly, 50 mM potassium phosphate, the optimal buffer, was used in determination of pH and temperature effects on the activity, and the optimal, pH (7.6) and temperature (37°C) established were used in subsequent experiments. Since phosphate forms insoluble salts with some metal ions, Tris-HCl buffer was used to determine the influence of ionic strength (sodium chloride and potassium chloride) and of divalent metal cations on the activity of the wild-type enzyme.

6.4.3.3 Substrate Preference

To determine the substrate preferences of the enzyme, twelve GAG substrates from different sources and manufacturers were used: CS-A from bovine trachea (Biocorp), DS from porcine intestinal mucosa (Sigma, Calbiochem), CS-C from shark cartilage (Sigma, Seikagaku), CS-D from skate (unpublished results), CS-E from squid (unpublished results), low molecular weight (100 kDa) and high molecular weight (1400 kDa) hyaluronan from chicken combs (Kibun Food Chemifa Co, Japan), acharan sulfate from giant snail (according to: (Kim *et al.* 1996c)), heparin from porcine intestinal mucosa (New Zealand Pharmaceuticals Ltd.), and heparan sulfate from bovine kidney (Seikagaku). Each substrate was dissolved at a concentration of 1 mg ml^{-1} in 50 mM phosphate buffer pH 7.6. Unless specified otherwise, the reaction was initiated by the addition of one microgram of wild type BactnABC and $\Delta A232$ min^{-1} was recorded over 300 sec at 37°C. Since the chondroitin sulfates substrates, derived from various natural sources, differ in the level and type of substitutions; we have characterized their disaccharide composition by a complete digestion using PvulABC. Disaccharide separation and identification was preformed using FPLC as described later in the product profile analysis section, composition analysis was calculated by integrating and comparing the peak areas (mAU min).

6.4.3.4 Optimal Temperature, pH, Ionic Strength, and Divalent Metal Cations

CS-A (Biocorp), DS (Sigma) and CS-C (Sigma), each at a concentration of 1 mg ml^{-1}, were assayed under various conditions in an attempt to determine the effects of temperature, pH, ionic strength and divalent metal cations on enzymatic activity. The temperature was varied between 25-50°C using a temperature-controlled UV spectrophotometer (DU 800, Beckmann-Coulter) and 50 mM phosphate; pH 7.6 was used as the reaction buffer. To determine BactABC's temperature stability we have re-measured the activity at 37^{o}C after pre-incubation for 5 min at each temperature in the 25-50^{o}C range. The influence of pH on activity was investigated in 50 mM phosphate buffer, pH 6-9. The relative effects of ionic strength, divalent metal ions, and titration of Ca^{2+} cations were measured in Tris-HCl buffer. The protein was dialyzed into 50 mM Tris-HCl, pH 7.6 and increasing concentrations of sodium, potassium and calcium

chloride were added to the reaction. To evaluate the effect of divalent metal ions on the activity of BactnABC, a constant concentration of various metal chloride-salts (10 mM) was added to the reaction.

6.4.3.5 Determination of Apparent T_m

A fluorescence microplate reader (FluoDia T70, Photon Technology International) was used to monitor the thermal denaturation of BactnABC and PvulABC using SYPRO orange fluorescence as a readout (Matulis *et al.* 2005; Vedadi *et al.* 2006). PvulABC was purchased from Seikagaku. A final concentration of 2 µM protein combined with SYPRO Orange (Invitrogen) in reaction volumes of 25 µl, were incubated in 50 mM phosphate, pH 7.6, and 50 mM Tris-HCl, pH 7.6, respectively. The temperatures were varied between 23-75°C using 1°C steps. The samples were excited at 465 nm, and the fluorescence emission recorded at 590 nm. The data was interpreted using the Levenberg-Marquardt nonlinear regression algorithm (Octave, http://www.octave.org). The inflection point (apparent T_m) was determined by fitting the data to the equation f(T)=A+(B-A)/(1+exp(Tm-T)/C), where A is the fluorescence minimum, B is the fluorescence maximum, C is the rate of fluorescence change around the inflection point, and T is the temperature.

6.4.3.6. Kinetic Analysis

The kinetic parameters of BactnABC and its mutants toward CS-A (Biocorp), DS (Sigma) and CS-C (Sigma) were determined spectrophotometrically. Various quantities of substrate (0.1-5 mg) were dissolved in 50 mM phosphate pH 7.6, at 37°C, and the initial reaction rates were recorded within 1 to 5 min. The data were interpreted using linear regression analysis (Hyper Version 1.0.0 http://wvlc.uwaterloo.ca/biology447/modules/module7/hyper/index.htm).

The kinetic parameters (k_{cat} and K_M) were determined by fitting the data to the equation: [S]/v = K_M/Vm + [S]/Vm (Hanes plot), where K_M represents the substrate concentration at half-saturation. The k_{cat} value was calculated using the equation V_{max} = k_{cat} x [E], where [E] represents total enzyme concentration.

6.4.3.7 Product Profile Analysis

The product profiles resulting from exhaustive digestion of various substrates by BactnABC and its corresponding mutant enzymes were analyzed using strong anion-exchange, high performance liquid chromatography (SAX-FPLC). For each reaction, 0.5 mg of CS-A (Biocorp), DS (Sigma), or CS-C (Sigma) were enzymatically depolymerised using 0.25 μg, 0.75 μg, and 0.5 μg enzyme respectively, in total volumes of 500 μl of 50 mM phosphate, pH 7.6. The reactions were incubated overnight at 37^{o}C and terminated by heating for 5 min at 100°C. Samples were applied to a Hypersil SAX column (250 x 4.6 mm, 5 μm, Bellefonte, PA, USA) operated using a ÄKTA purifier FPLC (GE Healthcare). The column was washed with one column volume of water (pH 3.5, adjusted with HCl), and a linear gradient of 0 - 1.0 M NaCl (pH 3.5, adjusted with HCl) over 10 column volumes was used to separate the products. Disaccharide standard mixtures (Seikagaku), analyzed in the same manner, were used to identify the digestion products by comparing retention times.

6.5 Results

6.5.1 Biochemical Characterization of BactnABC

6.5.1.1 Buffer Effects on Activity

Previous reports showed that the enzymatic activity of BactnABC is strongly influenced by the buffer used. Linn and coworkers reported maximal activity for BactnABC in phosphate buffer, while using Tris-HCl buffer required the addition of NaCl (Linn *et al.* 1983). To establish the appropriate buffer for recombinant BactnABC we compared its activity in Tris versus phosphate buffer system. The enzyme activity measured in phosphate is 77.6 U/mg for CS-A, 47.4 U/mg for CS-C and 9.1 U/mg for DS (Tables 6-1 and 6-2). When the activity is measured in Tris-HCl buffer, we observed a ~6 fold lower activity for the CS-A/CS-C substrates while a much larger, 80 fold decrease for DS (Table 6-2). Accordingly, 50 mM phosphate was used in determination of pH and temperature effects on the activity.

Table 6-1. Specific activity of BactnABC acting on various GAG substrates

One mg ml^{-1} of each substrate was dissolved in 50mM phosphate, pH 7.6, and the reaction monitored for 300 sec at 37°C following addition of the enzyme. The values for the specific activities are a mean ± S.D. of at least three measurements. Disaccharides composition of the substrates are: ΔDi0S = ΔUA-GalNAc; ΔDi6S = ΔUA-GalNAc6S; ΔDi4S = ΔUA-GalNAc4S; ΔDi2,6 diS = ΔUA2S-GalNAc6S; ΔDi4,6 diS = ΔUA-GalNAc4S6S; ΔDi2,4 diS = ΔUA2S-GalNAc4S.

Substrate	Average MW (kDa)	Specific activity (U/mg protein)	Relative activity (%)	ΔDi-0S (%)	ΔDi-6S (%)	ΔDi-4S (%)	ΔDi-2,6 diS (%)	ΔDi-4,6 diS (%)	ΔDi-2,4 diS (%)
CS-A bovine trachea (Biocorp)[1]	12	77.6±1.1	100	5.4	9.2	74.7	0.4	0.3	0
DS porcine intestinal mucosa (Sigma)[1]	16	9.1±0.8	11.7	1.4	5.5	85.7	0.7	6.7	0
DS porcine intestinal mucosa (Calbiochem)[1]	20	12.5±0.9	16.1	23.8	6.1	63.7	0	5.2	1.2
CS-C shark cartilage (Sigma)[1]	35	47.4±0.8	61.1	2.8	46.4	31.2	18.2	0.7	0.7
CS-C shark cartilage (Seikagaku)[1]	39	18.2±1.0	23.4	4.1	61.1	21.5	13.1	0.2	0
CS-D skate cartilage[1,4]	44	14.4±0.4	18.5	2.5	40.5	35.6	20.7	0.1	0.6
CS-E squid cartilage[1,5]	97	28.5±0.8	36.7	13.0	22.1	45.8	0	19.1	0
Low molecular weight Hyaluronic acid (Kibun Food Chemifa) [2]	100	0.8±0.1	1.1	-	-	-	-	-	-
High molecular weight Hyaluronic acid (Kibun Food Chemifa) [3]	1400	0.2±0.1	0.2	-	-	-	-	-	-
Acharan sulfate[3,6]	-	n.d.	n.d.	-	-	-	-	-	-
Heparin (New Zealand Pharmaceuticals) [3]	-	n.d.	n.d.	-	-	-	-	-	-
Heparan sulphate (Seikagaku) [3]	-	n.d.	n.d.	-	-	-	-	-	-

[1]1.0 μg of WT BactnABC was used in the assay
[2]7.5 μg WT BactnABC was used in the assay
[3]15 μg WT BactnABC was used in the assay
[4]CS-D was purified from skate cartilage
[5]CS-E was purified from squid cartilage
[6]Acharan sulfate was purified from Giant snail according to (Kim *et al.* 1996).

Table 6-2. Specific Activity of BactnABC in the Presence of Divalent Cations[a]

Divalent cation	Specific activity (U/mg)		
	CS-A	CS-C	DS
50 mM phosphate (pH 7.6)	77.6 ± 1.1	47.4 ± 0.8	9.1 ± 0.8
50 mM Tris-HCl (pH 7.6)	12.3 ± 0.7	8.1 ± 0.5	0.1 ± 0.1
+10 mM CaCl2	28.2 ± 1.0	14.6 ± 0.9	4.5 ± 0.3
+10 mM MgCl2	38.2 ± 0.4	19.3 ± 0.8	5.2 ± 1.8
+10 mM MnCl2	4.3 ± 1.6	4.0 ± 0.5	0.7 ± 0.2
+10 mM ZnCl2	1.3 ± 0.4	0.9 ± 0.6	0.7 ± 0.2
+10 mM BaCl2	18.6 ± 0.2	10.2 ± 0.6	4.4 ± 0.1
+10 mM CsCl	12.1 ± 0.2	8.2 ± 0.4	0.7 ± 0.1

[a] One mg/ml CSA (Biocorp)/CSC (Sigma)/ DS (Sigma) were assayed with 1 μg of BactnABC in 50 mM Tris-HCl (pH 7.6) containing various divalent cations at a final concentration of 10 mM. The values given are the mean ± S.D. for at least three measurements.

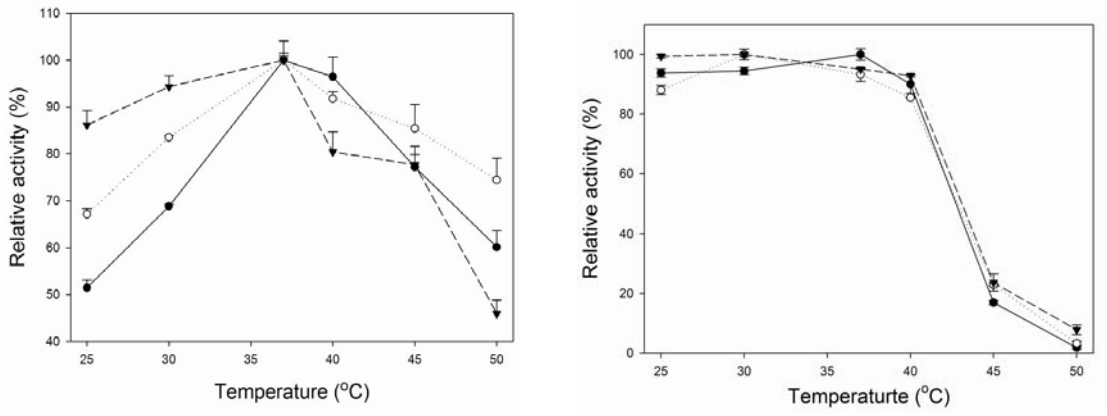
6.5.1.2 Temperature Dependence of Activity and Enzyme Stability

The temperature dependence of BactnABC activity and its thermal stability was tested in the range of 25-50^{o}C (Fig. 6-2). The enzyme displayed no significant reduction in activity at temperatures in the range of 25-40^{o}C. Maximal activity for all substrates was attained at 37^{o}C. At lower temperatures the activity toward DS drops least (85% at 25^{o}C) while the activity towards CS-A and CS-C was reduced to 50% and 65%, respectively. However, at higher temperatures the activity toward DS declines faster (45% at 50^{o}C for DS *vs*. 60% for CS-A and 75% for CS-C). To determine if this decrease of activity above 40^{o}C is caused in part by irreversible denaturation, we have incubated the enzyme for 5 min at each temperature and transferred the enzyme to 37^{o}C after for activity measurements (Fig. 6-2b). The results indicate that denaturation is, indeed, a factor at temperatures above 40^{o}C.

We further investigated the temperature stability of BactnABC by fluorescence. We measured the changes in fluorescence of the SYPRO Orange dye as a function of temperature. This fluorophore is sensitive to the change in hydrophobicity of the environment. The increased exposure of hydrophobic regions of the protein during

thermal denaturation leads to an increase in the fluorescence emitted from the fluorophore (Linn *et al.* 1983). Thermal unfolding was measured in two buffers, potassium phosphate and Tris-HCl, at pH 7.6. In both buffers the fluorescence emission as a function of temperature showed a bimodal melting curve with two inflection points representing two consecutive denaturing events. The first T_m corresponds to 44°C and the second to 50°C. Thermal denaturation was not visible below 41°C, in complete agreement with the activity-measured temperature stability (Fig. 6-2b). PvulABC, on the other hand, presented a unimodal curve with a T_m of 38°C for thermal denaturation.

Figure 6-1. Effect of temperature and pH on BactnABC activity.
a) Effect of temperature on BactnABC activity, buffer - 50 mM phosphate, pH 7.6; **b)** Temperature stability of BactnABC, buffer - 50 mM phosphate, pH 7.6, 5 min incubation of BactnABC in temperatures varying between 25 – 50ºC followed by activity measurement at 37ºC. All experiments contained 1 mg ml^{-1} of: CS-A with 0.5 µg BactnABC (●), DS with 1.5 µg BactnABC (▼), and CS-C with 1 µg BactnABC (○).

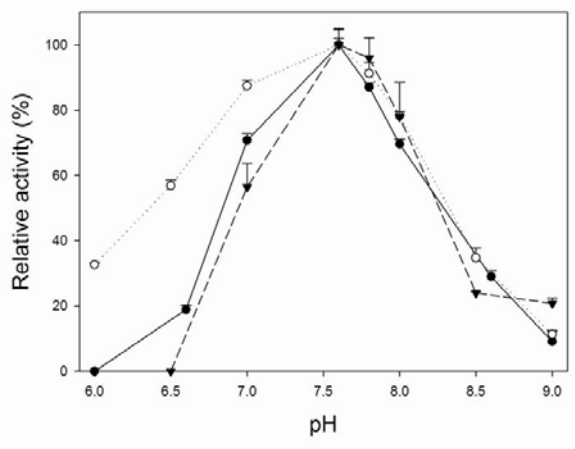


6.5.1.3 Effect of pH on Activity

The pH optimum for BactnABC was determined to be 7.6. The enzyme exhibited a bell-shaped, narrow pH-activity profile of ±~0.5 pH unit around the optimum, with marked decrease in activity at pH values below 7.0 and above 8.0 (Fig. 6-2b). Interestingly, the activity toward CS-C remained higher than with the other substrates at low pH. Thus, at pH 6.5 the enzyme exhibited almost 60% of its maximal activity toward CS-C, whereas only 20% activity toward CS-A, and no measurable activity toward DS. In contrast, at pH of 8.5 BactnABC displayed ~30% of its maximal activity toward all

substrates. Subsequent assays were performed at the optimum pH of 7.6, and at a temperature of 37°C.

Figure 6-2. pH effects on BactnABC activity
Buffer - 50 mM phosphate with varying pH. All experiments contained 1 mg ml^{-1} of: CS-A with 0.5 µg BactnABC (●), DS with 1.5 µg BactnABC (▼), and CS-C with 1 µg BactnABC (○).



6.5.1.4 Substrate Preference

To map the substrate preference of BactnABC, we compared the enzyme's specific activity toward a wide range of GAG substrates: CS-A, DS, CS-C, CS-D, CS-E, low and high molecular weight hyaluronic acid, acharan sulfate (AS), heparin, and heparan sulfate (Table 6-1). The enzyme displayed somewhat different activity toward the same substrate obtained from different suppliers. We have, therefore, analyzed the disaccharide composition of each substrate by exhaustive digestion with PvulABC and analysis of the disaccharides by FPLC (Table 6-1) and identified the content of various substitutions. The differences were relatively small for DS but substantially larger for CS-C, reflecting different levels of 4-*O*-sulfation in CS-C from Sigma compared to CS-C from Seikagaku. CS-E resembles the disaccharide composition of CS-A as compared to CS-D, which resembles the composition of CS-C from Sigma. BactnABC is highly active toward chondroitin substrates, whereas it shows no activity toward heparin substrates (AS, heparin and heparan sulfate; Table 6-1).

6.5.1.5 Mode of Action

To determine whether BactnABC acts in an endolytic or exolytic manner, we examined the product profile of BactnABC using SAX-FPLC. We observed only disaccharide products for each of the substrates tested (Fig. 6-3). CS-A and DS digestion afforded three disaccharide products ΔUA-GalNAc4S, ΔUA-GalNAc6S, and ΔUA-GalNAc (ΔUA – 4,5 unsaturated uronic acid); with the 4-sulfated disaccharide being the major product (Fig. 6-3b). The profile for digestion of CS-C revealed the same three products with the 6-sulfated disaccharide being the major product (Fig. 6-3c). CS-C digestion yielded also the di-sulphated disaccharide product ΔUA2S-GalNAc6S. No tetrasaccharide products were found during the course of digestion, consistent with an exolytic mode of action for this enzyme.

6.5.1.6 Effect of Ions on Activity

The concentration of mono and divalent ions had a clear effect on enzyme activity. These measurements were performed in Tris-HCl buffer. For the monovalent ions tested, increasing the concentration of sodium, potassium or cesium chloride from 0 to 0.1 M increased the activity against all substrates (Fig. 6-4a-c). Further increases in salt concentration reduced the activity dramatically, with 50% inhibition occurring at ~ 0.15 M and a nearly complete inhibition at 0.4 M salt. This general effect of ionic concentration on BactnABC's activity likely explains the elevated activity in the phosphate buffer versus that of the Tris-HCl buffer.

We have previously demonstrated that Ca^{2+}/Mg^{2+} cations play a role in catalysis, especially pronounced in the case of DS (Kim *et al.* 1996). We continued investigating the effect of other divalent ions, Mn^{2+}, Zn^{2+}, on the activity of BactnABC. As reported earlier, the addition of 10 mM Ca^{2+} or Mg^{2+} ions increased the activity toward CS-A/CS-C by 2-3 fold while the activity toward DS increased much more significantly, by 50-fold (Table 6-2). Addition of Mn^{2+} and Zn^{2+} reduced the activity against CS substrates but increased the activity toward DS (Table 6-2).

Figure 6-3. Product profile analysis using SAX-FPLC

a) Disaccharide standard mixtures from Seikagaku. Degradation profiles of wild-type BactnABC acting on: **b)** Biocorp CS-A, **c)** Sigma CS-C and **d)** Sigma DS. Disaccharides identified: ΔDi0S = ΔUA-GalNAc; ΔDi6S = ΔUA-GalNAc6S; ΔDi4S = ΔUA-GalNAc4S; ΔDi2,6diS = ΔUA2S-GalNAc6S; ΔDi4,6diS = ΔUA-GalNAc4S6S; ΔDi2,4diS = ΔUA2S-GalNAc4S.

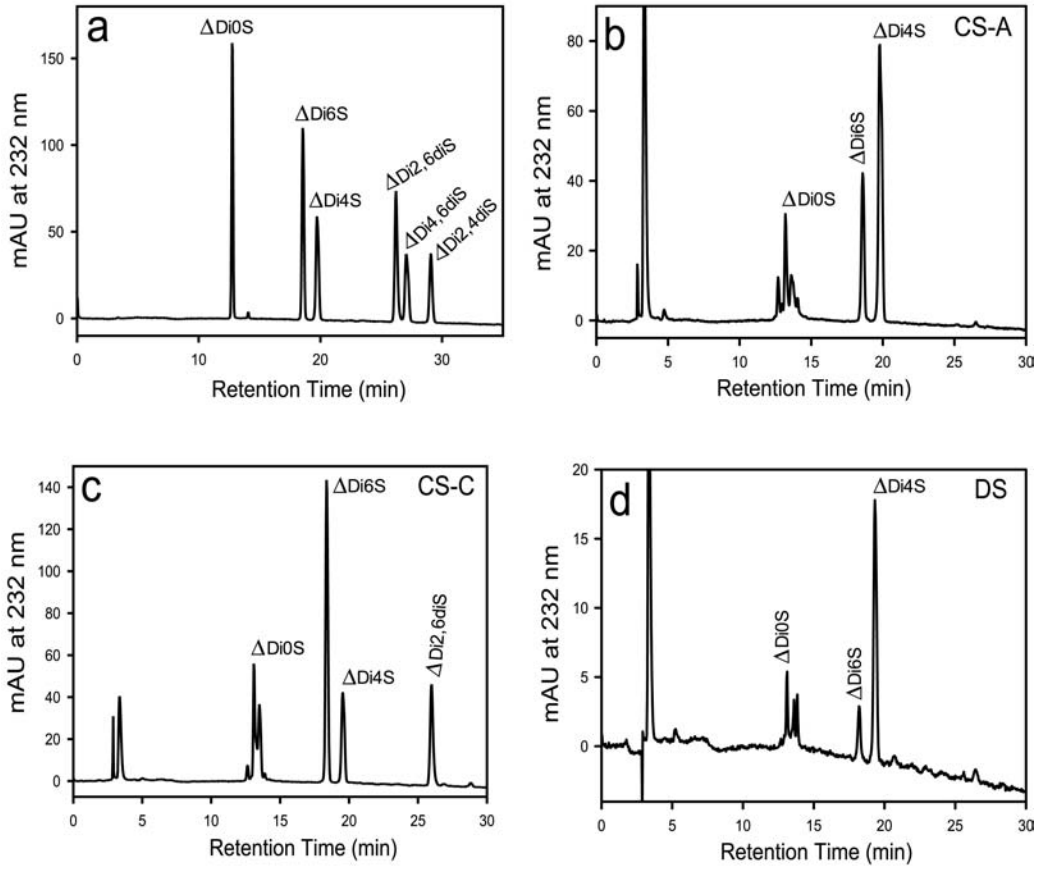


Figure 6-4. Influence of ionic concentration on BactnABC activity
Increasing concentrations of: **a)** KCl ,**b)** NaCl, **c)** CsCl, **d)** $CaCl_2$ and **e)** $BaCl_2$ in 50 mM Tris-HCl pH 7.6. The mixture contained 1 mg of CS-A (●, full lines), CS-C (○, dotted lines), DS (▼, dashed lines), and 0.5 µg, 1.0 µg, 1.5 µg BactnABC respectively in 1 ml volume.

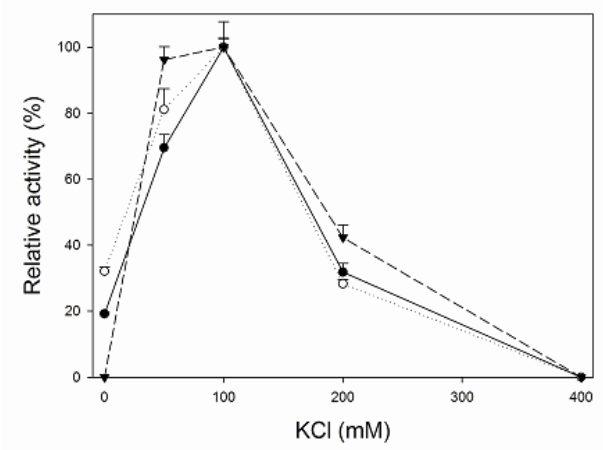


Figure 6-4a

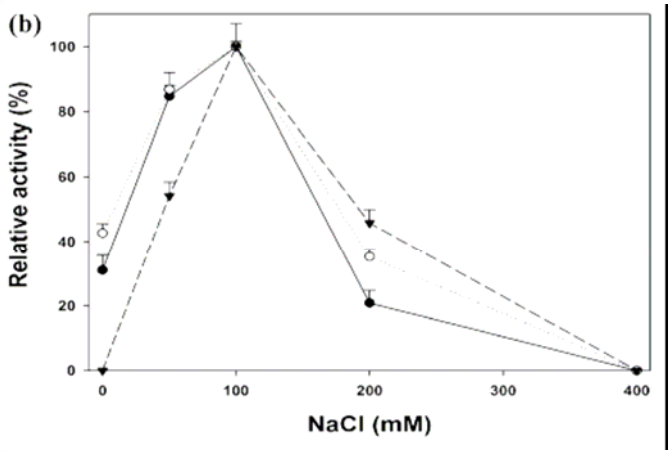


Figure 6-4b

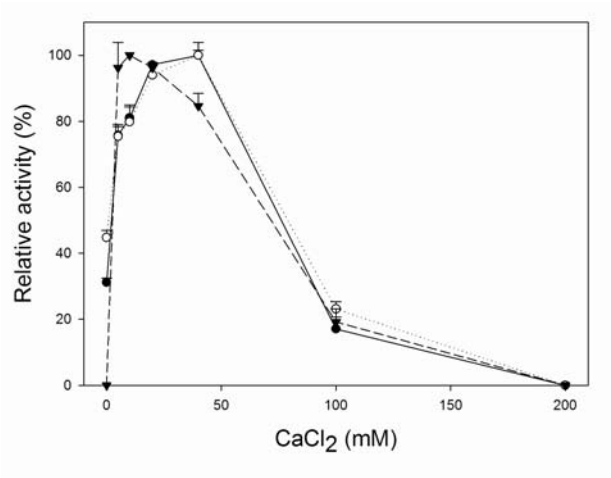


Figure 6-4c

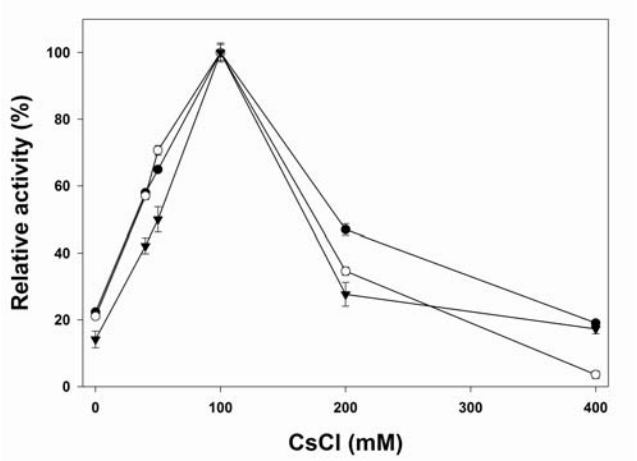


Figure 6-4d

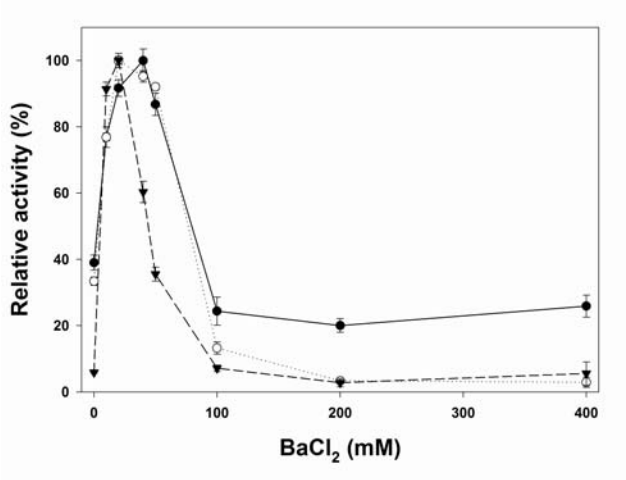


Figure 6-4e

6.5.1.7 Effect of Ca^{2+} on Activity

Having demonstrated the importance of Ca^{2+} for degradation we further investigated the effect of this cation on the activity of BactnABC. The addition of $CaCl_2$ to the reaction mixture had a different effect on CS degradation compared to DS degradation. The activity profile recorded with CS substrates was comparable to the effect observed with monovalent ions. Thus, increasing the concentration of $CaCl_2$ with both CS-A and CS-C from 0 to 0.04 M increased the activity, as in the case of the monovalent ions (Fig. 6-4d). A further increase of $CaCl_2$ concentration reduced the activity, with 50% inhibition at ~ 0.065-0.085 M, and a complete inhibition of the reaction at 0.2 M. The activity with DS displayed a different behaviour as a function of Ca^{2+} concentration, indicating a different role for Ca^{2+} with this substrate. The addition of a low concentration of $CaCl_2$ increased the activity dramatically from a basal level. The maximal activity was reached at 0.01 M $CaCl_2$ (Fig. 6-4d, Table 6-2). Further increasing the $CaCl_2$ concentration resulted in behaviour similar to that with other salts. A similar, specific effect on DS degradation was also observed with Ba^{2+} (Fig. 6-4e).

6.5.2 Kinetic Analysis of Active-Site Mutants

Using alanine-scanning site-directed mutagenesis, we previously identified two groups of amino acids involved in catalysis (Kim *et al.* 1996). In order to characterize the specific catalytic roles of these residues we have investigated the effects of conservative and non-conservative mutations on BactnABC activity. The first group of residues was identified by structural similarity to the active site of chondroitinase AC. Following the nomenclature used for chondroitinase AC (Féthière *et al.* 1999; Huang *et al.* 2001), we termed His454, Tyr461, Arg514 and Glu628 (Fig. 6-5) the catalytic tetrad (Féthière *et al.* 1999). Mutating any of these residues to an alanine was shown to have an acute effect on degradation of both CS and DS substrates (Féthière *et al.* 1999). We prepared the H454D mutant in order to ascertain its role as a general base in catalysis. In parallel, we replaced this histidine with an asparagine or a glutamine to evaluate the importance of hydrogen bonds maintained by either of its nitrogen atoms (ND isosteric to $-NH_2$ of asparagine, and NE isosteric to NH_2 of glutamine) for activity. No detectable activity was observed for any of these mutants (Table 6-3, Fig. 6-6, and Supplementary Fig. 6-S1). Another

residue, Tyr461, was previously proposed to serve as a general base in CS degradation, but not DS degradation (Féthière *et al.* 1999). In addition to the previously reported Y461F mutation that rendered the enzyme inactive toward all substrates, we constructed an Y461A mutant enzyme with the reasoning that a water molecule could enter the vacated space and partially replace the -OH group of the tyrosine sidechain. However, this mutant also showed no activity. The mutations E628D and E628Q were then prepared in order to examine the role of this residue in maintaining the integrity of hydrogen bonding interactions amongst the residues forming the tetrad. The E628D mutant rendered the enzyme inactive; however, an E628Q mutant, retained low levels of activity towards CS substrates but not towards DS (Table 6-3, Fig. 6-6). Previously, we investigated the functional role of His453, which precedes His454. Replacement of His454 by an alanine moderately reduced the activity toward all substrates (Shaya *et al.* 2008a). Surprisingly, the H453N mutant showed no activity toward DS while retaining about 10% of its catalytic efficiency (k_{cat}/K_M) against CS-A and CS-C (Table 6-3, Fig. 6-6). The second group of residues includes the basic patch of Arg172, His344 and His345 (Fig. 6-5). These residues were identified by structural conservation between BactnABC and PvulABCI (Shaya *et al.* 2008). These residues are located on the opposite rim of the substrate-binding site ~12 Å from the tetrad, and are implicated in degradation of DS substrates (Shaya *et al.* 2008). We mutated the two histidines to either aspartates or glutamates in order to ascertain their role in catalysis as potential general bases. Replacing either of these two histidine side chains by acidic residues rendered the enzyme inactive toward all substrates (Table 6-3, Fig. 6-7, and Supplementary Fig. 6-S2). We then replaced each of these histidines by either an asparagine or a glutamine in order to evaluate the importance of their cognate hydrogen bonds in catalysis. While the H344N mutant retained 5-25% catalytic efficiency towards all substrates, the H344Q mutant showed a very low k_{cat}/K_M (~5% compared to wt) only towards CS-A (Table 6-3). His345 mutants behaved differently. Both H345N and H345Q mutants displayed low levels of activity toward CS substrates (2-15% k_{cat}/K_M compared to wt) but no activity toward DS (Table 6-3, Fig. 6-7) suggesting that His345 plays an active role in DS catalysis.

Table 6-3. Kinetic Analysis of BactnABC and Its mutants [a]

	CS-A[b, c]			CS-C[b,d]			DS[b,e]		
Enzyme	*K*M (μM)	*k*cat (min^{-1})	*k*cat/*K*M ($\mu M^{-1}.min^{-1}$)	*K*M (μM)	*k*cat (min^{-1})	*k*cat/*K*M ($\mu M^{-1}.min^{-1}$)	*K*M (μM)	*k*cat (min^{-1})	*k*cat/*K*M ($\mu M^{-1}.min^{-1}$)
WT	67 ± 1	15792 ± 238	237.5	33 ± 3	10404 ± 38	321.1	61 ± 8	2307 ± 190	38.3
H453N	34 ± 1	789 ± 34	23.2	12 ± 0.5	413 ± 17	33.8	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]
H454/D/N/Q	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]
Y461A	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]
E628D	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]
E628Q	**158 ± 25**	**623 ± 11**	**4**	**35 ± 5**	**265 ± 39**	**7.6**	**n.d.**[f]	**n.d.**[f]	**n.d.**[f]
H344D	n.d.[f,g]	n.d.[f,g]	n.d.[f,g]	n.d.[f,g]	n.d.[f,g]	n.d.[f]	n.d.[f]	n.d.[f]	n.d.[ff]
H344E	n.d.[f,g]	n.d.[f,g]	n.d.[f,g]	n.d.[f]	n.d.[f]	n.d.[f]	n.d.[f]	n.d.[f]	n.d.[f]
H344N	39 ± 0.4	2532 ± 70	64.8	84 ± 18	1122 ± 66	13.7	90 ± 5	544 ± 10	6.1
H344Q	73 ± 3	1007 ± 91	13.8	n.d.[f,g]	n.d.[f,g]	n.d.[f,g]	n.d.[f]	n.d.[f]	n.d.[f]
H345D/E	n.d.[f,g]	n.d.[f,g]	n.d.[f,g]	n.d.[f]	n.d.[f]	n.d.[f]	n.d.[f]	n.d.[f]	n.d.[f]
H345N	61 ± 5	1260 ± 40	20.6	26 ± 0.1	396 ± 14	15.1	n.d.[f]	n.d.[f]	n.d.[f]
H345Q	45 ± 4	1408 ± 144	31.3	102 ± 6	651 ± 61	6.4	n.d.[f]	n.d.[f]	n.d.[f]

[a] Mutants of the tetrad residues are in bold font.
[b] Results are the mean ± SD for at least three measurements.
[c] 1 mg ml^{-1} CS-A from bovine trachea (Biocorp) in 50 mM phosphate buffer at pH 7.6, measured for 300 sec at 37 °C using 0.5 μg of enzyme.
[d] 1 mg ml^{-1} CS-C from shark cartilage (sigma) in 50 mM phosphate buffer at pH 7.6, measured for 300 s at 37 °C using 1.0 μg of enzyme.
[e] 1 mg ml^{-1} DS from porcine intestinal mucosa (sigma) in 50 mM phosphate buffer at pH 7.6, measured for 300 sec at 37 °C using 1.5 μg of enzyme.
[f] n.d., not detected.
[g] Showed residual activity on the overnight degradation at 37 °C, but the values were too low for kinetic calculations.

6.7 Discussion

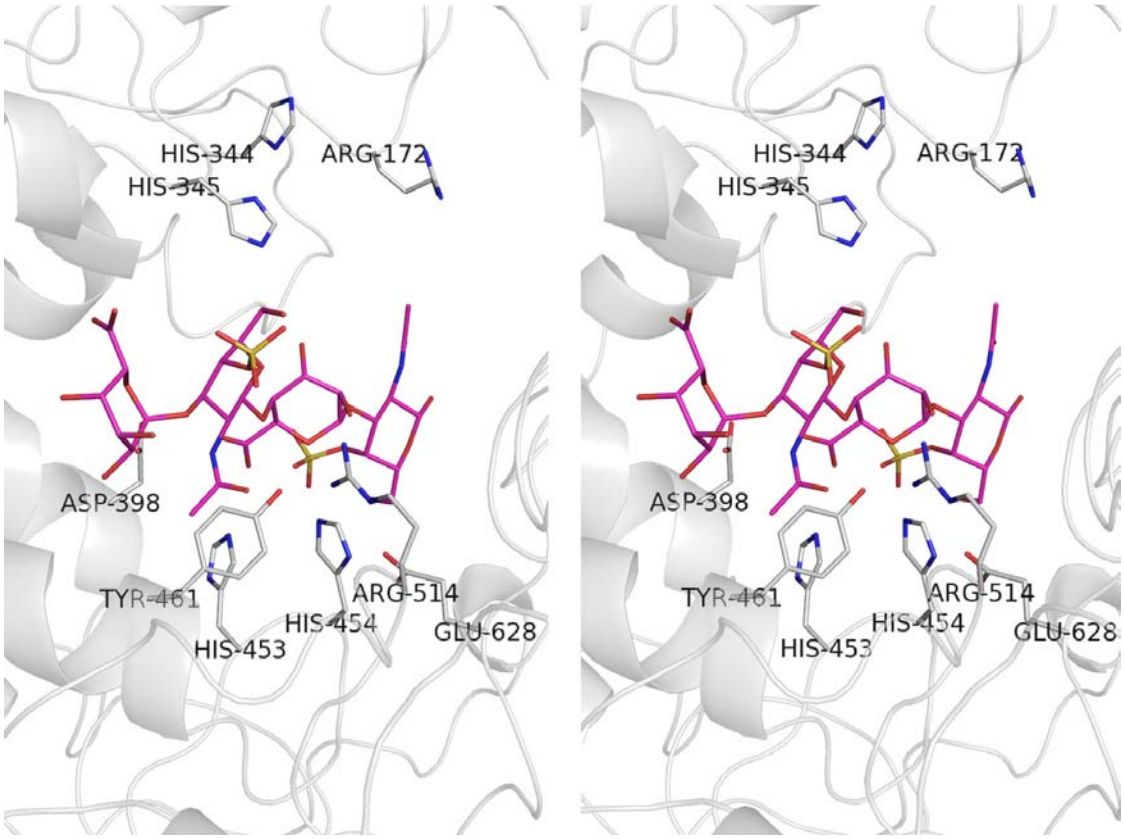
6.7.1 Comparison of BactnABC with *P. vulgaris* Chondroitin Lyases ABC

BactnABC displays low sequence identity with the two isoforms of chondroitinase ABC from *Proteus vulgaris*. It is somewhat more similar to PvulABCII (30% identity) than to PvulABCI (23% identity). The *P. vulgaris* enzymes are well characterized (Hamai *et al.* 1997; Prabhakar *et al.* 2005b; Prabhakar *et al.* 2005c; Prabhakar *et al.* 2006) and have found a variety of applications as analytical tools in identification and analysis of GAG samples. More recently, this class of enzymes showed promise in treatment of spinal cord injury (Bradbury *et al.* 2002; Lemons *et al.* 2003). Therefore, it merits comparing the properties of BactnABC to the *P. vulgaris* enzymes, and investigating the biochemical conditions for their optimal activity *in vitro*.

It is noteworthy that the C-terminal, β-sheet domain of BactnABC also shows a recognizable structural similarity, albeit with low sequence identity, to the C-terminal domains of chondroitinases AC and hyaluronan lyases (Lemons *et al.* 2003).

Clear differences in behaviour between BactnABC and PvulABCI were noted regarding the buffer conditions for their optimal enzymatic activity. While BactnABC displays optimal activity in phosphate buffer, particularly with DS as substrate, PvulABCI shows optimal activity in Tris buffer and has reduced activity toward CS-A in phosphate buffer (Prabhakar *et al.* 2005; Prabhakar *et al.* 2005). The effect of ions for both *B. thetaiotaomicron* and *P. vulgaris* ChonABC enzymes is similar (Prabhakar *et al.* 2005b). Both enzymes attain maximal activity in the presence of 0.1 M monovalent ions, whereas at 0.4 M salt completely inactivates the enzymes.

Figure 6-5. Stereo view of ChonABC's active site
The dermatan sulfate tetrasaccharide substrate is modeled based on structural superposition of the Tyr-His-Glu-Arg catalytic tetrads of BactnABC and *P. heparinus* chondroitinase AC – DS complex (PDB code 1HM2). BactnABC sidechains discussed in this work are depicted as white sticks, DS tetrasaccharide substrate as pink sticks.



Substrate preferences of GAG lyases are usually evaluated using naturally occurring GAGs due to the lack of homogeneously modified polysaccharides (Ernst *et al.* 1995). We have measured the activity of *B. thetaiotaomicron* enzyme using a variety of GAGs from different suppliers and sources. BactnABC and PvulABCI share a similar preference profile toward various GAG substrates, with specific activities decreasing in the order CS-A > CS-C > DS > CS-E, CS-D >> HA (Table 6-1) (Prabhakar *et al.* 2005). Additional information about the preference for 4- or 6-sulfation can be extracted from the differences in specific activities toward similar substrates from different suppliers (Table 6-1). The disaccharides compositional analysis in various GAG substrates shows

clearly a preference of BactnABC for chondroitin with 4-sulfation relative to 6-sulfation, while a non-sulfated dermatan is preferred.

In comparison to PvulABCI, the respective activities of BactnABC are lower toward DS, higher toward CS-E, and lower toward HA. BactnABC depolymerization of GAGs yielded only disaccharide products, indicating an exolytic mode of action, as was shown for PvulABCII (Hamai *et al.* 1997). In contrast, the products released by the endolytic PvulABCI were a mixture of tetrasaccharides and disaccharides (Prabhakar *et al.* 2005). Thus, the higher sequence similarity of BactnABC to PvulABCII and the shared mode of action show that these two enzymes are functional homologs.

Since *B. thetaiotaomicron* resides in the human gut, the enzyme's optimum activity toward all substrates at 37^{o}C was expected. BactnABC displays higher thermal stability and activity at temperatures above the optimal 37^{o}C, as compared to PvulABCI. The enzymatic activity of BactnABC was unaffected by incubation at temperatures as high as 40 ^{o}C. It retains significant activity toward CS-A, DS and CS-C at 45^{o}C (Fig. 6-2a), however, the activity drops down with pre-incubation, indicating partial denaturation at this temperature. PvulABCI retains only 10% of its maximal activity toward CS-C and less than 40% activity toward DS at that temperature (Prabhakar *et al.* 2005). BactnABC displays higher apparent T_m for thermal denaturation of 44^{o}C as apposed to 38^{o}C measured for PvulABC.

The narrow pH-activity profile of BactnABC reflects the importance of the protonation state of ionizable groups identified in the active site, including His344, His345, His453, His454, and Tyr461. These residues might participate in general acid/base catalysis or take part in precise hydrogen-bonding networks important for correct positioning of active site residues. Both roles would be affected by the protonation state of these amino acids as demonstrated by the His344 and His345 mutant enzymes. Furthermore, the bell shape of the pH profile attests for the general base/acid catalytic functions discussed below for the histidine and tyrosine side-chains. The marked decrease in activity below pH 6.5 is consistent with the need for deprotonated histidine side chain acting as a general base whereas the low activity recorded above pH 8.5 is consistent with the importance of a protonated tyrosine acting as a general acid.

6.7.2 Toward Understanding of the Catalytic Mechanism

Of the five divalent cations tested, Ca^{2+}, Mg^{2+} and Ba^{2+} increase the specific activity of BactnABC toward all substrates, whereas Zn^{2+} and Mn^{2+} decrease the activity against CS while mildly increasing the activity toward DS. Ca^{2+}/Ba^{2+} cations affect the catalytic activity towards CS in a different manner than towards DS. The effect on CS is predominantly related to ionic strength in a fashion similar to that exerted by the monovalent salts (Fig. 6-4). On the other hand, the activity toward DS is dramatically augmented in the presence of divalent cations, with maximal effect at a much lower ion concentration, indicating selective differences in the effect of the metal in the degradation of these two substrates. This behaviour is reminiscent of the narrower specificity lyases, chondroitinase AC and chondroitinase B. These two enzymes present different three-dimensional folds and catalytic site arrangements (Huang *et al.* 1999; Féthière *et al.* 1999; Michel *et al.* 2004; Lunin *et al.* 2004). Chondroitinase AC requires no divalent cations to degrade CS, while chondroitinase B requires the presence of Ca^{2+} to break down DS. The Ca^{2+} is coordinated between the enzyme and the iduronic acid moiety of DS, neutralizing the negative charge of the acidic group (Michel *et al.* 2004). We propose that BactnABC requires a metal cation for binding DS, whereas it is not necessary for the binding of CS. Furthermore, we propose that BactnABC uses the metal to neutralize the negative charge of the IdoA carboxylate moiety, as has been found for DS in chondroitinase B. This effect is metal-specific, as neither Zn^{2+} nor Mn^{2+} enhance the activity significantly towards DS.

The β-elimination reaction mechanism requires charge neutralization of the negative carboxylic group of the uronic acid in order to reduce the p*Ka* of the C5-bound proton (Gacesa 1987). The next step of catalysis involves a residue acting as a general base, triggering the elimination by removal of the C5 proton from the uronic acid with the formation of a C4-C5 double bond in the uronic acid ring. Concomitantly or subsequently, an amino acid acting as a general acid protonates the bridging oxygen between the uronic acid and glucosamine thus breaking the glycosidic bond, completing the degradation and releasing the products (Gacesa 1987).

We have previously postulated that substrate binding involves a conformational rearrangement, bringing the tetrad and residues on the opposite rim of the binding site

into proximity creating one composite active site with different sidechains participating in the consecutive reaction steps for different substrate (Shaya 2008).

In addition to the alanine replacements other, more conservative mutations of the tetrad components inactivated the enzyme towards all substrates. An exception is the E628Q mutant, which retained a low catalytic efficiency toward CS substrates. The low tolerance for chemical changes in the tetrad implies the importance of a precise hydrogen-bonding network between the residues comprising the tetrad. The specific interactions between those residues are vital for correct placement of the chemical groups necessary for the degradation to take place. The acute influence of the tetrad residues on the degradation of both substrates indicates that these residues take part in the final step of degradation – proton donation and product release, which we propose is shared by both substrates. In the case of CS degradation, the tetrad residues are proposed to function in the first step of degradation as well, paralleling chondroitinase AC (Lunin *et al.* 2004). The lack of activity for all mutants of either His454 or Tyr461 does not permit distinguishing which residue acts as the general base abstracting the C5 proton from the glucuronic acid of CS. The low catalytic efficiency of the E628Q mutant is in agreement with the chondroitinase AC crystal structure (Lunin *et al.* 2004), which showed that Glu628 is important for positioning of Arg514 and His454 rather than being directly involved in catalysis. The replacement of Glu628 by a glutamine would preserve the hydrogen bonds to Arg514 and His454, although to the latter only in its neutral state. The generally low activity of BactnABC in the acidic range may therefore be related in part to the protonation of this glutamate residue, mimicking glutamine in this position.

The H453N mutant shows no activity toward DS, while retaining about 10% of its catalytic efficiency against CS-A and CS-C. This decrease is mostly related to a reduction of k_{cat} for these substrates. Comparing the BactnABC crystal structure (PDB code 2Q1F) to chondroitinase AC – CS tetrasaccharide complex (PDB code 1RWH) we find the side chain of His453 located close to the predicted location of the uronate carboxylic group. We suggest that this histidine is involved in charge neutralization for CS and DS substrates, either directly, by hydrogen bonding to the carboxylic group in CS degradation, or indirectly, by coordinating the Ca^{2+} (possibly with Asp398) in DS degradation.

We have previously demonstrated by alanine mutations that His344 and His345 play key roles in DS degradation but are not absolutely essential for degradation of CS (Shaya 2008). Replacement of His345 by either asparagine or glutamine had a similar effect as the alanine replacement – the mutants are inactive toward DS while the catalytic efficiency toward CS decreases to 5-10% of the wild type enzyme. Therefore, a histidine sidechain with its potential to abstract a proton, rather than asparagine or glutamine, with nitrogen atoms isosteric to the histidine nitrogen atoms, is essential for DS degradation. His344 plays a different role. The H344N mutant retains ~5-25% catalytic efficiency toward all substrates while H344Q displays low levels of activity only towards CS-A but not towards CS-C or DS. Therefore, asparagine but not glutamine can partially replace His344, indicating that the ND nitrogen of His344 is important. Indeed, in the crystal structure (PDB code 2Q1F) this nitrogen atom is involved in a hydrogen bond to His345, likely anchoring the latter in a proper orientation for catalysis. Finally, mutation of either of these two histidines to an acidic residue, aspartate or glutamate, inactivates the enzyme toward all substrates. This substitution introduces a negative charge in the substrate-binding site and most likely adversely affects binding of the highly negatively charged substrates. Together, the kinetic properties of the site-specific mutants suggest that with CS as a substrate these two histidines participate in substrate binding, whereas for DS His345 serves as the catalytic base and His344 is essential for maintaining the correct positioning and specific protonation of His345.

Our results provide additional support for the notion of a composite active site in BactnABC, allowing for degradation of GAGs possessing both uronic acid epimers. The residue that performs the first step of degradation, proton abstraction from the C5 carbon, is different for CS (GlcA-containing) than for DS (IdoA-containing) substrate as a result of the different configuration of the substituents of the C5 carbon. In contrast, proton donation to the leaving group involves a planar intermediate common for both epimers, and consequently the same residue can carry out the reaction with either substrate. The tetrad participates in both acid and base catalysis for CS but only in donating a proton to the leaving group with DS as substrate. His345 from the rim of the substrate-binding cleft opposite to the tetrad was identified as the most likely candidate for proton abstraction of DS substrate. In addition, we provide a comprehensive biochemical characterization of

BactnABC establishing optimal conditions for *in vitro* degradation of CS-A, CS-C and DS.

Figure 6-6. Product profile analysis using SAX-FPLC
BactnABC mutants of the tetrad residues H454N, E628Q, and H453N acting on: **a)** Biocorp CS-A, **b)** Sigma CS-C and **c)** Sigma DS. Identified disaccharides marked as in figure 6-3.

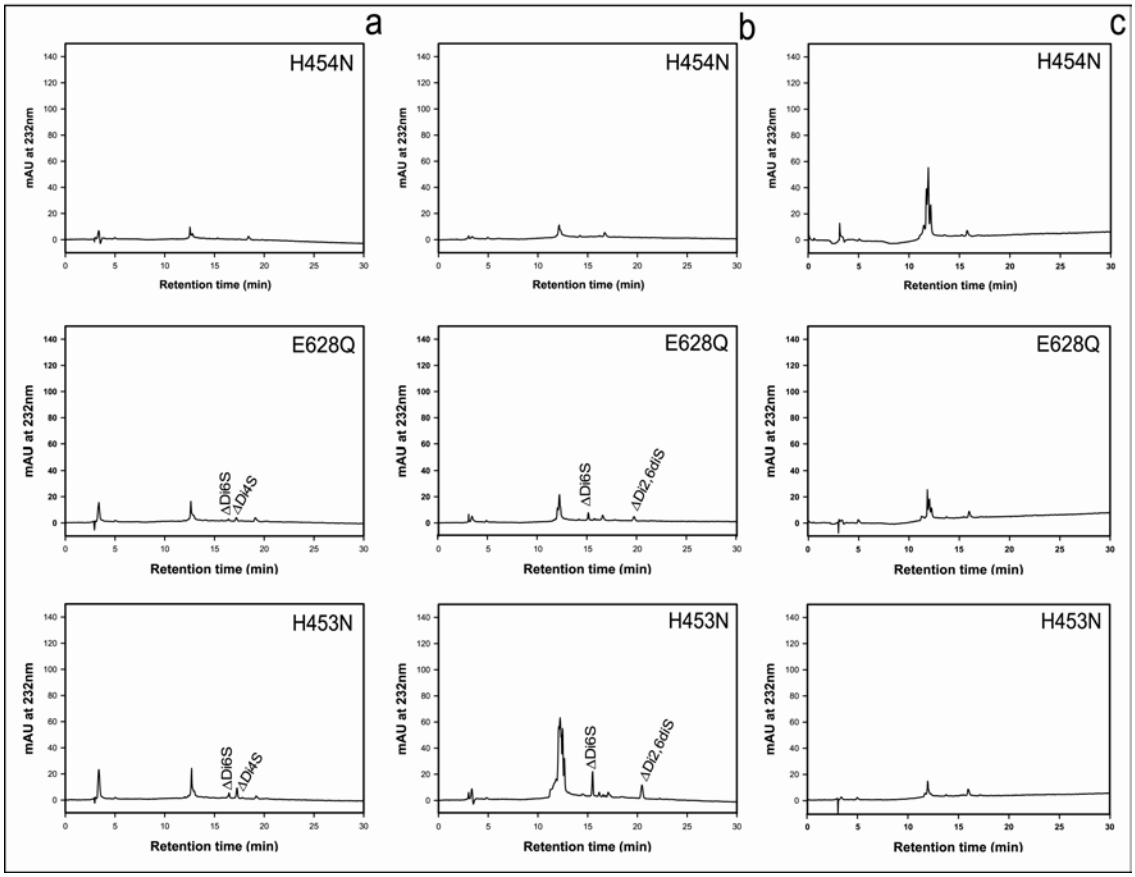
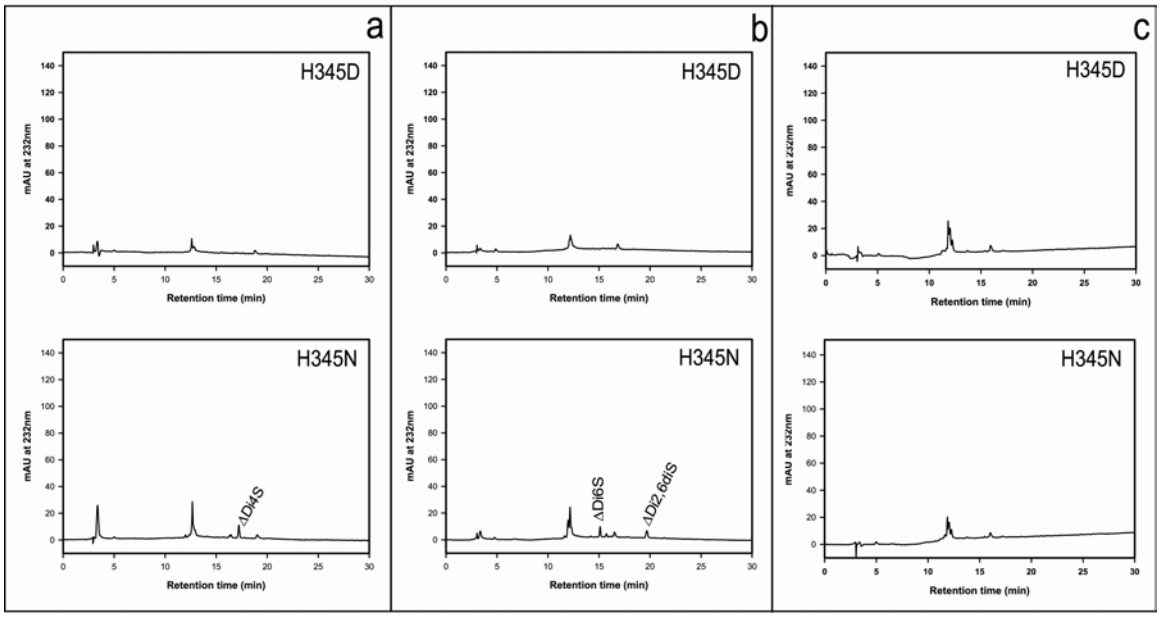


Figure 6-7. Mutants of the second active site
H345D/N acting on: **a)** Biocorp CS-A, **b)** Sigma CS-C and **c)** Sigma DS. Identified disaccharides marked as in Figure 6-3.



6.8 Acknowledgements

We would like to thank Drs. Holger Lindner and Allan Matte for their helpful comments preparing this manuscript.

6.3 Appendix A – Supplementary Material

Table 6-S1. Sequence of BactnABCI.
Amino acids differing from the NCBI sequence highlighted in grey.

```
  1    MLILSFLCPAFLNAQIVTDERMFSFEEPQLPACITGVQSQLGISGAHYKDGKHSLEWTFEPN
 63    GRLELRKDLKFEKKDPTGKDLYLSAFIVWIYNEQPQDAAIEFEFLKDGRKCASFPFGINFKG
125    WRAAWVCYERDMQGTPEEGMNELRIVAPDAKGRLFIDHLITATKVDARQQTADLQVPFVNAG
187    TTNHWLVLYKHSLLKPDIELTPVSDKQRQEMKLLEKRFRDMIYTKGKVTEKEAETIRKKYDL
249    YQITYKDGQVSGVPVFMVRASEAYERMIPDWDKDMLTKMGIEMRAYFDLMKRIAVAYNNSEA
311    GSPIRKEMRRKFLAMYDHITDQGVAYGSCWGNIHHYGYSVRGLYPAYFLMKDVLREEGKLLE
373    AERTLRWYAITNEVYPKPEGNGIDMDSFNTQTTGRIASILMMEDTPEKLQYLKSFSRWIDYG
453    CRPAPGLAGSFKVDGGAFHHRNNYPAYAVGGLDGATNMIYLFSRTSLAVSELAHRTVKDVLL
496    AMRFYCNKLNFPLSMSGRHPDGKGKLVPMHYAIMAIAGTPDGKGDFDKEMASAYLRLVSSDS
559    SSAEQAPEYMPKVSNAQERKIAKRLVENGFRAEPDPQGNLSLGYGCVSVQRRENWSAVARGH
621    SRYLWAAEHYLGHNLYGRYLAHGSLQILTAPPGQTVTPTTSGWQQEGFDWNRIPGVTSIHLP
683    LDLLKANVLNVDTFSGMEEMLYSDEAFAGGLSQGKMNGNFGMKLHEHDKYNGTHRARKSFHF
745    IDGMIVCLGSDIENTNMDYPTETTIFQLAVTDKAAHDYWKNNAGEGKVWMDHLGTGYYVPVA
807    ARFEKNFPQYSRMQDTGKETKGDWVSLIIDHGKAPKAGSYEYAILPGTDRKTMTAFAKKPAY
869    SVLQQDRNAHILESPSDRITSYVLFETPQSLLPGGLLQRTDTSCLVMVRKESADKVLLTVAQ
931    PDLALYRGPSDEAFDKDGKRMERSIYSRPWIDNESGEIPVTVTLKGRWKVVETPYCKVVSED
993    KKQTVLRFLCKDGASYEVELEK
```

Table 6-S2. Summary of primer sequences for site-directed mutagenesis. Mutation triplet codons are indicated in bold; bases modified in order to create the desired point mutations are underlined. Forward primers for each mutant are provided.

Mutant	Primer Sequence
His^{344}→Asp	5' GTTGCTGGGGCAATATT**G**<u>**A**</u>**T**CATTACGGATATAGCG 3'
His^{344}→Glu	5' GTAGTTGCTGGGGCAATATT**G**<u>**A**</u>**G**CATTACGGATATAGCGTGCG 3'
His^{344}→Asn	5' GTTGCTGGGGCAATATT<u>**A**</u>**A**TCATTACGGATATAGCG 3'
His^{344}→Gln	5' GCTGGGGCAATATT**CA**<u>**G**</u>CATTACGGATATAGCG 3'
His^{345}→Asp	5' GCTGGGGCAATATTCAT<u>**G**</u>**A**TTACGGATATAGCGTG 3'
His^{345}→Glu	5' GTTGCTGGGGCAATATTCAT<u>**G**</u>**A**<u>**G**</u>ACGGATATAGCGTGCGCG 3'
His^{345}→Asn	5' GCTGGGGCAATATTCAT<u>**A**</u>**A**TTACGGATATAGCGTGC 3'
His^{345}→Gln	5' GGGGCAATATTCAT**CA**<u>**G**</u>TACGGATATAGCGTGC 3'
His^{453}→Asn	5' GGACGGAGGAGCTTTC<u>**A**</u>**A**TCACCGTAATAACTAC 3'
His^{454}→Asp	5' GGACGGAGGAGCTTTCCAT<u>**G**</u>**AC**CGTAATAACTACCCTGCC 3'
His^{454}→Asn	5' GGACGGAGGAGCTTTCCAT<u>**A**</u>**AC**CGTAATAACTACCCTGCC 3'
His^{454}→Gln	5' GAGGAGCTTTCCAT**CA**<u>**G**</u>CGTAATAACTACCCTGC 3'
Tyr^{461}→Ala	5' GTAATAACTACCCTGCC<u>**GC**</u>**T**GCCGTTGGAGGACTGG 3'
Glu^{628}→Asp	5' CCTTTGGGCAGCC**GA**<u>**T**</u>CATTATCTGGGACAC 3'
Glu^{628}→Gln	5' CGCTACCTTTGGGCAGCC<u>**C**</u>**AA**CATTATCTGGGACAC 3'

Figure 6-S1. Product profile analysis using SAX-FPLC

BactnABC mutants of the tetrad residues - H454D/Q, Y461A, and E628D that presented no activity toward all substrates acting on: **a)** BioCorp CS-A, **b)** Sigma CS-C and **c)** Sigma DS. Disaccharides identified: ΔDi0S = ΔUA-GalNAc; ΔDi6S = ΔUA-GalNAc6S; ΔDi4S = ΔUA-GalNAc4S; ΔDi2,6diS = ΔUA2S-GalNAc6S; ΔDi4,6diS = ΔUA-GalNAc4S6S; ΔDi2,4diS = ΔUA2S-GalNAc4S.

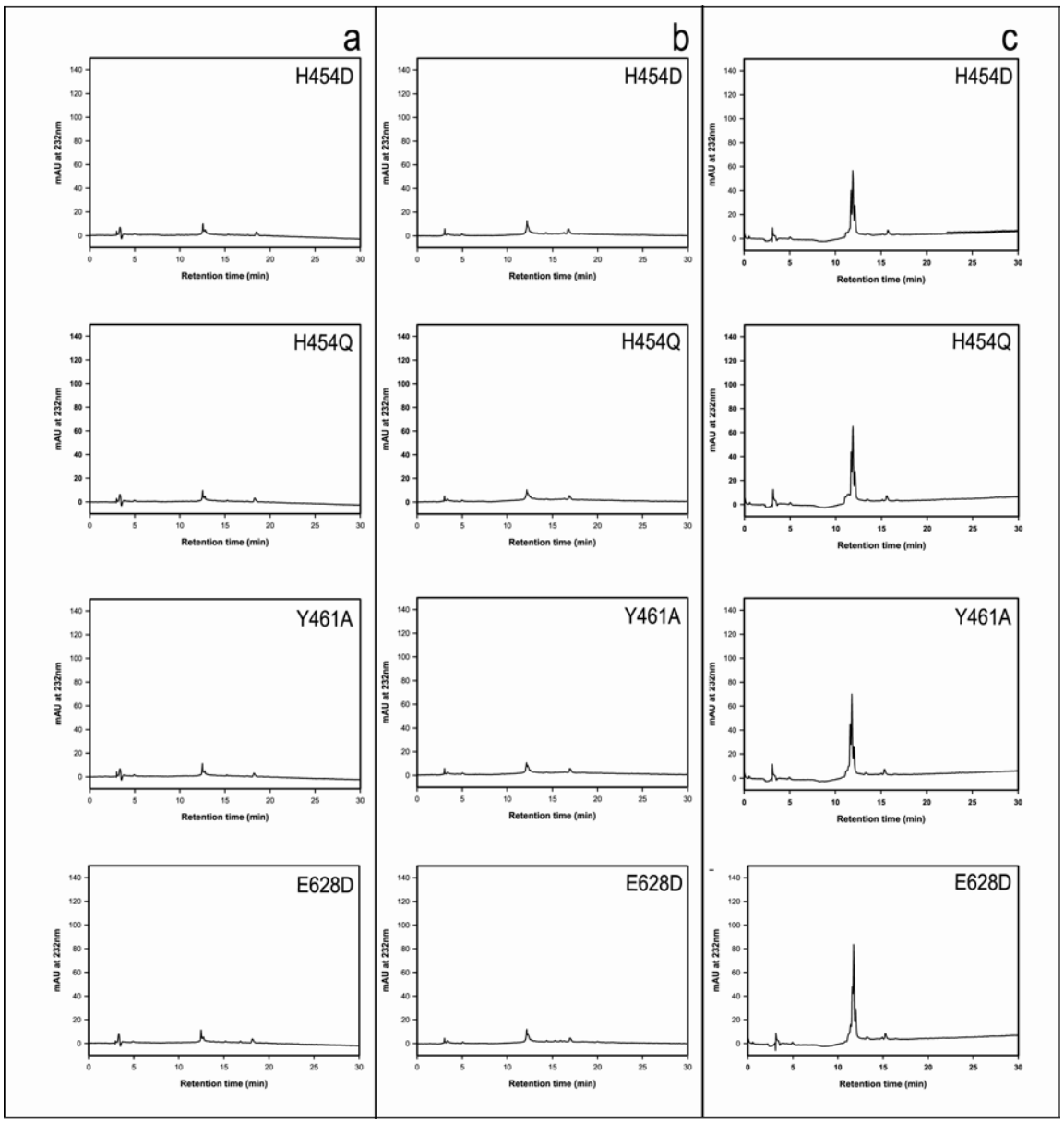
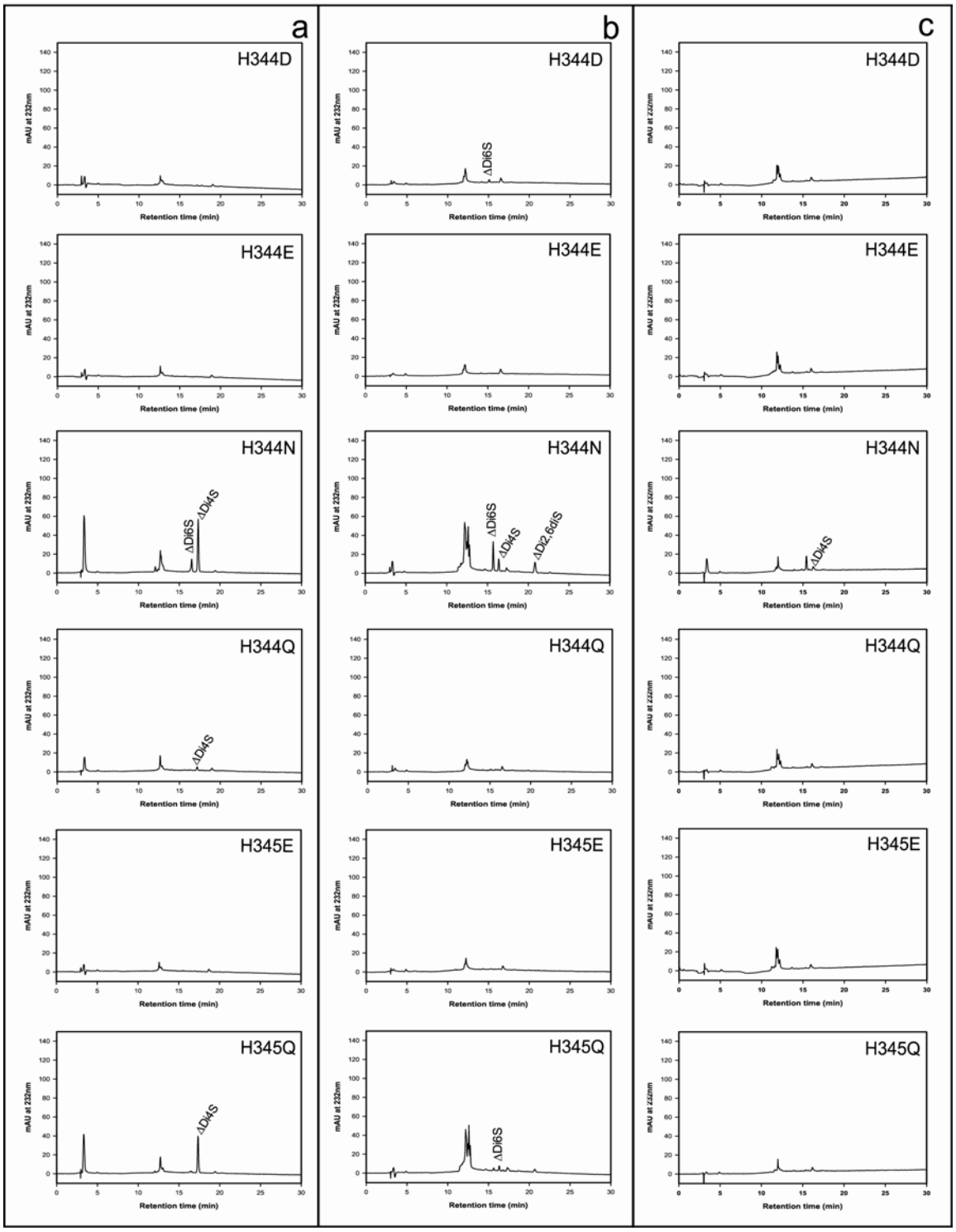


Figure 6-S2. Product profile analysis using SAX-FPLC for BactnABC

Mutants of the identified second active site - H344D/E/N/Q, and H345E/Q acting on: **a)** BioCorp CS-A, **b)** Sigma CS-C and **c)** Sigma DS. Disaccharides identified: ΔDi0S = ΔUA-GalNAc; ΔDi6S = ΔUA-GalNAc6S; ΔDi4S = ΔUA-GalNAc4S; ΔDi2,6diS = ΔUA2S-GalNAc6S; ΔDi4,6diS = ΔUA-GalNAc4S6S; ΔDi2,4diS = ΔUA2S-GalNAc4S.



Chapter 7 Summary and Conclusions

7.1 Summary of the Experimental Results

7.1.1 Heparinase II

The first question this research has answered regarding HepII has been to provide a structural description of its overall topology and architecture. HepII was shown to form dimers, with each monomer comprised of the α+β lyase fold found in PL8 family enzymes such as chondroitinase AC, chondroitinase ABC, and hyaluronan lyase. Each monomer of HepII contains three subdomains, an N-terminal $(\alpha/\alpha)_6$ helical toroid, a double-layered β-sandwich central domain and a C-terminal small β domain. The N-terminal toroidal fold is found as well as a stand-alone one-active-domain enzyme in alginate lyases belonging to family PL5 and pectate lyases from family PL10. Similarly to family PL8 enzymes, the HepII structure contains a surface exposed, positively charged cleft traversing the protein, where the substrate-binding site has been identified. On the other hand, HepII differs in the disposition and orientation of the three domains from family PL8 proteins. It differs in the relative location of its active site cleft with regard to the three sub-domains and its unique structural Zn^{2+} ion. The Zn^{2+} ion was found at the interface between the N-terminal and central domains, maintaining their relative orientation. This ion is primarily bound by the extended loops belonging to the central β-sheet domain that create one rim of the substrate-binding cleft. In contrast, family PL8 proteins lack the metal ion in this position; however, their active site clefts are created at the top of the incomplete toroidal domain and contain residues belonging primarily to this domain. HepII also differs from the PL8 lyases in its C-terminal small domain, which forms a β-barrel 'lectin-like' fold. The β-strands counterparts of this domain in PL8 lyases present somewhat different conctivity and a different topology, stacking more closely to the central domain and constitute a continuation of this domain.

The second question this research has answered has been associated with the broad specificity of HepII with regard to the sulfate modifications, as well as its mode of action. The surface-exposed cleft harboring the active site of HepII is divided into two parts, a narrow part with the capacity of binding 3-4 monosaccharides, containing the "+" subsites, and a wider part harboring the "–" subsites. From the structures of HepII-oligosaccharide complexes, it is clear that the "+" subsites represent the primary oligosaccharide-binding site whereas the "–" subsites represent a less specific binding

site, with fewer direct contacts between the protein and the oligosaccharide. The structures revealed another interesting feature, namely, sugars in the "+" subsites bound with two of their sulfate groups (position 2 of the GlcA and 6 of GlcNS) pointing toward the solvent, maintaining few interactions with the protein, mainly through water molecules. This explains the lack of specificity displayed by HepII for the substrate sulfation state. The monosaccharides in the "–" subsites are bound with their N-acetyl group pointing towards the solvent (opposite to the "+" subsites); however, the cleft at this location is very wide and can easily accommodate sulfate groups. At position "–1", a depression formed by the sidechains of Arg148, Trp126 (Nε1), and His347, that is likely to provide a binding pocket for the possible 6-*O* sulfation of the glucosamine.

The overall structure of the enzyme shows an open cleft that can accommodate 2-3 disaccharide units on either side of the catalytic site, allowing easy access for substrate binding and product release. This structure of the cleft seems to be perfectly adapted to the endolytic, non-processive mode of action exhibited by HepII.

This study also yielded a similarity in the strategies for substrate binding used by the family PL8 enzymes, as represented by the structure of chondroitin lyase AC from *A. aurescens* complexed with its substrate, (PDB code 1RWG; (Lunin *et al.* 2004)) and HepII, namely, the bending of the oligosaccharide backbone around the scissile glycosidic bond that contributes to a higher energy conformation and torsion strain on the uronic acid at position "+1". It has been argued that this higher energy conformation eases the elimination of the glycosidic bond. Although the structure of HepII H202A/Y257A mutant with the bound substrate revealed only bending of the polysaccharide backbone around the sessile bond but not the higher energy boat conformation of the uronate in position "+" (as found in the ChonAC-substrate complex structure), it is clear that the reason for this is the void created by the Tyr257Ala mutation, and I postulated that binding to the native enzyme involves distortion of uronic acid from the chair conformation.

Finally, I have addressed the issue of the uronic acid epimer-independent activity of HepII by demonstrating that the enzyme has adapted a unique, symmetric arrangement around the chiral reaction center, the C5 carbon. I have demonstrated that a low energy-barrier hydrogen bond, maintained between one carboxylate oxygen atom and Glu205, is

responsible for the negative charge neutralization. This hydrogen bond is characterized by sharing of the proton between the two oxygen atoms involved, thus achieving protonation of the carboxylate oxygen. On the other side of the carboxylate, the second oxygen atom is hydrogen bonded to His406, an interaction that has been proven to be highly important for the catalytic activity of HepII by the lack of activity presented by the His406Ala mutant in the real-time assay. Residual activity displayed using the over-night saturation assay demonstrates that His406 works in an environment of active side-chains, and that even in its absence, the enzyme is capable of low-level degradation of the substrate. Furthermore, I have shown that to overcome the second stereochemical hurdle, namely abstracting a proton from either side of the chiral center, HepII has adopted two-fold symmetry around C5 of the uronate, placing His202 and Tyr257 at equal positions from the chiral center (Fig. 4-2). Both residues are proposed to serve as catalytic bases, His202 for the degradation of IdoA-containing glycosidic bonds (typically found in heparin) and Tyr257 in a deprotonated state abstracting the proton from the GlcA epimer that is more commonly found in HS. For the general acid role, I propose Tyr257 as the functional residue. Since both substrates lead to a planar intermediate containing C4-C5 double bond at the uronate, stereo-chemically, those intermediates are similar around C5, and hence in both scenarios Tyr257 can donate a proton to the leaving group.

Recently, we solved the structure of the gut mutualist *B. thetaiotaomicron* HepI and its complex with two disaccharide degradation products (Garron M.L., Shaya D., and Cygler M., unpublished data). HepI structures further support the assignment of catalytic roles for the active site residues of HepII proposed in the text above and in Chapter 5. HepI displays markedly different architecture (β-jelly roll containing a single domain), molecular weight (34.5 kDa), substrate specificity (stringent specificity, depolymerizing only the tri-sulfated glycosidic bond –[IdoA2s (1→4) GlcNS6S]), and a different mode of action (processive exolytic enzyme) from HepII. Nevertheless, a structural overlay of the catalytic sites of both enzymes reveals remarkable similarity. The amino acid differences between the two active sites stress the activity HepII possesses toward glycosidic bonds containing GlcA, which HepI is devoid of (Fig. 4-4). Thus, the catalytic residues involved in degradation of IdoA-containing glycosidic bonds in HepII, His202 and Tyr257, superpose with His151 and Tyr357 in HepI. His 406 (HepII) is replaced by

the side chain of Arg83, excluding binding of GlcA since it is a longer and bulkier residue, and occupies the space that the C5-bound proton of GlcA in HepII is proposed to occupy. Interestingly, HepI utilizes a glutamine residue whereas HepII uses Glu205 for performing charge neutralization. The same phenomenon was observed comparing ChonAC to ChonABC. The epimer-specific ChonAC utilizes an asparagine side chain whereas the epimer non-specific ChonABC uses an aspartate. This later point is discussed in more detail in the following text.

7.1.2 Chondroitin Lyase ABC

The crystal structure of ChonABC form *B. thetaiotaomicron* revealed a three-dimensional architecture typical for PL8 family enzymes, i.e., a α+β lyase fold similar to the published structure of the *P. heparinus* ChonABCI (Huang *et al.* 2003). The enzyme adopts a "W" shape and is composed of three structural domains. The N-terminal domain adopts a two-layered β-jellyroll fold denoted commonly as a carbohydrate binding module in enzymes such as cellulase and endo 1,4-α-D-xylanases both contain a structural Ca^{2+} ion located in an equivalent location as identified in BactnABC. The central domain, a double-layered $(\alpha/\alpha)_5$ incomplete toroid, is structurally similar to the catalytic domains of PL8 lyases such as *P. heparinus* and *A. aurescens* ChonAC, hyaluronate lyases, xanthan lyase, alginate lyases, and *P. heparinus* HepII. The C-terminal domain contains a stack of four antiparallel β-sheets, contributes Glu628 to the catalytic tetrad, and superposes very well with PL8 lyase structures.

A noticeable difference between ChonABC and the other lyases is found in the dimensions of the substrate-binding cleft located in the central domain. ChonABC displays a wider, more open cleft in comparison to PL8 lyases. Site-directed mutagenesis, combined with the kinetic analysis presented in Chapters 5 and 6, together support the hypothesis put forward in this part of the dessertation. I have demonstrated that BactnABC undergoes a conformational change upon forming the ternary complex with the polysaccharide substrate and the metal ion, which narrows the cleft, reducing the distance between its opposite rims. The mutagenesis data show that amino acids from both rims participate in binding and degradation of both CS and DS.

The chemical roles attained by the catalytic tetrad residues during degradation of CS were assigned based on structural comparison with *A. aurescens* ChonAC, namely Tyr461, His454, Arg514 and Glu628. The proposed mechanistic scheme identifies Tyr461, in a deprotonated state, as the catalytic base abstracting the C5-bound proton from GlcA and as the catalytic acid donating the proton to the glucosamine leaving group. The possibility of His454 serving as a catalytic base could not be ruled out by our data, although the structure of *A. aurescens* ChonAC complexed with its substrate suggests that the parallel histidine residue in ChonAC participates in charge neutralization *via* hydrogen bonding with the carboxylate ion. The kinetic analysis demonstrates that Glu628, as in ChonAC, plays a part in positioning both His454 and Arg514 in a precise orientation necessary for effective polymer degradation, but is not directly involved in catalysis. The arginine side chain of the tetrad was shown, in ChonAC, to bind the glycosidic oxygen and stabilize the negative charge on this oxygen during formation of the reaction intermediate.

The role of charge neutralization is more difficult to assign unambiguously at this point. His453, preceding His454 of the catalytic tetrad, abolishes DS activity when mutated to asparagine, while reducing CS degradation to 10% compared to the wild type. Structural comparison to ChonAC suggested that this side chain resides in the vicinity of the substrate carboxylate and we proposed that this side chain participates in charge neutralization, either by direct binding to the substrate's carboxylate or indirectly by binding a metal ion. The residue directly involved in binding and neutralization of the substrate carboxylate in ChonAC, asparagine, finds its counterpart in the structure of BActnABC in the form of Asp398. My results indicate the importance of metal ion binding, mainly for the degradation of the IdoA-containing DS, which can be argued, is coordinated by this aspartate and His453. A recent report, from the laboratory of R. Sasisekharan, identifies the Ca^{2+} chelating residues in *P. vulgaris* ChonABCI as Asp444 (Asp398 in BactnABC), Asp442 (Asp396 in BactnABC), and the oxygen atom of either or both of Tyr392 (Ser348 in BactnABCII) or Asn447 (Asn401 BactnABC) (Prabhakar 2006). From the point of view of conservation between the two ChonABC structures, Tyr392 can be ruled out, leaving the two aspartate residues and the asparagine as

candidates to comprise the coordination sphere surrounding the Ca^{2+}, along with the carboxylate oxygen atoms.

By structural comparison with *P. vulgaris,* I identified a basic patch of residues located on the rim of the cleft opposite to the catalytic tetrad and located 12 Å away from them. Kinetic analysis led me to propose that His345 of this patch serves as that catalytic base, abstracting the proton from DS while its preceding His344 plays an important role in orienting His345 and maintaining the protonation state of its nitrogen atoms. As well, I have identified an arginine residue from this patch that, when mutated to Ala, abolishes the activity completely toward CS as well as DS. This side chain is most likely involved in binding and recognition of both CS and DS by BactnABC, and plays an important role in maintaining its closed conformation, which may occur upon substrate binding.

Our results presented in Chapter 6 clearly identify BactnABC as an exolytic enzyme removing dissacharide units, one at a time, from the polysaccharide ends. As well, we established the optimal conditions for its *in vitro* catalysis as 37°C, pH 7.6 in phosphate buffer. Addition of 0.1 M of various salts increased the activity toward all substrates. Finally, we demonstrated that BactnABCII displays higher temperature stability than PvulABCI.

7.2 Conclusions

This dissertation presents the structure-function studies of two GAG degrading enzymes, namely heparin lyase II form *P. heparinus* and chondroitin lyase ABC form *B. thetaiotaomicron*. The overall structural architecture of both enzymes display known folds with variation in the fine details, conferring their unique activities. HepII and ChonABCII present architecture typical for PL8 lyases, composed of an α+β three domain arrangement, with a wide positively charged cleft transversing the protein surface. This cleft harbors the substrate binding / active site.

A closer inspection revealed that both proteins utilize one active site to degrade all the possible glycosidic bonds related to their activity. Their general tolerance toward different sulfation patterns in their substrates is achieved by orienting the polysaccharides such that the majority of the sulfo groups are accommodated at the opening of the cleft.

These sulfo modifications maintain few hydrogen bonds to the protein, mostly mediated by water molecules.

A thorough analysis of the active site revealed that both proteins have adapted to cleave the glycosidic bonds on the non-reducing and of either uronic acid epimer, be that GlcA or IdoA, by placing basic residues in a symmetric arrangement around the chiral center of the uronic acid, carbon 5. Upon the abstraction of the relatively acidic proton bound to the chiral center, a C4-C5 double bond with sp^2 planar hybridization is formed. This intermediate has no chirality and affords the utilization of the same residues for the completion of the β-elimination.

The structure of HepII displayed only one amino acid, His202, located on the rim of the active site responsible for the IdoA containing substrate depolymerization. In a similar fashion, BactnABCII utilizes His345 for the degradation of DS. The structures of the binding site reveal that both display the majority of the amino acids responsible for the specificity and degradation of the GAG substrates on one rim of the cleft, the rim oriented toward the C5 bound proton of GlcA. This observation, combined with the lower rates of catalysis for the IdoA-containing linkages displayed by both enzymes might indicate a possible origin for these enzymes, originating from GlcA specific enzymes (like ChonAC for example), later evolved to degrade IdoA-containing substrates.

Depolymerization of DS by BactnABC is greatly enhanced by the addition of divalent metal cations. The presence of divalent metal ions was argued to play a crucial role maintaining the orientation of the flexible IdoA containing substrates in the active site. Furthermore, the β-elimination performed by ChonB, several pectate lyases, and few alginate lyases was demonstrated to relay on metal divalent cations for charge neutralization, the catalytic step that triggers the elimination. Although HepII contains a glutamate side chain (commonly found in metal coordination spheres in protein structures) in close contact of the uronic acid carboxylate group, the structure does not indicate any free space for divalent cation coordination. Moreover, through the years numerous studies have been published describing biochemical analysis of HepII; none have reported a requirement for divalent cations. Structural investigations of ChonAC presented a catalytic scheme devoid of divalent metal cations utilization. To clarify the issue of charge neutralization in both systems more enzyme-substrate complex structures,

crystallized using inactive mutants of the enzymes and both different substrates (GlcA and IdoA bound in "+1" catalytic subsite) are required.

In the biological context of depolymerization of GAG polysaccharides to their disaccharide building blocks by microorganisms for the purposes of digestion and increased virulence through invasion of host tissues, both enzymes should be considered in the milieu of GAG degrading enzymes. Lyases are secreted out of the cytosol by their bacterial host. HepII resides in the periplasmic space of *P. heparinus* whereas BactnABCII is secreted to the gut lumen by *B. thetaiotaomicron*. Complete depolymerization of GAG is probably the result of combined action of the GAG lyases. It is easy to envision a simplistic scheme where heparin / HS degradation being the result of several lyases, among which HepII with its general, endolytic activity as the "first line of attack", chopping the long polysaccharide into shorter oligosaccharides serving as substrates for the exolytic degradation of more specific lyases i.e. HepI and III assisted by HepII itself. In the case of degradation of CS and DS species the picture was originally simplistic as well. *P. vulgaris* was shown to contain ChonABCI and II, on the other hand, *A. aurescens* and *P. heparinum* were shown to contain ChonAC and ChonB. Thus one could have argued that the combined action of ChonABCI and II is parallel to the combined actions of ConAC and ChonB, and sufficient for the complete depolymerization of these GAG species. The recently published genomic sequence of *B. thetaiotaomicron*, containing ChonABCI and II and two more sequences with similarities to ChonAC but with no established activities. This observation, together with the lack of complete genomic sequences for *P. vulgaris* and *A. aurescens* restrict the discussion about these systems until further data is accumulated.

7.3 Future Direction - Perspectives

Structure-function studies with GAG lyases, in particular HepII and ChonABC, have matured to the point where the actual roles of specific amino acid during catalysis are assigned and discussed. Enzyme-substrate complex structures reflecting the degradation of all types of substrates with HepII will be the next and maybe final step in the investigation of this lyase. A ternary complex of enzyme-substrate-metal ion is

required for ChonABC to resolve the issue of charge neutralization in this system and to capture the transient closed conformation that is argued through this dissertation.

The future of GAG research is not restricted to the elucidation of their catalytic mechanisms. The exciting new developments in the clinical applications of GAG have led a collaboration between the laboratory of Linhardt R. J. and our laboratory in order to improve their medical administration and clinical application. In that light, mutants of the known GAG lyases with altered specificities and modified catalysis might prove to be useful. Designing mutants can be achieved rationally based on the structures of the native enzymes.

7.4 Original Contributions to Knowledge

This dissertation presented the native structure of heparinase II, the first structure to be published of a heparin / HS degrading enzyme, which displayed:

1. The overall architecture of HepII resembles that of the known PL8 family enzymes
2. HepII is a dimer, as apposed to all the known lyase structures
3. A structural Zn^{2+} ion is present in the structure of HepII
4. Despite the structural differences the active site of HepII is located is a similar position as found in the known lyases with α-toroidal fold

After the native structure I presented the complex structures of HepII with its substrate, and products, which demonstrated:

1. HepII uses one active site to degrade both substrates
2. Four binding subsites and the binding characteristics were revealed
3. Active site residues were assigned roles in catalysis
4. HepII utilizes a lyase known strategy – higher energy conformation of substrate sugar ring to ease the degradation
5. Albeit acute differences HepII and HepI display similar active site residues arrangement, displaying convergent evolution

In its second part, this dissertation addressed ChonABC and revealed:

1. Although there is less than 22% identity between the *P. vulgaris* and *B. thetaiotaomicron* both enzyme display similar architecture presenting an excellent example for the conservation of structure, which conveys fuction.
2. The overall structure of ChonABC is typical for PL8 family enzyme
3. CS degradation is achieved much like the manner of ChonAC using one rim of the active site
4. New active site residue responsible for DS degradation reside on the opposite side of the active site, among them some have been assigned catalytic roles
5. ChonABC requires a bivalent metal, identified as Ca^{2+}, for DS degradation
6. The newly identified residue demonstrate that the enzyme undergoes conformational change reducing the size of the catalytic site
7. The work with this enzyme demonstrates an alternative approach to structure elucidation of enzyme-substrate complex to identify active site residues and even conformational changes through site directed mutagenesis

This dissertation has demonstrated that the lyase degrading glycosidic bonds with stereo-epimers of the uronic acid utilize one active site, by placing amino acids on both sides of the chiral center to perform both degradation processes

LIST OF REFERENCES

Aguiar JA, Lima CR, Berto AG & Michelacci YM 2003 An improved methodology to produce Flavobacterium heparinum chondroitinases, important instruments for diagnosis of diseases. *Biotechnol.Appl.Biochem* **37** 115-127.

Aigner T & McKenna L 2002 Molecular pathology and pathobiology of osteoarthritic cartilage. *Cell Mol.Life Sci.* **59** 5-18.

Allard ST, Giraud MF & Naismith JH 2001 Epimerases: structure, function and mechanism. *Cell Mol.Life Sci.* **58** 1650-1665.

Alvarez-Dominguez C, Vazquez-Boland JA, Carrasco-Marin E, Lopez-Mato P & Leyva-Cobian F 1997 Host cell heparan sulfate proteoglycans mediate attachment and entry of Listeria monocytogenes, and the listerial surface protein ActA is involved in heparan sulfate receptor recognition. *Infect.Immun.* **65** 78-88.

Ameer GA, Barabino G, Sasisekharan R, Harmon W, Cooney CL & Langer R 1999 Ex vivo evaluation of a Taylor-Couette flow, immobilized heparinase I device for clinical application. *Proc.Natl.Acad.Sci.U.S.A* **96** 2350-2355.

Ameye L & Young MF 2002 Mice deficient in small leucine-rich proteoglycans: novel in vivo models for osteoporosis, osteoarthritis, Ehlers-Danlos syndrome, muscular dystrophy, and corneal diseases. *Glycobiology* **12** 107R-116R.

Arya M, Patel HR & Williamson M 2003 Chemokines: key players in cancer. *Curr.Med.Res.Opin.* **19** 557-564.

Asher RA, Morgenstern DA, Moon LD & Fawcett JW 2001 Chondroitin sulphate proteoglycans: inhibitory components of the glial scar. *Prog.Brain Res.* **132** 611-619.

Backhed F, Ley RE, Sonnenburg JL, Peterson DA & Gordon JI 2005 Host-bacterial mutualism in the human intestine. *Science* **307** 1915-1920.

Bali JP, Cousse H & Neuzil E 2001 Biochemical basis of the pharmacologic action of chondroitin sulfates on the osteoarticular system. *Semin.Arthritis Rheum.* **31** 58-68.

Ban N, Nissen P, Hansen J, Moore PB & Steitz TA 2000 The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science* **289** 905-920.

Bandtlow CE & Zimmermann DR 2000 Proteoglycans in the developing brain: new conceptual insights for old proteins. *Physiol Rev.* **80** 1267-1290.

Bao X, Nishimura S, Mikami T, Yamada S, Itoh N & Sugahara K 2004 Chondroitin sulfate/dermatan sulfate hybrid chains from embryonic pig brain, which contain a higher proportion of L-iduronic acid than those from adult pig brain, exhibit neuritogenic and growth factor binding activities. *J Biol Chem* **279** 9765-9776.

Barth H, Schafer C, Adah MI, Zhang F, Linhardt RJ, Toyoda H, Kinoshita-Toyoda A, Toida T, Van Kuppevelt TH, Depla E, Von WF, Blum HE & Baumert TF 2003 Cellular binding of hepatitis C virus envelope glycoprotein E2 requires cell surface heparan sulfate. *J Biol Chem* **278** 41003-41012.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN & Bourne PE 2000 The Protein Data Bank. *Nucleic Acids Res* **28** 235-242.

Bernfield M, Gotte M, Park PW, Reizes O, Fitzgerald ML, Lincecum J & Zako M 1999 Functions of cell surface heparan sulfate proteoglycans. *Annu.Rev.Biochem.* **68** 729-777.

Bernfield M, Kokenyesi R, Kato M, Hinkes MT, Spring J, Gallo RL & Lose EJ 1992 Biology of the syndecans: a family of transmembrane heparan sulfate proteoglycans. *Annu.Rev.Cell Biol.* **8** 365-393.

Bernfield M & Sanderson RD 1990 Syndecan, a developmentally regulated cell surface proteoglycan that binds extracellular matrix and growth factors. *Philos.Trans.R.Soc.Lond B Biol.Sci.* **327** 171-186.

Bettelheim FA & Plessy B 1975 The hydration of proteoglycans of bovine cornea. *Biochim Biophys Acta* **381** 203-214.

Blackhall FH, Merry CL, Davies EJ & Jayson GC 2001 Heparan sulfate proteoglycans and cancer. *Br.J.Cancer* **85** 1094-1098.

Blain F, Tkalec AL, Shao Z, Poulin C, Pedneault M, Gu K, Eggimann B, Zimmermann J & Su H 2002 Expression system for high levels of GAG lyase gene expression and study of the hepA upstream region in Flavobacterium heparinum. *J Bacteriol.* **184** 3242-3252.

Blundell TL & Johnson LN 1976 *Protein Crystalography*. London: Academic Press Inc.

Bouchet V, Hood DW, Li J, Brisson JR, Randle GA, Martin A, Li Z, Goldstein R, Schweda EK, Pelton SI, Richards JC & Moxon ER 2003a Host-derived sialic acid is incorporated into Haemophilus influenzae lipopolysaccharide and is a major virulence factor in experimental otitis media. *Proc.Natl.Acad.Sci.U.S.A* **100** 8898-8903.

Bourdon MA, Krusius T, Campbell S, Schwartz NB & Ruoslahti E 1987 Identification and synthesis of a recognition signal for the attachment of glycosaminoglycans to proteins. *Proc.Natl.Acad.Sci.U.S.A* **84** 3194-3198.

Bourin MC & Lindahl U 1993 Glycosaminoglycans and the regulation of blood coagulation. *Biochem J* **289 (Pt 2)** 313-330.

Bradbury EJ, Moon LD, Popat RJ, King VR, Bennett GS, Patel PN, Fawcett JW & McMahon SB 2002 Chondroitinase ABC promotes functional recovery after spinal cord injury. *Nature* **416** 636-640.

Bradford MM 1976 A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72:248-54** 248-254.

Bragg WL 1913 *Proc.Camb.Phil.Soc.* **17**.

Bricogne G, Vonrhein C, Flensburg C, Schiltz M & Paciorek W 2003 Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr.D.Biol.Crystallogr.* **59** 2023-2030.

Brittis PA & Silver J 1995 Multiple factors govern intraretinal axon guidance: a time-lapse study. *Mol.Cell Neurosci.* **6** 413-432.

Brown EL, Wooten RM, Johnson BJ, Iozzo RV, Smith A, Dolan MC, Guo BP, Weis JJ & Hook M 2001 Resistance to Lyme disease in decorin-deficient mice. *J.Clin.Invest* **107** 845-852.

Brown GM, Huckerby TN, Abram BL & Nieduszynski IA 1996 Characterization of a non-reducing terminal fragment from bovine articular cartilage keratan sulphates containing alpha(2-3)-linked sialic acid and alpha(1-3)-linked fucose. A sulphated variant of the VIM-2 epitope. *Biochem J* **319 (Pt 1)** 137-141.

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T & Warren GL 1998 Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr.D.Biol.Crystallogr.* **54** 905-921.

Brunger AT, Kuriyan J & Karplus M 1987 Crystallographic R Factor Refinement by Molecular Dynamics. *Science* **235** 458-460.

Bulow HE & Hobert O 2004 Differential sulfations and epimerization define heparan sulfate specificity in nervous system development. *Neuron* **41** 723-736.

Bulow HE & Hobert O 2006 The molecular diversity of glycosaminoglycans shapes animal development. *Annu.Rev.Cell Dev.Biol.* **22** 375-407.

Burg MA & Cole GJ 1994 Claustrin, an antiadhesive neural keratan sulfate proteoglycan, is structurally related to MAP1B. *J.Neurobiol.* **25** 1-22.

Busse M, Feta A, Presto J, Wilen M, Gronning M, Kjellen L & Kusche-Gullberg M 2007 Contribution of EXT1, EXT2, and EXTL3 to heparan sulfate chain elongation. *J Biol Chem* **282** 32802-32810.

Butcher BA, Sklar LA, Seamer LC & Glew RH 1992 Heparin enhances the interaction of infective Leishmania donovani promastigotes with mouse peritoneal macrophages. A fluorescence flow cytometric analysis. *J Immunol.* **148** 2879-2886.

Caggiano AO, Zimber MP, Ganguly A, Blight AR & Gruskin EA 2005 Chondroitinase ABCI improves locomotion and bladder function following contusion injury of the rat spinal cord. *J.Neurotrauma* **22** 226-239.

Capila I & Linhardt RJ 2002 Heparin-protein interactions. *Angew.Chem Int.Ed Engl.* **41** 391-412.

Capila I, Wu Y, Rethwisch DW, Matte A, Cygler M & Linhardt RJ 2002 Role of arginine 292 in the catalytic activity of chondroitin AC lyase from Flavobacterium heparinum. *Biochim Biophys Acta* **1597** 260-270.

Carlsson P, Presto J, Spillmann D, Lindahl U & Kjellen L 2008 Heparin/heparan sulfate biosynthesis: processive formation of N-sulfated domains. *J Biol Chem* **283** 20008-20014.

Carter CW, Jr. & Sweet RM 1997 *Macromolecular Crystallography*. San Diego: Academic Press.

Carvalho AL, Goyal A, Prates JA, Bolam DN, Gilbert HJ, Pires VM, Ferreira LM, Planas A, Romao MJ & Fontes CM 2004 The family 11 carbohydrate-binding module of Clostridium thermocellum Lic26A-Cel5E accommodates beta-1,4- and beta-1,3-1,4-mixed linked glucans at a single binding site. *J Biol Chem* **279** 34785-34793.

Charnock SJ, Bolam DN, Turkenburg JP, Gilbert HJ, Ferreira LM, Davies GJ & Fontes CM 2000 The X6 "thermostabilizing" domains of xylanases are carbohydrate-binding modules: structure and biochemistry of the Clostridium thermocellum X6b domain. *Biochemistry* **39** 5013-5021.

Chau CH, Shum DK, Li H, Pei J, Lui YY, Wirthlin L, Chan YS & Xu XM 2004 Chondroitinase ABC enhances axonal regrowth through Schwann cell-seeded guidance channels after spinal cord injury. *FASEB J* **18** 194-196.

Chen JC & Stephens RS 1997 Chlamydia trachomatis glycosaminoglycan-dependent and independent attachment to eukaryotic cells. *Microb.Pathog.* **22** 23-30.

Chen Y, Maguire T, Hileman RE, Fromm JR, Esko JD, Linhardt RJ & Marks RM 1997 Dengue virus infectivity depends on envelope protein binding to target cell heparan sulfate. *Nat Med* **3** 866-871.

Cleland WW, Frey PA & Gerlt JA 1998 The low barrier hydrogen bond in enzymatic catalysis. *J Biol Chem* **273** 25529-25532.

Comstock LE & Coyne MJ 2003 Bacteroides thetaiotaomicron: a dynamic, niche-adapted human symbiont. *Bioessays* **25** 926-929.

Comstock LE & Kasper DL 2006 Bacterial glycans: key mediators of diverse host immune responses. *Cell* **126** 847-850.

Condic ML, Snow DM & Letourneau PC 1999 Embryonic neurons adapt to the inhibitory proteoglycan aggrecan by increasing integrin expression. *J.Neurosci.* **19** 10036-10043.

Coutinho PM & Henrissat B 1999 Carbohydrate-active enzymes: an integrated database approach. In *Recent Advances in Carbohydrate Bioengineering*, pp 3-12. Eds HJ Gilbert, G Davies, B Henrissat & B Svensson. Cambridge, The Royal Society of Chemistry.

D'Amore PA 1990 Modes of FGF release in vivo and in vitro. *Cancer Metastasis Rev.* **9** 227-238.

Dai Y, Yang Y, MacLeod V, Yue X, Rapraeger AC, Shriver Z, Venkataraman G, Sasisekharan R & Sanderson RD 2005 HSulf-1 and HSulf-2 are potent inhibitors of myeloma tumor growth in vivo. *J Biol Chem* **280** 40066-40073.

Davies G & Henrissat B 1995 Structures and mechanisms of glycosyl hydrolases. *Structure* **3** 853-859.

Davies GJ, Wilson KS & Henrissat B 1997 Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem J* **321 (Pt 2)** 557-559.

Davies Y, Lewis D, Fullwood NJ, Nieduszynski IA, Marcyniuk B, Albon J & Tullo A 1999 Proteoglycans on normal and migrating human corneal endothelium. *Exp.Eye Res.* **68** 303-311.

DeAngelis PL 2002 Microbial glycosaminoglycan glycosyltransferases. *Glycobiology* **12** 9R-16R.

Degroote S, Lo-Guidice JM, Strecker G, Ducourouble MP, Roussel P & Lamblin G 1997 Characterization of an N-acetylglucosamine-6-O-sulfotransferase from human respiratory mucosa active on mucin carbohydrate chains. *J Biol Chem* **272** 29493-29501.

Delpech B, Girard N, Bertrand P, Courel MN, Chauzy C & Delpech A 1997 Hyaluronan: fundamental principles and applications in cancer. *J.Intern.Med.* **242** 41-48.

Desai UR, Wang HM & Linhardt RJ 1993a Specificity studies on the heparin lyases from Flavobacterium heparinum. *Biochemistry* **32** 8140-8145.

Desai UR, Wang HM & Linhardt RJ 1993b Substrate specificity of the heparin lyases from Flavobacterium heparinum. *Arch.Biochem.Biophys.* **306** 461-468.

Deutscher MP 1990 *Protein Purification*. San-Diego: Academic Press.

Dietrich CP, Tersariol IL, Toma L, Moraes CT, Porcionatto MA, Oliveira FW & Nader HB 1998 Structure of heparan sulfate: identification of variable and constant

oligosaccharide domains in eight heparan sulfates of different origins. *Cell Mol.Biol.(Noisy.-le-grand)* **44** 417-429.

Dityatev A & Schachner M 2003 Extracellular matrix molecules and synaptic plasticity. *Nat.Rev.Neurosci.* **4** 456-468.

Dong J, Peters-Libeu CA, Weisgraber KH, Segelke BW, Rupp B, Capila I, Hernaiz MJ, LeBrun LA & Linhardt RJ 2001 Interaction of the N-terminal domain of apolipoprotein E4 with heparin. *Biochemistry* **40** 2826-2834.

Doublie S 1997 Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol* **276** 523-530.

Drenth J 1994 *Principles of Protein X-ray Crystalography*. New York: Springer-Verlag.

Duncan G, McCormick C & Tufaro F 2001 The link between heparan sulfate and hereditary bone disease: finding a function for the EXT family of putative tumor suppressor proteins. *J.Clin.Invest* **108** 511-516.

Edge AS & Spiro RG 1990 Characterization of novel sequences containing 3-O-sulfated glucosamine in glomerular basement membrane heparan sulfate and localization of sulfated disaccharides to a peripheral domain. *J Biol Chem* **265** 15874-15881.

Elefteriou F, Exposito JY, Garrone R & Lethias C 2001 Binding of tenascin-X to decorin. *FEBS Lett* **495** 44-47.

Emsley P & Cowtan K 2004 Coot: model-building tools for molecular graphics. *Acta Crystallogr.D.Biol.Crystallogr.* **60** 2126-2132.

Ernst S, Langer R, Cooney CL & Sasisekharan R 1995a Enzymatic degradation of glycosaminoglycans. *Crit.Rev.Biochem.Mol.Biol.* **30** 387-444.

Ernst S, Rhomberg AJ, Biemann K & Sasisekharan R 1998 Direct evidence for a predominantly exolytic processive mechanism for depolymerization of heparin-like glycosaminoglycans by heparinase I. *Proc.Natl.Acad.Sci.U.S.A* **95** 4182-4187.

Esko JD & Lindahl U 2001 Molecular diversity of heparan sulfate. *J.Clin.Invest* **108** 169-173.

Esko JD & Selleck SB 2002 Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu.Rev.Biochem* **71** 435-471.

Esko JD & Zhang L 1996 Influence of core protein sequence on glycosaminoglycan assembly. *Curr.Opin.Struct.Biol* **6** 663-670.

Faissner A, Clement A, Lochter A, Streit A, Mandl C & Schachner M 1994 Isolation of a neural chondroitin sulfate proteoglycan with neurite outgrowth promoting properties. *J Cell Biol* **126** 783-799.

Fethiere J, Eggimann B & Cygler M 1999 Crystal structure of chondroitin AC lyase, a representative of a family of glycosaminoglycan degrading enzymes. *J Mol Biol* **288** 635-647.

Feyerabend TB, Li JP, Lindahl U & Rodewald HR 2006 Heparan sulfate C5-epimerase is essential for heparin biosynthesis in mast cells. *Nat.Chem.Biol.* **2** 195-196.

Feyzi E, Saldeen T, Larsson E, Lindahl U & Salmivirta M 1998 Age-dependent modulation of heparan sulfate structure and function. *J Biol Chem* **273** 13395-13398.

Filmus J 2001 Glypicans in growth control and cancer. *Glycobiology* **11** 19R-23R.

Folkman J & Klagsbrun M 1987 Angiogenic factors. *Science* **235** 442-447.

Fransson LA, Belting M, Jonsson M, Mani K, Moses J & Oldberg A 2000 Biosynthesis of decorin and glypican. *Matrix Biology* **19** 367-376.

Freissler E, Meyer auf der HA, David G, Meyer TF & Dehio C 2000 Syndecan-1 and syndecan-4 can mediate the invasion of OpaHSPG-expressing Neisseria gonorrhoeae into epithelial cells. *Cell Microbiol.* **2** 69-82.

Frevert U, Sinnis P, Cerami C, Shreffler W, Takacs B & Nussenzweig V 1993 Malaria circumsporozoite protein binds to heparan sulfate proteoglycans associated with the surface membrane of hepatocytes. *J Exp Med* **177** 1287-1298.

Frick IM, Schmidtchen A & Sjobring U 2003 Interactions between M proteins of Streptococcus pyogenes and glycosaminoglycans promote bacterial adhesion to host cells. *Eur J Biochem* **270** 2303-2311.

Fukuta M, Inazawa J, Torii T, Tsuzuki K, Shimada E & Habuchi O 1997 Molecular cloning and characterization of human keratan sulfate Gal-6-sulfotransferase. *J Biol Chem* **272** 32321-32328.

Funderburgh JL 2000 Keratan sulfate: structure, biosynthesis, and function. *Glycobiology* **10** 951-958.

Funderburgh JL, Mitschler RR, Funderburgh ML, Roth MR, Chapes SK & Conrad GW 1997 Macrophage receptors for lumican. A corneal keratan sulfate proteoglycan. *Invest Ophthalmol.Vis.Sci.* **38** 1159-1167.

Gacesa P 1987 Alginate-modifying enzymes. A proposed unified mechanism of action for the lyases and epimerases. *FEBS Lett* **212** 199-202.

Gailit J & Clark RA 1994 Wound repair in the context of extracellular matrix. *Curr.Opin.Cell Biol* **6** 717-725.

Gallagher JT & Turnbull JE 1992 Heparan sulphate in the binding and activation of basic fibroblast growth factor. *Glycobiology* **2** 523-528.

Gallo R, Kim C, Kokenyesi R, Adzick NS & Bernfield M 1996 Syndecans-1 and -4 are induced during wound repair of neonatal but not fetal skin. *J.Invest Dermatol.* **107** 676-683.

Galtrey CM & Fawcett JW 2007 The role of chondroitin sulfate proteoglycans in regeneration and plasticity in the central nervous system. *Brain Res.Rev.* **54** 1-18.

Godavarti R & Sasisekharan R 1996 A comparative analysis of the primary sequences and characteristics of heparinases I, II, and III from *Flavobacterium heparinum*. *Biochem Biophys Res Commun* **229** 770-777.

Graham RA, Li TC, Cooke ID & Aplin JD 1994 Keratan sulphate as a secretory product of human endometrium: cyclic expression in normal women. *Hum.Reprod.* **9** 926-930.

Greenwood JA & Murphy-Ullrich JE 1998 Signaling of de-adhesion in cellular regulation and motility. *Microsc.Res.Tech.* **43** 420-432.

Gu K, Linhardt RJ, Laliberte M & Zimmermann JJ 1995 Purification, characterization and specificity of chondroitin lyases and glycuronidase from *Flavobacterium heparinum*. *Biochem J* **312** 569-577.

Guerassimov A, Zhang Y, Cartman A, Rosenberg LC, Esdaile J, Fitzcharles MA & Poole AR 1999 Immune responses to cartilage link protein and the G1 domain of proteoglycan aggrecan in patients with osteoarthritis. *Arthritis Rheum.* **42** 527-533.

Habuchi O 2000 Diversity and functions of glycosaminoglycan sulfotransferases. *Biochim Biophys Acta* **1474** 115-127.

Habuchi O, Hirahara Y, Uchimura K & Fukuta M 1996 Enzymatic sulfation of galactose residue of keratan sulfate by chondroitin 6-sulfotransferase. *Glycobiology* **6** 51-57.

Halldorsdottir AM, Zhang L & Tollefsen DM 2006 N-Acetylgalactosamine 4,6-O-sulfate residues mediate binding and activation of heparin cofactor II by porcine mucosal dermatan sulfate. *Glycobiology* **16** 693-701.

Hamai A, Hashimoto N, Mochizuki H, Kato F, Makiguchi Y, Horie K & Suzuki S 1997 Two distinct chondroitin sulfate ABC lyases. An endoeliminase yielding tetrasaccharides and an exoeliminase preferentially acting on oligosaccharides. *J Biol Chem* **272** 9123-9130.

Hashimoto W, Nankai H, Mikami B & Murata K 2003 Crystal structure of Bacillus sp. GL1 xanthan lyase, which acts on the side chains of xanthan. *J Biol Chem* **278** 7663-7673.

Hassell JR, Newsome DA, Krachmer JH & Rodrigues MM 1980 Macular corneal dystrophy: failure to synthesize a mature keratan sulfate proteoglycan. *Proc.Natl.Acad.Sci.U.S.A* **77** 3705-3709.

Heldermon C, DeAngelis PL & Weigel PH 2001 Topological organization of the hyaluronan synthase from Streptococcus pyogenes. *J Biol Chem* **276** 2037-2046.

Hemmerich S, Leffler H & Rosen SD 1995 Structure of the O-glycans in GlyCAM-1, an endothelial-derived ligand for L-selectin. *J Biol Chem* **270** 12035-12047.

Hendrickson WA, Horton JR & LeMaster DM 1990 Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J* **9** 1665-1672.

Herron SR, Benen JA, Scavetta RD, Visser J & Jurnak F 2000 Structure and function of pectic enzymes: virulence factors of plant pathogens. *Proc.Natl.Acad.Sci.U.S.A* **97** 8762-8769.

Herron SR, Scavetta RD, Garrett M, Legner M & Jurnak F 2003 Characterization and implications of Ca2+ binding to pectate lyase C. *J Biol Chem* **278** 12271-12277.

Hileman RE, Fromm JR, Weiler JM & Linhardt RJ 1998 Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins. *Bioessays* **20** 156-167.

Hirsh J, Anand SS, Halperin JL & Fuster V 2001 Guide to anticoagulant therapy: Heparin a statement for healthcare professionals from the American Heart Association. *Circulation* **103** 2994-3018.

Hirsh J & Raschke R 2004 Heparin and low-molecular-weight heparin: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest* **126** 188S-203S.

Hiyama K & Okada S 1976 Action of chondroitinases. I. The mode of action of two chondroitinase- AC preparations of different origin. *J.Biochem.(Tokyo)* **80** 1201-1207.

Holmborn K, Ledin J, Smeds E, Eriksson I, Kusche-Gullberg M & Kjellen L 2004 Heparan sulfate synthesized by mouse embryonic stem cells deficient in NDST1 and NDST2 is 6-O-sulfated but contains no N-sulfate groups. *J Biol Chem* **279** 42355-42358.

Huang L., van Halbeek, H., Eggimann, B., and Zimmermann, J. Structural Characterization of the Novel O-Linked Carbohydrate Structure of *Flavobacterium Heparinum* Heparinase I. 5 No. 7, 709-713. 1995a. Oxford Journals. Abstracts form the 3rd International Glycobiology Symposium: Current Analytical Methods: 2. Mass Spectrometry and NMR. Conference Proceeding

Huang W, Boju L, Tkalec L, Su H, Yang HO, Gunay NS, Linhardt RJ, Kim YS, Matte A & Cygler M 2001 Active site of chondroitin AC lyase revealed by the structure of enzyme-oligosaccharide complexes and mutagenesis. *Biochemistry* **40** 2359-2372.

Huang W, Lunin VV, Li Y, Suzuki S, Sugiura N, Miyazono H & Cygler M 2003 Crystal structure of Proteus vulgaris chondroitin sulfate ABC lyase I at 1.9A resolution. *J Mol Biol* **328** 623-634.

Huang W, Matte A, Li Y, Kim YS, Linhardt RJ, Su H & Cygler M 1999 Crystal Structure of Chondroitinase B from *Flavobacterium heparinum* and its Complex with a Disaccharide Product at 1.7Å Resolution. *J Mol Biol* **294** 1257-1269.

Humphries DE, Wong GW, Friend DS, Gurish MF, Qiu WT, Huang C, Sharpe AH & Stevens RL 1999 Heparin is essential for the storage of specific granule proteases in mast cells. *Nature* **400** 769-772.

Ibrahimi OA, Zhang F, Hrstka SC, Mohammadi M & Linhardt RJ 2004 Kinetic model for FGF, FGFR, and proteoglycan signal transduction complex assembly. *Biochemistry* **43** 4724-4730.

Iozzo RV 1998 Matrix proteoglycans: from molecular design to cellular function. *Annu.Rev.Biochem.* **67** 609-652.

Iozzo RV & Cohen I 1993 Altered proteoglycan gene expression and the tumor stroma. *Experientia* **49** 447-455.

Iozzo RV & Murdoch AD 1996 Proteoglycans of the extracellular environment: clues from the gene and protein side offer novel perspectives in molecular diversity and function. *FASEB J* **10** 598-614.

Jackson RL, Busch SJ & Cardin AD 1991 Glycosaminoglycans: molecular properties, protein interactions, and role in physiological processes. *Physiol.Rev.* **71** 481-539.

Jancarik J & Kim SH 1991 Sparse matrix sampling: a screening method for crystallization of proteins. *J Appl Crystallogr* **24** 409-414.

Jandik KA, Gu K & Linhardt RJ 1994 Action pattern of polysaccharide lyases on glycosaminoglycans. *Glycobiology* **4** 289-296.

Jedrzejas MJ & Chantalat L 2000 Structural studies of Streptococcus agalactiae hyaluronate lyase. *Acta Crystallogr.* **D56** 460-463.

Jedrzejas MJ, Chantalat L & Mewbourne RB 1998 Crystallization and preliminary X-ray analysis of *Streptococcus pneumoniae* hyaluronate lyase. *J.Struct.Biol.* **121** 75.

Jenkins J, Mayans O & Pickersgill R 1998 Structure and evolution of parallel beta-helix proteins. *J.Struct.Biol.* **122** 236-246.

Johnson AR 1993 Contact inhibition in the failure of mammalian CNS axonal regeneration. *Bioessays* **15** 807-813.

Johnson KA, Hulse DA, Hart RC, Kochevar D & Chu Q 2001 Effects of an orally administered mixture of chondroitin sulfate, glucosamine hydrochloride and manganese ascorbate on synovial fluid chondroitin sulfate 3B3 and 7D4 epitope in a canine cruciate ligament transection model of osteoarthritis. *Osteoarthritis.Cartilage.* **9** 14-21.

Johnson P, Maiti A, Brown KL & Li R 2000 A role for the cell adhesion molecule CD44 and sulfation in leukocyte-endothelial cell adhesion during an inflammatory response? *Biochem.Pharmacol.* **59** 455-465.

Jones TA, Kleywegt GJ & Brunger AT 1996 Storing diffraction data. *Nature* **383** 18-19.

Kato M, Wang H, Bernfield M, Gallagher JT & Turnbull JE 1994 Cell surface syndecan-1 on distinct cell types differs in fine structure and ligand binding of its heparan sulfate chains. *J Biol Chem* **269** 18881-18890.

Kawashima H, Atarashi K, Hirose M, Hirose J, Yamada S, Sugahara K & Miyasaka M 2002 Oversulfated chondroitin/dermatan sulfates containing GlcAbeta1/IdoAalpha1-3GalNAc(4,6-O-disulfate) interact wi. *J Biol Chem* **277** 12921-12930.

Keiser N, Venkataraman G, Shriver Z & Sasisekharan R 2001 Direct isolation and sequencing of specific protein-binding glycosaminoglycans. *Nat Med* **7** 123-128.

KENDREW JC, BODO G, DINTZIS HM, PARRISH RG, WYCKOFF H & Phillips DC 1958 A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181** 662-666.

Kiani C, Chen L, Wu YJ, Yee AJ & Yang BB 2002 Structure and function of aggrecan. *Cell Res.* **12** 19-32.

Kim CW, Goldberger OA, Gallo RL & Bernfield M 1994 Members of the syndecan family of heparan sulfate proteoglycans are expressed in distinct cell-, tissue-, and development-specific patterns. *Mol.Biol Cell* **5** 797-805.

Kim YS, Jo YY, Chang IM, Toida T, Park Y & Linhardt RJ 1996c A new glycosaminoglycan from the giant African snail Achatina fulica. *J Biol Chem* **271** 11750-11755.

Kim YS, Jo YY, Chang IM, Toida T, Park Y & Linhardt RJ 1996b A new glycosaminoglycan from the giant African snail Achatina fulica. *J Biol Chem* **271** 11750-11755.

Kim YS, Jo YY, Chang IM, Toida T, Park Y & Linhardt RJ 1996a A new glycosaminoglycan from the giant African snail Achatina fulica. *J Biol Chem* **271** 11750-11755.

Kitagawa H, Izumikawa T, Uyama T & Sugahara K 2003 Molecular cloning of a chondroitin polymerizing factor that cooperates with chondroitin synthase for chondroitin polymerization. *J Biol Chem*.

Kitagawa H, Uyama T & Sugahara K 2001 Molecular cloning and expression of a human chondroitin synthase. *J Biol Chem* **276** 38721-38726.

Kjellen L & Lindahl U 1991 Proteoglycans: structures and interactions. *Annu.Rev.Biochem.* **60** 443-475.

Knudson CB & Knudson W 1993 Hyaluronan-binding proteins in development, tissue homeostasis, and disease. *FASEB J* **7** 1233-1241.

Knudson CB & Knudson W 2001 Cartilage proteoglycans. *Semin.Cell Dev.Biol.* **12** 69-78.

Kobayashi M, Sugumaran G, Liu J, Shworak NW, Silbert JE & Rosenberg RD 1999 Molecular cloning and characterization of a human uronyl 2-sulfotransferase that sulfates iduronyl and glucuronyl residues in dermatan/chondroitin sulfate. *J Biol Chem* **274** 10474-10480.

Kolset SO & Gallagher JT 1990 Proteoglycans in haemopoietic cells. *Biochim Biophys Acta* **1032** 191-211.

Kresse H & Schonherr E 2001 Proteoglycans of the extracellular matrix and growth control. *J.Cell Physiol* **189** 266-274.

Krusius T, Finne J, Margolis RK & Margolis RU 1986 Identification of an O-glycosidic mannose-linked sialylated tetrasaccharide and keratan sulfate oligosaccharides in the chondroitin sulfate proteoglycan of brain. *J Biol Chem* **261** 8237-8242.

Kuroki R, Weaver LH & Matthews BW 1993 A covalent enzyme-substrate intermediate with saccharide distortion in a mutant T4 lysozyme. *Science* **262** 2030-2033.

Kussie PH, Hulmes JD, Ludwig DL, Patel S, Navarro EC, Seddon AP, Giorgio NA & Bohlen P 1999 Cloning and functional expression of a human heparanase gene. *Biochem Biophys Res Commun* **261** 183-187.

Lairson LL, Henrissat B, Davies GJ & Withers SG 2008 Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu.Rev.Biochem.* **77** 521-555.

Lamanna WC, Baldwin RJ, Padva M, Kalus I, Ten DG, Van Kuppevelt TH, Gallagher JT, von FK, Dierks T & Merry CL 2006 Heparan sulfate 6-O-endosulfatases: discrete in vivo activities and functional co-operativity. *Biochem J* **400** 63-73.

Lamanna WC, Kalus I, Padva M, Baldwin RJ, Merry CL & Dierks T 2007 The heparanome--the enigma of encoding and decoding heparan sulfate sulfation. *J.Biotechnol.* **129** 290-307.

Lamzin VS, Dauter Z & Wilson KS 1994 Dehydrogenation through the looking-glass. *Nat Struct Biol* **1** 281-282.

Lamzin VS, Dauter Z & Wilson KS 1995 How nature deals with stereoisomers. *Curr.Opin.Struct.Biol* **5** 830-836.

Lander AD, Nie Q & Wan FY 2002 Do morphogen gradients arise by diffusion? *Dev.Cell* **2** 785-796.

Lander AD & Selleck SB 2000 The elusive functions of proteoglycans: in vivo veritas. *J Cell Biol* **148** 227-232.

Langer R, Linhardt RJ, Hoffberg S, Larsen AK, Cooney CL, Tapper D & Klein M 1982 An enzymatic system for removing heparin in extracorporeal therapy. *Science* **217** 261-263.

Laskowski RA, MacArthur MW, Moss DS & Thornton JM 1993 PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* **26** 283-291.

LeBrun LA & Linhardt RJ 2001 Degradation of heparan sulfate with heparin lyases. *Methods Mol.Biol.* **171** 353-361.

Ledin J, Ringvall M, Thuveson M, Eriksson I, Wilen M, Kusche-Gullberg M, Forsberg E & Kjellen L 2006 Enzymatically active N-deacetylase/N-sulfotransferase-2 is present in liver but does not contribute to heparan sulfate N-sulfation. *J Biol Chem* **281** 35727-35734.

Lee B & Richards FM 1971 The interpretation of protein structures: estimation of static accessibility. *J Mol Biol* **55** 379-400.

Lemons ML, Sandy JD, Anderson DK & Howland DR 2003b Intact aggrecan and chondroitin sulfate-depleted aggrecan core glycoprotein inhibit axon growth in the adult rat spinal cord. *Exp.Neurol.* **184** 981-990.

Li F, Shetty AK & Sugahara K 2007 Neuritogenic activity of chondroitin/dermatan sulfate hybrid chains of embryonic pig brain and their mimicry from shark liver. Involvement of the pleiotrophin and hepatocyte growth factor signaling pathways. *J Biol Chem* **282** 2956-2966.

Li S & Jedrzejas MJ 2001 Hyaluronan binding and degradation by *Streptococcus agalactiae* hyaluronate lyase. *J Biol Chem* **276** 41407-41416.

Li S, Kelly SJ, Lamani E, Ferraroni M & Jedrzejas MJ 2000 Structural basis of hyaluronan degradation by *Streptococcus pneumoniae* hyaluronate lyase. *EMBO J* **19** 1228-1240.

Liaw PC, Austin RC, Fredenburgh JC, Stafford AR & Weitz JI 1999 Comparison of heparin- and dermatan sulfate-mediated catalysis of thrombin inactivation by heparin cofactor II. *J Biol Chem* **274** 27597-27604.

Liaw PC, Becker DL, Stafford AR, Fredenburgh JC & Weitz JI 2001 Molecular basis for the susceptibility of fibrin-bound thrombin to inactivation by heparin cofactor ii in the presence of dermatan sulfate but not heparin. *J Biol Chem* **276** 20959-20965.

Lindahl B, Eriksson L & Lindahl U 1995 Structure of heparan sulphate from human brain, with special regard to Alzheimer's disease. *Biochem J* **306 (Pt 1)** 177-184.

Lindahl U, Kusche-Gullberg M & Kjellen L 1998 Regulated diversity of heparan sulfate. *J Biol Chem* **273** 24979-24982.

Lindahl U, Lidholt K, Spillmann D & Kjellen L 1994 More to "heparin" than anticoagulation. *Thromb.Res.* **75** 1-32.

Linhardt RJ 2003 2003 Claude S. Hudson Award address in carbohydrate chemistry. Heparin: structure and activity. *J Med Chem* **46** 2551-2564.

Linhardt RJ, Galliher PM & Cooney CL 1986 Polysaccharide lyases. *Appl.Biochem.Biotechnol.* **12** 135-176.

Linhardt RJ & Toida T 2004 Role of glycosaminoglycans in cellular communication. *Acc Chem Res* **37** 431-438.

Linhardt RJ, Wang HM, Loganathan D & Bae JH 1992 Search for the heparin antithrombin III-binding site precursor. *J Biol Chem* **267** 2380-2387.

Linn S, Chan T, Lipeski L & Salyers AA 1983a Isolation and characterization of two chondroitin lyases from *Bacteroides thetaiotaomicron*. *J Bacteriol.* **156** 859-866.

Litwack ED, Ivins JK, Kumbasar A, Paine-Saunders S, Stipp CS & Lander AD 1998 Expression of the heparan sulfate proteoglycan glypican-1 in the developing rodent. *Dev.Dyn.* **211** 72-87.

Liu D, Shriver Z, Venkataraman G, El SY & Sasisekharan R 2002 Tumor cell surface heparan sulfate as cryptic promoters or inhibitors of tumor growth and metastasis. *Proc.Natl.Acad.Sci.U.S.A* **99** 568-573.

Lohse DL & Linhardt RJ 1992 Purification and characterization of heparin lyases from Flavobacterium heparinum. *J Biol Chem* **267** 24347-24355.

Lunin VV, Li Y, Linhardt RJ, Miyazono H, Kyogashima M, Kaneko T, Bell AW & Cygler M 2004 High Resolution Crystal Structure of *Arthrobacter aurescens* Chondroitin AC Lyase: Enzyme-Substrate Complex Defines the Catalytic Mechanism. *J Mol Biol* **337** 367-386.

Lyon M, Deakin JA & Gallagher JT 1994 Liver heparan sulfate structure. A novel molecular design. *J Biol Chem* **269** 11208-11215.

Lyon M, Deakin JA, Rahmoune H, Fernig DG, Nakamura T & Gallagher JT 1998 Hepatocyte growth factor/scatter factor binds with high affinity to dermatan sulfate. *J Biol Chem* **273** 271-278.

Maccarana M, Casu B & Lindahl U 1993 Minimal sequence in heparin/heparan sulfate required for binding of basic fibroblast growth factor. *J Biol Chem* **268** 23898-23905.

Maccarana M, Sakura Y, Tawada A, Yoshida K & Lindahl U 1996 Domain structure of heparan sulfates from bovine organs. *J Biol Chem* **271** 17804-17810.

Mascellani G, Liverani L, Bianchini P, Parma B, Torri G, Bisio A, Guerrini M & Casu B 1993 Structure and contribution to the heparin cofactor II-mediated inhibition of thrombin of naturally oversulphated sequences of dermatan sulphate. *Biochem J* **296 (Pt 3)** 639-648.

Mattevi A, Vanoni MA, Todone F, Rizzi M, Teplyakov A, Coda A, Bolognesi M & Curti B 1996 Crystal structure of D-amino acid oxidase: a case of active site mirror-image convergent evolution with flavocytochrome b2. *Proc.Natl.Acad.Sci.U.S.A* **93** 7496-7501.

Matthews BW 1968 Solvent content of protein crystals. *J Mol Biol* **33** 491-497.

Matulis D, Kranz JK, Salemme FR & Todd MJ 2005 Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor. *Biochemistry* **44** 5258-5266.

Mayans O, Scott M, Connerton I, Gravesen T, Benen J, Visser J, Pickersgill R & Jenkins J 1997 Two crystal structures of pectin lyase A from *Aspergillus* reveal a pH driven conformational change and striking divergence in the substrate- binding clefts of pectin and pectate lyases. *Structure* **5** 677-689.

McAlindon TE, LaValley MP, Gulin JP & Felson DT 2000 Glucosamine and chondroitin for treatment of osteoarthritis: a systematic quality assessment and meta-analysis. *JAMA* **283** 1469-1475.

McCormick C, Duncan G, Goutsos KT & 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. *Proc.Natl.Acad.Sci.U.S.A* **97** 668-673.

McKeon RJ, Schreiber RC, Rudge JS & Silver J 1991 Reduction of neurite outgrowth in a model of glial scarring following CNS injury is correlated with the expression of inhibitory molecules on reactive astrocytes. *J.Neurosci.* **11** 3398-3411.

McPherson A 2004 Protein crystallization in the structural genomics era. *J.Struct.Funct.Genomics* **5** 3-12.

Mello LV, De Groot BL, Li S & Jedrzejas MJ 2002 Structure and flexibility of Streptococcus agalactiae hyaluronate lyase complex with its substrate. Insights into the mechanism of processive degradation of hyaluronan. *J Biol Chem* **277** 36678-36688.

Menozzi FD, Pethe K, Bifani P, Soncin F, Brennan MJ & Locht C 2002 Enhanced bacterial virulence through exploitation of host glycosaminoglycans. *Mol.Microbiol.* **43** 1379-1386.

Michel G, Pojasek K, Li Y, Sulea T, Linhardt RJ, Raman R, Prabhakar V, Sasisekharan R & Cygler M 2004 The structure of chondroitin B lyase complexed with glycosaminoglycan oligosaccharides unravels a calcium-dependent catalytic machinery. *J Biol Chem* **279** 32882-32896.

Michelacci YM & Dietrich CP 1974 Isolation and partial characterization of an induced chondroitinase B from Flavobacterium heparinum. *Biochem.Biophys.Res.Commun* **56** 973-980.

Michelacci YM & Dietrich CP 1975 A comparative study between a chondroitinase B and a chondroitinase AC from Flavobacterium heparinum: Isolation of a chondroitinase AC- susceptible dodecasaccharide from chondroitin sulphate B. *Biochem J* **151** 121-129.

Miller BE, Guzzetta NA, Tosone SR & Levy JH 2000 Rapid evaluation of coagulopathies after cardiopulmonary bypass in children using modified thromboelastography. *Anesth.Analg.* **90** 1324-1330.

Moffat CF, McLean MW, Long WF & Williamson FB 1991a Heparinase II from *Flavobacterium heparinum*. Action on chemically modified heparins. *Eur J Biochem* **197** 449-459.

Moffat CF, McLean MW, Long WF & Williamson FB 1991b Heparinase II from *Flavobacterium heparinum*. HPLC analysis of the saccharides generated from chemically modified heparins. *Eur J Biochem* **202** 531-541.

Moon LD, Asher RA, Rhodes KE & Fawcett JW 2001 Regeneration of CNS axons back to their target following treatment of adult rat brain with chondroitinase ABC. *Nat.Neurosci.* **4** 465-466.

Morgenstern DA, Asher RA & Fawcett JW 2002 Chondroitin sulphate proteoglycans in the CNS injury response. *Prog.Brain Res.* **137** 313-332.

Mummert ME 2005 Immunologic roles of hyaluronan. *Immunol.Res.* **31** 189-206.

Muraki M 2002 The importance of CH/pi interactions to the function of carbohydrate binding proteins. *Protein Pept.Lett.* **9** 195-209.

Murphy KJ, Merry CL, Lyon M, Thompson JE, Roberts IS & Gallagher JT 2004 A new model for the domain structure of heparan sulfate based on the novel specificity of K5 lyase. *J Biol Chem* **279** 27239-27245.

Murshudov GN, Vagin AA & Dodson EJ 1997 Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr.D.Biol.Crystallogr.* **53** 240-255.

Myette JR, Shriver Z, Kiziltepe T, McLean MW, Venkataraman G & Sasisekharan R 2002 Molecular cloning of the heparin/heparan sulfate delta 4,5 unsaturated glycuronidase from Flavobacterium heparinum, its recombinant expression in

Escherichia coli, and biochemical determination of its unique substrate specificity. *Biochemistry* **41** 7424-7434.

Nadanaka S & Sugahara K 1997 The unusual tetrasaccharide sequence GlcA beta 1-3GalNAc(4-sulfate)beta 1-4GlcA(2-sulfate)beta 1-3GalNAc(6-sulfate) found in the hexasaccharides prepared by testicular hyaluronidase digestion of shark cartilage chondroitin sulfate D. *Glycobiology* **7** 253-263.

Nader HB, Chavante SF, dos-Santos EA, Oliveira TW, de-Paiva JF, Jeronimo SM, Medeiros GF, de-Abreu LR, Leite EL, de-Sousa-Filho JF, Castro RA, Toma L, Tersariol IL, Porcionatto MA & Dietrich CP 1999 Heparan sulfates and heparins: similar compounds performing the same functions in vertebrates and invertebrates? *Braz.J.Med.Biol.Res.* **32** 529-538.

Nader HB, Porcionatto MA, Tersariol IL, Pinhal MA, Oliveira FW, Moraes CT & Dietrich CP 1990 Purification and substrate specificity of heparitinase I and heparitinase II from Flavobacterium heparinum. Analyses of the heparin and heparan sulfate degradation products by 13C NMR spectroscopy. *J Biol Chem* **265** 16807-16813.

Nagy T, Simpson P, Williamson MP, Hazlewood GP, Gilbert HJ & Orosz L 1998 All three surface tryptophans in Type IIa cellulose binding domains play a pivotal role in binding both soluble and insoluble ligands. *FEBS Lett* **429** 312-316.

Najmudin S, Andersen JT, Patkar SA, Borchert TV, Crout DH & Fulop V 2003 Purification, crystallization and preliminary X-ray crystallographic studies on acetolactate decarboxylase. *Acta Crystallogr.D.Biol.Crystallogr.* **59** 1073-1075.

Nelson RM, Venot A, Bevilacqua MP, Linhardt RJ & Stamenkovic I 1995 Carbohydrate-protein interactions in vascular biology. *Annu.Rev.Cell Dev.Biol.* **11** 601-631.

Nilsson B, Nakazawa K, Hassell JR, Newsome DA & Hascall VC 1983 Structure of oligosaccharides and the linkage region between keratan sulfate and the core protein on proteoglycans from monkey cornea. *J Biol Chem* **258** 6056-6063.

Oldberg A, Antonsson P, Hedbom E & Heinegard D 1990 Structure and function of extracellular matrix proteoglycans. *Biochem Soc Trans* **18** 789-792.

Olsson L, Stigson M, Perris R, Sorrell JM & Lofberg J 1996 Distribution of keratan sulphate and chondroitin sulphate in wild type and white mutant axolotl embryos during neural crest cell migration. *Pigment Cell Res.* **9** 5-17.

Otwinowski Z & Minor W 1997 Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276** 307-326.

Ovodov YS 2006a Bacterial capsular antigens. Structural patterns of capsular antigens. *Biochemistry (Mosc.)* **71** 937-954.

Ovodov YS 2006b Capsular antigens of bacteria. Capsular antigens as the basis of vaccines against pathogenic bacteria. *Biochemistry (Mosc.)* **71** 955-961.

Painter J & Merritt EA 2005 A molecular viewer for the analysis of TLS rigid-body motion in macromolecules. *Acta Crystallogr.D.Biol.Crystallogr.* **61** 465-471.

Pawelek PD, Cheah J, Coulombe R, Macheroux P, Ghisla S & Vrielink A 2000 The structure of L-amino acid oxidase reveals the substrate trajectory into an enantiomerically conserved active site. *EMBO J* **19** 4204-4215.

Penc SF, Pomahac B, Winkler T, Dorschner RA, Eriksson E, Herndon M & Gallo RL 1998 Dermatan sulfate released after injury is a potent promoter of fibroblast growth factor-2 function. *J Biol Chem* **273** 28116-28121.

Perrakis A, Morris R & Lamzin VS 1999 Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* **6** 458-463.

Perrimon N & Bernfield M 2000 Specificities of heparan sulphate proteoglycans in developmental processes. *Nature* **404** 725-728.

Perrimon N & Bernfield M 2001 Cellular functions of proteoglycans--an overview. *Semin.Cell Dev.Biol.* **12** 65-67.

Perutz MF 1960 Structure of hemoglobin. *Brookhaven.Symp.Biol.* **13** 165-183.

Pervin A, al Hakim A & Linhardt RJ 1994 Separation of glycosaminoglycan-derived oligosaccharides by capillary electrophoresis using reverse polarity. *Anal Biochem* **221** 182-188.

Pervin A, Gallo C, Jandik KA, Han XJ & Linhardt RJ 1995 Preparation and structural characterization of large heparin-derived oligosaccharides. *Glycobiology* **5** 83-95.

Peterson FC, Elgin ES, Nelson TJ, Zhang F, Hoeger TJ, Linhardt RJ & Volkman BF 2004 Identification and characterization of a glycosaminoglycan recognition element of the C chemokine lymphotactin. *J Biol Chem* **279** 12598-12604.

Petitou M, Herault JP, Bernat A, Driguez PA, Duchaussoy P, Lormeau JC & Herbert JM 1999 Synthesis of thrombin-inhibiting heparin mimetics without side effects. *Nature* **398** 417-422.

Pinhal MA, Smith B, Olson S, Aikawa J, Kimata K & Esko JD 2001 Enzyme interactions in heparan sulfate biosynthesis: uronosyl 5-epimerase and 2-O-sulfotransferase interact in vivo. *Proc.Natl.Acad.Sci.U.S.A* **98** 12984-12989.

Pojasek K, Raman R, Kiley P, Venkataraman G & Sasisekharan R 2002 Biochemical characterization of the chondroitinase B active site. *J Biol Chem* **277** 31179-31186.

Pojasek K, Shriver Z, Kiley P, Venkataraman G & Sasisekharan R 2001 Recombinant expression, purification, and kinetic characterization of chondroitinase AC and chondroitinase B from Flavobacterium heparinum. *Biochem.Biophys.Res.Commun* **286** 343-351.

Ponnuraj K & Jedrzejas MJ 2000 Mechanism of hyaluronan binding and degradation: structure of Streptococcus pneumoniae hyaluronate lyase in complex with hyaluronic acid disaccharide at 1.7 A resolution. *J Mol Biol* **299** 885-895.

Ponyi T, Szabo L, Nagy T, Orosz L, Simpson PJ, Williamson MP & Gilbert HJ 2000 Trp22, Trp24, and Tyr8 play a pivotal role in the binding of the family 10 cellulose-binding module from Pseudomonas xylanase A to insoluble ligands. *Biochemistry* **39** 985-991.

Prabhakar V, Capila I, Bosques CJ, Pojasek K & Sasisekharan R 2005 Chondroitinase ABC I from Proteus vulgaris: cloning, recombinant expression and active site identification. *Biochem J* **386** 103-112.

Prabhakar V, Capila I, Raman R, Srinivasan A, Bosques CJ, Pojasek K, Wrick MA & Sasisekharan R 2006 The catalytic machinery of chondroitinase ABC I utilizes a calcium coordination strategy to optimally process dermatan sulfate. *Biochemistry* **45** 11130-11139.

Prabhakar V, Raman R, Capila I, Bosques CJ, Pojasek K & Sasisekharan R 2005c Biochemical characterization of the chondroitinase ABC I active site. *Biochem J* **390** 395-405.

Prehm P 1983 Synthesis of hyaluronate in differentiated teratocarcinoma cells. Characterization of the synthase. *Biochem J* **211** 181-189.

Presto J, Thuveson M, Carlsson P, Busse M, Wilen M, Eriksson I, Kusche-Gullberg M & Kjellen L 2008 Heparan sulfate biosynthesis enzymes EXT1 and EXT2 affect NDST1 expression and heparan sulfate sulfation. *Proc.Natl.Acad.Sci.U.S.A* **105** 4751-4756.

Pritchard DG, Trent JO, Li X, Zhang P, Egan ML & Baker JR 2000 Characterization of the active site of group B streptococcal hyaluronan lyase. *Proteins* **40** 126-134.

Purushothaman A, Fukuda J, Mizumoto S, ten Dam GB, Van Kuppevelt TH, Kitagawa H, Mikami T & Sugahara K 2007 Functions of chondroitin sulfate/dermatan sulfate chains in brain development. Critical roles of E and iE disaccharide units recognized by a single chain antibody GD3G7. *J Biol Chem* **282** 19442-19452.

Quiocho FA, Spurlino JC & Rodseth LE 1997 Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. *Structure.* **5** 997-1015.

Raman R, Sasisekharan V & Sasisekharan R 2005 Structural insights into biological roles of protein-glycosaminoglycan interactions. *Chem Biol* **12** 267-277.

Rathore D, McCutchan TF, Garboczi DN, Toida T, Hernaiz MJ, LeBrun LA, Lang SC & Linhardt RJ 2001 Direct measurement of the interactions of glycosaminoglycans and a heparin decasaccharide with the malaria circumsporozoite protein. *Biochemistry* **40** 11518-11524.

Reinhold BB, Hauer CR, Plummer TH & Reinhold VN 1995 Detailed structural analysis of a novel, specific O-linked glycan from the prokaryote *Flavobacterium meningosepticum*. *J Biol Chem* **270** 13197-13203.

Rhomberg AJ, Ernst S, Sasisekharan R & Biemann K 1998a Mass spectrometric and capillary electrophoretic investigation of the enzymatic degradation of heparin-like glycosaminoglycans. *Proc.Natl.Acad.Sci.U.S.A* **95** 4176-4181.

Rhomberg AJ, Shriver Z, Biemann K & Sasisekharan R 1998b Mass spectrometric evidence for the enzymatic mechanism of the depolymerization of heparin-like glycosaminoglycans by heparinase II. *Proc.Natl.Acad.Sci.U.S.A* **95** 12232-12237.

Rider CC 1997 The potential for heparin and its derivatives in the therapy and prevention of HIV-1 infection. *Glycoconj.J.* **14** 639-642.

Rosenberg RD, Shworak NW, Liu J, Schwartz JJ & Zhang L 1997 Heparan sulfate proteoglycans of the cardiovascular system. Specific structures emerge but how is synthesis regulated? *J.Clin.Invest* **100** S67-S75.

Rostand KS & Esko JD 1997 Microbial adherence to and invasion through proteoglycans. *Infect.Immun.* **65** 1-8.

Ruoslahti E 1988 Structure and biology of proteoglycans. *Annu.Rev.Cell Biol.* **4** 229-255.

Rusnati M, Oreste P, Zoppetti G & Presta M 2005 Biotechnological engineering of heparin/heparan sulphate: a novel area of multi-target drug discovery. *Curr.Pharm.Des* **11** 2489-2499.

Ruter ER & Kresse H 1984 Partial purification and characterization of 3'-phosphoadenylylsulfate:keratan sulfate sulfotransferases. *J Biol Chem* **259** 11771-11776.

Rye CS, Matte A, Cygler M & Withers SG 2006 An atypical approach identifies TYR234 as the key base catalyst in chondroitin AC lyase. *Chembiochem.* **7** 631-637.

Saadi S, Wrenshall LE & Platt JL 2002 Regional manifestations and control of the immune system. *FASEB J* **16** 849-856.

Salmivirta M, Lidholt K & Lindahl U 1996 Heparan sulfate: a piece of information. *FASEB J* **10** 1270-1279.

Sambrook J & Russell D 2001 *Molecular Cloning: A Laboratory Manual (Third Edition)*. New York: Cold Spring Harbor Laboratory Press.

Sasisekharan R, Moses MA, Nugent MA, Cooney CL & Langer R 1994 Heparinase inhibits neovascularization. *Proc.Natl.Acad.Sci.U.S.A* **91** 1524-1528.

Sasisekharan R, Shriver Z, Venkataraman G & Narayanasami U 2002 Roles of heparan-sulphate glycosaminoglycans in cancer. *Nat.Rev.Cancer* **2** 521-528.

Sasisekharan R, Venkataraman G, Godavarti R, Ernst S, Cooney CL & Langer R 1996 Heparinase I from Flavobacterium heparinum. Mapping and characterization of the heparin binding domain. *J Biol Chem* **271** 3124-3131.

Sawitzky D 1996 Protein-glycosaminoglycan interactions: infectiological aspects. *Med.Microbiol.Immunol.* **184** 155-161.

Scavetta RD, Herron SR, Hotchkiss AT, Kita N, Keen NT, Benen JA, Kester HC, Visser J & Jurnak F 1999 Structure of a plant cell wall fragment complexed to pectate lyase C. *Plant Cell* **11** 1081-1092.

Schluenzen F, Tocilj A, Zarivach R, Harms J, Gluehmann M, Janell D, Bashan A, Bartels H, Agmon I, Franceschi F & Yonath A 2000 Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell* **102** 615-623.

Schmidt A, Gübitz GM & Kratky C 1999 Xylan binding subsite mapping in the xylanase from Penicillium simplicissimum using xylooligosaccharides as cryo-protectant. *Biochemistry* **38** 2403-2412.

Schmidtchen A, Frick IM & Bjorck L 2001 Dermatan sulphate is released by proteinases of common pathogenic bacteria and inactivates antibacterial alpha-defensin. *Mol.Microbiol.* **39** 708-713.

Schwartz NB & Domowicz M 2002 Chondrodysplasias due to proteoglycan defects. *Glycobiology* **12** 57R-68R.

Senay C, Lind T, Muguruma K, Tone Y, Kitagawa H, Sugahara K, Lidholt K, Lindahl U & Kusche-Gullberg M 2000 The EXT1/EXT2 tumor suppressors: catalytic activities and role in heparan sulfate biosynthesis. *EMBO Rep.* **1** 282-286.

Sharon N & Lis H 2002 How proteins bind carbohydrates: lessons from legume lectins. *J.Agric.Food Chem.* **50** 6586-6591.

Shaya D, Hahn BS, Bjerkan TM, Kim WS, Park NY, Sim JS, Kim YS & Cygler M 2008 Composite Active Site of Chondroitin Lyase ABC Accepting Both Epimers of Uronic Acid. *Glycobiology* **18** 270-277.

Shaya D, Li Y & Cygler M 2004 Crystallization and preliminary X-ray analysis of heparinase II from Pedobacter heparinus. *Acta Crystallogr.D.Biol.Crystallogr.* **60** 1644-1646.

Shaya D, Tocilj A, Li Y, Myette J, Venkataraman G, Sasisekharan R & Cygler M 2006 Crystal structure of heparinase II from Pedobacter heparinus and its complex with a disaccharide product. *J Biol Chem* **281** 15525-15535.

Sheldrick GM 1991 Heavy atom location using SHELX-90. In *Isomorphous replacement and Anomalous Scattering*, pp 80-86. Eds W Wolf, PR Evans & AG Leslie. Warrington: SERC Daresbury Laboratory.

Shriver Z, Hu Y, Pojasek K & Sasisekharan R 1998a Heparinase II from *Flavobacterium heparinum*. Role of cysteine in enzymatic activity as probed by chemical modification and site-directed mutagenesis. *J Biol Chem* **273** 22904-22912.

Shriver Z, Hu Y & Sasisekharan R 1998b Heparinase II from *Flavobacterium heparinum*. Role of histidine residues in enzymatic activity as probed by chemical modification and site-directed mutagenesis. *J Biol Chem* **273** 10160-10167.

Shriver Z, Raguram S & Sasisekharan R 2004 Glycomics: a pathway to a class of new and improved therapeutics. *Nat.Rev.Drug Discov.* **3** 863-873.

Shukla D, Liu J, Blaiklock P, Shworak NW, Bai X, Esko JD, Cohen GH, Eisenberg RJ, Rosenberg RD & Spear PG 1999 A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell* **99** 13-22.

Siegelman MH, DeGrendele HC & Estess P 1999 Activation and interaction of CD44 and hyaluronan in immunological systems. *J.Leukoc.Biol.* **66** 315-321.

Steyn PL, Segers P, Vancanneyt M, Sandra P, Kersters K & Joubert JJ 1998 Classification of heparinolytic bacteria into a new genus, Pedobacter, comprising four species: Pedobacter heparinus comb. nov., Pedobacter piscium comb. nov., Pedobacter africanus sp. nov. and Pedobacter saltans sp. nov. proposal of the family Sphingobacteriaceae fam. nov. *Int.J Syst.Bacteriol.* **48 Pt 1** 165-177.

Stout GH & Jensen LH 1989 *X-Ray structure Determination*. Seattle: John Wiley and Sons.

Stringer SE & Gallagher JT 1997 Heparan sulphate. *Int J Biochem Cell Biol* **29** 709-714.

Strynadka NC & James MN 1991 Lysozyme revisited: crystallographic evidence for distortion of an N-acetylmuramic acid residue bound in site D. *J Mol Biol* **220** 401-424.

Strynadka NC & James MN 1996 Lysozyme: a model enzyme in protein crystallography. *EXS* **75** 185-222.

Su H, Blain F, Musil RA, Zimmermann JJ, Gu K & Bennett DC 1996 Isolation and expression in *Escherichia coli* of hepB and hepC, genes coding for the glycosaminoglycan-degrading enzymes heparinase II and heparinase III, respectively, from *Flavobacterium heparinum*. *Appl.Environ.Microbiol.* **62** 2723-2734.

Su H, Shao Z, Tkalec L, Blain F & Zimmermann J 2001 Development of a genetic system for the transfer of DNA into Flavobacterium heparinum. *Microbiology* **147** 581-589.

Sugahara K & Kitagawa H 2000 Recent advances in the study of the biosynthesis and functions of sulfated glycosaminoglycans. *Curr.Opin.Struct.Biol* **10** 518-527.

Sugahara K & Kitagawa H 2002 Heparin and heparan sulfate biosynthesis. *IUBMB.Life* **54** 163-175.

Sugahara K & Mikami T 2007 Chondroitin/dermatan sulfate in the central nervous system. *Curr.Opin.Struct.Biol* **17** 536-545.

Sugahara K, Mikami T, Uyama T, Mizuguchi S, Nomura K & Kitagawa H 2003 Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate. *Curr.Opin.Struct.Biol* **13** 612-620.

Sugahara K, Tanaka Y, Yamada S, Seno N, Kitagawa H, Haslam SM, Morris HR & Dell A 1996 Novel sulfated oligosaccharides containing 3-O-sulfated glucuronic acid from king crab cartilage chondroitin sulfate K. Unexpected degradation by chondroitinase ABC. *J Biol Chem* **271** 26745-26754.

Tai GH, Huckerby TN & Nieduszynski IA 1996 Multiple non-reducing chain termini isolated from bovine corneal keratan sulfates. *J Biol Chem* **271** 23535-23546.

Tai GH, Nieduszynski IA, Fullwood NJ & Huckerby TN 1997 Human corneal keratan sulfates. *J Biol Chem* **272** 28227-28231.

Tanner ME 2002 Understanding nature's strategies for enzyme-catalyzed racemization and epimerization. *Acc Chem Res* **35** 237-246.

Taylor KR & Gallo RL 2006 Glycosaminoglycans and their proteoglycans: host-associated molecular patterns for initiation and modulation of inflammation. *FASEB J* **20** 9-22.

Tekotte H, Engel M, Margolis RU & Margolis RK 1994 Disaccharide composition of heparan sulfates: brain, nervous tissue storage organelles, kidney, and lung. *J.Neurochem.* **62** 1126-1130.

Termeer C, Sleeman JP & Simon JC 2003 Hyaluronan--magic glue for the regulation of the immune response? *Trends Immunol.* **24** 112-114.

Terwilliger TC 2002 Automated structure solution, density modification and model building. *Acta Crystallogr.D Biol.Crystallogr.* **58** 1937-1940.

Toole BP 2001 Hyaluronan in morphogenesis. *Semin.Cell Dev.Biol.* **12** 79-87.

Toole BP 2004 Hyaluronan: from extracellular glue to pericellular cue. *Nat.Rev.Cancer* **4** 528-539.

Toole BP & Slomiany MG 2008 Hyaluronan: a constitutive regulator of chemoresistance and malignancy in cancer cells. *Semin.Cancer Biol.* **18** 244-250.

Toyoshima M & Nakajima M 1999 Human heparanase. Purification, characterization, cloning, and expression. *J Biol Chem* **274** 24153-24160.

Tropea D, Caleo M & Maffei L 2003 Synergistic effects of brain-derived neurotrophic factor and chondroitinase ABC on retinal fiber sprouting after denervation of the superior colliculus in adult rats. *J.Neurosci.* **23** 7034-7044.

Trowbridge JM & Gallo RL 2002 Dermatan sulfate: new functions from an old glycosaminoglycan. *Glycobiology* **12** 117R-125R.

Uchimura K, Muramatsu H, Kadomatsu K, Fan QW, Kurosawa N, Mitsuoka C, Kannagi R, Habuchi O & Muramatsu T 1998 Molecular cloning and characterization of an N-acetylglucosamine-6-O-sulfotransferase. *J Biol Chem* **273** 22577-22583.

Vagin A & Teplyakov A 1997 MOLREP: an Automated Program for Molecular Replacement. *J Appl Crystallogr* **30** 1022-1025.

Valla S, Li J, Ertesvag H, Barbeyron T & Lindahl U 2001 Hexuronyl C5-epimerases in alginate and glycosaminoglycan biosynthesis. *Biochimie* **83** 819-830.

Varki A 1993 Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3** 97-130.

Vedadi M, Niesen FH, lali-Hassani A, Fedorov OY, Finerty PJ, Jr., Wasney GA, Yeung R, Arrowsmith C, Ball LJ, Berglund H, Hui R, Marsden BD, Nordlund P, Sundstrom M, Weigelt J & Edwards AM 2006 Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc.Natl.Acad.Sci.U.S.A* **103** 15835-15840.

Venkataraman G, Shriver Z, Raman R & Sasisekharan R 1999 Sequencing complex polysaccharides. *Science* **286** 537-542.

Vimr ER, Kalivoda KA, Deszo EL & Steenbergen SM 2004 Diversity of microbial sialic acid metabolism. *Microbiol.Mol.Biol.Rev.* **68** 132-153.

Vlodavsky I, bboud-Jarrous G, Elkin M, Naggi A, Casu B, Sasisekharan R & Ilan N 2006 The impact of heparanese and heparin on cancer metastasis and angiogenesis. *Pathophysiol.Haemost.Thromb.* **35** 116-127.

Vlodavsky I, Fuks Z, Ishai-Michaeli R, Bashkin P, Levi E, Korner G, Bar-Shavit R & Klagsbrun M 1991 Extracellular matrix-resident basic fibroblast growth factor: implication for the control of angiogenesis. *J.Cell Biochem.* **45** 167-176.

von Laue M 1912 *Sitz.Math.Phys.Klasee Bayer Akad.Wiss.* **303**.

Vyas NK, Vyas MN & Quiocho FA 1991 Comparison of the periplasmic receptors for L-arabinose, D-glucose/D-galactose, and D-ribose. Structural and Functional Similarity. *J Biol Chem* **266** 5226-5237.

Weigel PH & DeAngelis PL 2007 Hyaluronan synthases: a decade-plus of novel glycosyltransferases. *J Biol Chem* **282** 36777-36781.

Weigel PH, Fuller GM & LeBoeuf RD 1986 A model for the role of hyaluronic acid and fibrin in the early events during the inflammatory response and wound healing. *J.Theor.Biol.* **119** 219-234.

Weigel PH, Hascall VC & Tammi M 1997 Hyaluronan synthases. *J Biol Chem* **272** 13997-14000.

Wolff JJ, Amster IJ, Chi L & Linhardt RJ 2007 Electron detachment dissociation of glycosaminoglycan tetrasaccharides. *J.Am.Soc.Mass Spectrom.* **18** 234-244.

Wormald MR, Petrescu AJ, Pao YL, Glithero A, Elliott T & Dwek RA 2002 Conformational studies of oligosaccharides and glycopeptides: complementarity of NMR, X-ray crystallography, and molecular modelling. *Chem.Rev.* **102** 371-386.

Xu J, Bjursell MK, Himrod J, Deng S, Carmichael LK, Chiang HC, Hooper LV & Gordon JI 2003 A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. *Science* **299** 2074-2076.

Yamagata T, Saito H, Habuchi O & Suzuki S 1968 Purification and properties of bacterial chondroitinases and chondrosulfatases. *J Biol Chem* **243** 1523-1535.

Yamasaki M, Ogura K, Hashimoto W, Mikami B & Murata K 2005 A structural basis for depolymerization of alginate by polysaccharide lyase family-7. *J Mol Biol* **352** 11-21.

Yanagishita M & Hascall VC 1992 Cell surface heparan sulfate proteoglycans. *J Biol Chem* **267** 9451-9454.

Yick LW, Wu W, So KF, Yip HK & Shum DK 2000 Chondroitinase ABC promotes axonal regeneration of Clarke's neurons after spinal cord injury. *Neuroreport* **11** 1063-1067.

Yoon HJ, Hashimoto W, Miyake O, Murata K & Mikami B 2001 Crystal structure of alginate lyase A1-III complexed with trisaccharide product at 2.0 A resolution. *J Mol Biol* **307** 9-16.

Yoon HJ, Mikami B, Hashimoto W & Murata K 1999 Crystal structure of alginate lyase A1-III from Sphingomonas species A1 at 1.78 A resolution. *J Mol Biol* **290** 505-514.

Zhang F, McLellan JS, Ayala AM, Leahy DJ & Linhardt RJ 2007 Kinetic and structural studies on interactions between heparin or heparan sulfate and proteins of the hedgehog signaling pathway. *Biochemistry* **46** 3933-3941.

Zhang JP & Stephens RS 1992 Mechanism of C. trachomatis attachment to eukaryotic host cells. *Cell* **69** 861-869.

Appendix

Copyright Waivers

permissions requests for chapter 2

Reuse permissions

Requests for permission to reuse material from this journal should be addressed to:

Managing Editor
IUCr Journals
5 Abbey Square
Chester CH1 2HU
UK

Email: med@iucr.org

Please be sure to include the following information:

- The title, author(s) and page extent of the article you wish to republish, or, in the case where you do not wish to reproduce the whole article, details of the section(s) to be reproduced.
- The journal title, volume/issue number and year of publication in which the article appeared, and the page numbers of the article.
- Details of the publication in which you wish to reprint the Journal material.
- If the material is being edited or amended in any way, a copy of the final material as it will appear in your publication.
- Contact details for yourself, including a postal address.

If you wish to include material from **your own article** in a publication, prior permission is not required, subject to the following conditions:

- Reproduction is intended in a primary journal, secondary journal, CD-ROM, book or thesis.
- The original article in which the material appeared is cited.
- IUCr's copyright permission is indicated by the wording "Reproduced with permission of the International Union of Crystallography". In electronic form, this acknowledgement must be visible at the same time as the reused materials, and must be hyperlinked to **Crystallography Journals Online** (http://journals.iucr.org/).

Photocopying permissions

Teaching institutions with a current paid subscription to the journal may make multiple copies for teaching purposes without charge, provided such copies are not resold. In all other cases, permission should be obtained from a reproduction rights organisation (see below) or directly from the Publisher.

Permission to photocopy print versions of this journal varies by territory.

USA: Please obtain permission via the CCC:

Copyright Clearance Center
222 Rosewood Drive

Danvers, MA 01923
USA
Tel: +1(508) 750 8400
Fax: +1 (508) 750 4744

This journal is registered with the Copyright Licensing Agency, and licensed institutions may make copies under the terms of their licence.

For multiple copies and coursepacks permission should be cleared through CLARCS, the CLA Rapid Clearance Service:

Education, tel: 020 7631 5560 Business and Government, tel: 020 7631 5561; email: CLARCS

If you do not have a CLA licence, please apply direct to the Publisher.

Other Territories: Please contact your local reproduction rights organisation (see below), or apply direct to the Publisher.

- Australia* Copyright Agency Ltd (CAL)
- Austria (Literar-Mechana)
- Belgium (ReproBel)
- Brazil (ABDR)
- Canada* Access Copyright and Quebec* (COPIBEC)
- Denmark* (CopyDan)
- France* (C.F.C.)
- Finland* (Kopiosto)
- Germany* (VGWort & VG Musikedition/VG Bild-Kunst)
- Greece* (OSDEL)
- Italy (AIDRO)
- Iceland* (Fjolis)
- Ireland* (ICLA)
- Japan (JRRC)
- Kenya (KOPIKEN)
- Malta (Kopjamalt)
- The Netherlands* (Stichting Reprorecht)
- New Zealand* (CLL)
- Norway* Kopinor
- South Africa* (DALRO)
- Spain* (CEDRO)
- Sweden* (BONUS)
- Switzerland* (Pro Litteris)
- Zimbabwe (ZIMCOPY)

Copyright Permission Policy for chapter 3

ASBMB Journals
Journal of Biological Chemistry
Molecular and Cellular Proteomics
Journal of Lipid Research
ASBMB Today

ASBMB does not charge for and grants use without requiring your copyright permission request for:

- The Journal of Biological Chemistry is copyrighted by the American Society for Biochemistry and Molecular Biology, Inc. ASBMB grants use without requiring your copyright permission request for: Original authors wanting to reproduce figures or tables from their own work to publish in not-for-profit formats or venues, provided that full acknowledgment of the source is provided in the new work.
- Students wanting to reproduce or republish their work for educational purposes.
- Students using other authors' material for their theses.
- Reproduction or republication of abstracts only.
- Photocopying up to 5 copies for personal use.
- Non-profit educational institutions making multiple photocopies of articles for classroom use; all such reproduction must utilize institutionally owned equipment for this purpose.

Use of copyrighted material requires proper citation.

For all other uses, contact Copyright Clearance Center.

August 23, 2008

McGill University
6100 Royalmount Avenue
Montreal, Quebec
H4P 2R2, Canada

Re: David Shaya

To whom it may concern:

I hereby authorize David Shaya to include all the data in our upcoming joint publications including unpublished results associated with this study on heparinase mutants in his Doctoral Thesis.

Sincerely,

Robert J. Linhardt, Ph.D.
Acting Director of the Center for Biotechnology and Interdisciplinary Studies
Ann and John H. Broadbent, Jr. '59
Senior Constellation Professor,
Biocatalysis and Metabolic Engineering
Professor of Chemistry and Chemical Biology,
Biology and Chemical and Biological Engineering

Rensselaer Polytechnic Institute
Department of Chemistry, Biotech Building 4005
110 8th Street Troy, NY 12180-3590
Phone (518) 276-3404 Fax (518) 276-3405

McGill

Department of Biochemistry

McGill University
McIntyre Medical Sciences Building

3655 Promenade Sir William Osler
Montreal, Quebec, Canada H3G 1Y6

Tel.: (514) 398-
Fax: (514) 398-738

August 25, 2008

McGill University
6100 Royalmount Avenue
Montreal, Quebec
H4P 2R2, Canada

Re: David Shaya

To whom it may concern:

I hereby authorize David Shaya to include all the data in our upcoming joint publications including unpublished results associated with this study on heparinase mutants in his Doctoral Thesis.

Sincerely,

Mirek Cygler
Adjunct Professor, Departments of Biochemistry,
Anatomy and Cell Biology, McGill University,
Principal Research Officer and Head
Macromolecular Structure Group
Biotechnology Research Institute, NRC
6100 Royalmount Avenue, Montreal, Quebec H4P 2R2, Canada

OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS

for chapter 5

Jun 04, 2008

This is a License Agreement between David Shaya ("You") and Oxford University Press ("Oxford University Press"). The license consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

License Number	1962050379469
License date	Jun 04, 2008
Licensed content publisher	Oxford University Press
Licensed content publication	Glycobiology
Licensed content title	Composite active site of chondroitin lyase ABC accepting both epimers of uronic acid
Licensed content author	D. Shaya, et. al.
Licensed content date	March 2008
Type of Use	Thesis / Dissertation
Institution name	McGill University
Title of your work	Glycoseaminoglycan degradation - structural studies of heparinase II and chondroitin lyase ABC II
Publisher of your work	McGill University
Expected publication date	12/01/2008
Permissions cost	0.00 USD
Value added tax	0.00 USD
Total	0.00 USD

Terms and Conditions

STANDARD TERMS AND CONDITIONS FOR REPRODUCTION OF MATERIAL FROM AN OXFORD UNIVERSITY PRESS JOURNAL

1. Use of the material is restricted to your license details specified during the order process.

2. This permission covers the use of the material in the English language in the following territory: world. For permission to translate any material from an Oxford University Press journal into another language, please email journals.permissions@oxfordjournals.org

3. This permission is limited to the particular use authorized in (1) above and does not allow you to sanction its use elsewhere in any other format other than specified above, nor does it apply to quotations, images, artistic works etc that have been reproduced from other sources which may be part of the material to be used.

4. No alteration, omission or addition is made to the material without our written consent.

Permission must be re-cleared with Oxford University Press if/when you decide to reprint.

5. The following credit line appears wherever the material is used: author, title, journal, year, volume, issue number, pagination, by permission of Oxford University Press or the sponsoring society if the journal is a society journal. Where a journal is being published on behalf of a learned society, the details of that society must be included in the credit line.

6. For the reproduction of a full article from an Oxford University Press journal for whatever purpose, the corresponding author of the material concerned should be informed of the proposed use. Contact details for the corresponding authors of all Oxford University Press journal contact can be found alongside either the abstract or full text of the article concerned, accessible from www.oxfordjournals.org. Should there be a problem clearing these rights, please contact journals.permissions@oxfordjournals.org

7. If the credit line or acknowledgement in our publication indicates that any of the figures, images or photos was reproduced, drawn or modified from an earlier source it will be necessary for you to clear this permission with the original publisher as well. If this permission has not been obtained, please note that this material cannot be included in your publication/photocopies.

8. While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by Oxford University Press or by Copyright Clearance Center (CCC)) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and Oxford University Press reserves the right to take any and all action to protect its copyright in the materials.

9. This license is personal to you and may not be sublicensed, assigned or transferred by you to any other person without Oxford University Press's written permission.

10. Oxford University Press reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

11. You hereby indemnify and agree to hold harmless Oxford University Press and CCC, and their respective officers, directors, employs and agents, from and against any and all claims arising our to your use of the licensed material other than as specifically authorized pursuant to this license.

v1.1

Other Terms and Conditions

American Chemical Society

for chpter6

Publications Division
Copyright Office

1155 Sixteenth Street, NW
Washington, DC 20036
Phone: (1) 202-872-4368 or -4367
Fax: (1) 202-776-8112 E-mail: copyright@acs.org

VIA FAX: 514-496-5143 DATE: June 11, 2008

TO: David Shaya, McGill University Macromolecular Group
Biotechnology Research Institute, NRC, 6100 Royal Mount Ave., Montreal, QC H4P 2R2, Canada

FROM: C. Arleen Courtney, Copyright Associate C. Arleen Courtney

Thank you for your request for permission to include **your** paper(s) or portions of text from **your** paper(s) in your thesis. Permission is now automatically granted; please pay special attention to the implications paragraph below. The Copyright Subcommittee of the Joint Board/Council Committees on Publications approved the following:

Copyright permission for published and submitted material from theses and dissertations
ACS extends blanket permission to students to include in their theses and dissertations their own articles, or portions thereof, that have been published in ACS journals or submitted to ACS journals for publication, provided that the ACS copyright credit line is noted on the appropriate page(s).

Publishing implications of electronic publication of theses and dissertation material
Students and their mentors should be aware that posting of theses and dissertation material on the Web prior to submission of material from that thesis or dissertation to an ACS journal may affect publication in that journal. Whether Web posting is considered prior publication may be evaluated on a case-by-case basis by the journal's editor. If an ACS journal editor considers Web posting to be "prior publication", the paper will not be accepted for publication in that journal. If you intend to submit your unpublished paper to ACS for publication, check with the appropriate editor prior to posting your manuscript electronically.

If your paper has not yet been published by ACS, we have no objection to your including the text or portions of the text in your thesis/dissertation in **print and microfilm formats**; please note, however, that electronic distribution or Web posting of the unpublished paper as part of your thesis in electronic formats might jeopardize publication of your paper by ACS. Please print the following credit line on the first page of your article: "Reproduced (or 'Reproduced in part') with permission from [JOURNAL NAME], in press (or 'submitted for publication'). Unpublished work copyright [CURRENT YEAR] American Chemical Society." Include appropriate information.

If your paper has already been published by ACS and you want to include the text or portions of the text in your thesis/dissertation in **print or microfilm formats**, please print the ACS copyright credit line on the first page of your article: "Reproduced (or 'Reproduced in part') with permission from [FULL REFERENCE CITATION.] Copyright [YEAR] American Chemical Society." Include appropriate information.

Submission to a Dissertation Distributor: If you plan to submit your thesis to UMI or to another dissertation distributor, you should not include the unpublished ACS paper in your thesis if the thesis will be disseminated electronically, until ACS has published your paper. After publication of the paper by ACS, you may release the entire thesis (**not the individual ACS article by itself**) for electronic dissemination through the distributor; ACS's copyright credit line should be printed on the first page of the ACS paper.

Use on an Intranet: The inclusion of your ACS unpublished or published manuscript is permitted in your thesis in print and microfilm formats. If ACS has published your paper you may include the manuscript in your thesis on an intranet that is not publicly available. Your ACS article cannot be posted electronically on a publicly available medium (i.e. one that is not password protected), such as but not limited to, electronic archives, Internet, library server, etc. The only material from your paper that can be posted on a public electronic medium is the article abstract, figures, and tables, and you may link to the article's DOI or post the article's author-directed URL link provided by ACS. This paragraph does not pertain to the dissertation distributor paragraph above.

06/07/06

Karen Buehler

From: David Shaya [david.shaya@bri.nrc.ca]
Sent: Wednesday, June 04, 2008 3:55 PM
To: Copyright
Subject: copyright permission for publication for a manuscript published in "Biochemistry"

To whom it may concern,
I would like to include the text, tables, figures and supplementary material from the following publication in my thesis. I am the first author on this publication. The paper is:

ASAP Biochemistry, **ASAP Article**, 10.1021/bi800353g
Web Release Date: May 31, 2008

Characterization of Chondroitin Sulfate Lyase ABC from Bacteroides thetaiotaomicron WAL2926

David Shaya, Bum-Soo Hahn, Nam Young Park, Joon-Soo Sim, Yeong Shik Kim, and Miroslaw Cygler

Department of Biochemistry, McGill University, Montreal, Quebec, Canada, National Institute of Agricultural Biotechnology, 225 Seodun-Dong, Suwon 441-707, South Korea, Natural Products Research Institute, College of Pharmacy, Seoul National University, Seoul 151-742, South Korea, and Biotechnology Research Institute, NRC, 6100 Royalmount Avenue, Montre´al, Que´bec, Canada H4P 2R2

Received February 29, 2008

Revised Manuscript Received May 14, 2008

My contacts are :

David Shaya
Graduate student, McGill University
Macromolecular Structure Group
Biotechnology Research Institute, NRC
6100 Royal mount Av., Montreal, Quebec
H4P 2R2, Canada
e-mail : david.shaya@bri.nrc.ca
Phone : (514) 496-6173
Fax: (514) 496-5143
Mobile: (514) 824-5809

Regards
David

American Chemical Society's Policy on Theses and Dissertations

If your university requires a signed copy of this letter see contact information below.

Thank you for your request for permission to include your paper(s) or portions of text from your paper(s) in your thesis. Permission is now automatically granted; please pay special attention to the implications paragraph below. The Copyright Subcommittee of the Joint Board/Council Committees on Publications approved the following:

Copyright permission for published and submitted material from theses and dissertations
ACS extends blanket permission to students to include in their theses and dissertations their own articles, or portions thereof, that have been published in ACS journals or submitted to ACS journals for publication, provided that the ACS copyright credit line is noted on the appropriate page(s).

Publishing implications of electronic publication of theses and dissertation material
Students and their mentors should be aware that posting of theses and dissertation material on the Web prior to submission of material from that thesis or dissertation to an ACS journal may affect publication in that journal. Whether Web posting is considered prior publication may be evaluated on a case-by-case basis by the journal's editor. If an ACS journal editor considers Web posting to be "prior publication", the paper will not be accepted for publication in that journal. If you intend to submit your unpublished paper to ACS for publication, check with the appropriate editor prior to posting your manuscript electronically.

If your paper has **not** yet been published by ACS, we have no objection to your including the text or portions of the text in your thesis/dissertation in **print and microfilm formats**; please note, however, that electronic distribution or Web posting of the unpublished paper as part of your thesis in electronic formats might jeopardize publication of your paper by ACS. Please print the following credit line on the first page of your article: "Reproduced (or 'Reproduced in part') with permission from [JOURNAL NAME], in press (or 'submitted for publication'). Unpublished work copyright [CURRENT YEAR] American Chemical Society." Include appropriate information.

If your paper has already been published by ACS and you want to include the text or portions of the text in your thesis/dissertation in **print or microfilm formats**, please print the ACS copyright credit line on the first page of your article: "Reproduced (or 'Reproduced in part') with permission from [FULL REFERENCE CITATION.] Copyright [YEAR] American Chemical Society." Include appropriate information.

Submission to a Dissertation Distributor: If you plan to submit your thesis to UMI or to another dissertation distributor, you should not include the unpublished ACS paper in your thesis if the thesis will be disseminated electronically, until ACS has published your paper. After publication of the paper by ACS, you may release the entire thesis (**not the individual ACS article by itself**) for electronic dissemination through the distributor; ACS's copyright credit line should be printed on the first page of the ACS paper.

Use on an Intranet: The inclusion of your ACS unpublished or published manuscript is permitted in your thesis in print and microfilm formats. If ACS has published your paper you may include the manuscript in your thesis on an intranet that is not publicly available. Your ACS article cannot be posted electronically on a publicly available medium (i.e. one that is not password protected), such as but not limited to, electronic archives, Internet, library server, etc. The only material from your paper that can be posted on a public electronic medium is the article abstract, figures, and tables, and you may link to the article's DOI or post the article's author-directed URL link provided by ACS. This paragraph does not pertain to the dissertation distributor paragraph above.

Questions? Call +1 202-872-4368/4367. Send e-mail to copyright@acs.org or fax to +1 202-776-8112. 10/10/03, 01/15/04, 06/07/06