Protein Engineering of Heparinase I—
Elucidation of Structure-Activity Relationships

by

Ranganathan S. Godavarti

Submitted to the Department of Chemical Engineering

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Science

in

Chemical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 1996

Author _____

Department of Chemical Engineering

Certified by _____ _____

Prof. Charles L. Cooney

Thesis Supervisor, Department of Chemical Engineering

Accepted by _____ _____

Prof. Robert E. Cohen

Chairman, Committee for Graduate Students

Protein Engineering of Heparinase I—
Elucidation of Structure-Activity Relationships

by

Ranga Godavarti

Submitted to the Department of Chemical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Science in Chemical Engineering

ABSTRACT

Heparin-like molecules are complex polysaccharides that interact with various proteins and regulate several important physiological functions in the body. Heparinases are heparin degrading enzymes that cleave heparin like molecules with a high degree of specificity. They have significant therapeutic and clinical applications. Heparinase I is in phase I clinical trials for systemic heparin neutralization. The use of heparinase I *in vitro*, for clinical monitoring of heparin concentration has been approved by the FDA. However, the molecular basis for heparin-heparinase interactions are not understood. Understanding the structure-activity relationships of heparinase I would provide a framework towards the development of enzymes with improved properties for a variety of applications. This work investigates structure-activity relationships of heparinase I.

Previous studies identified cysteine-135 in the active site. In this study, first the role of other residues critical for catalytic activity of heparinase I was investigated. Heparinase I was inactivated in the presence of reversible histidine modifying diethylpyrocarbonate (DEPC) in a concentration dependent manner; 0.3 mM DEPC resulted in 95 % of heparinase I inactivation in less than 3 min and as low as 10 μM DEPC resulted in a 85 % loss of heparinase I activity in 15 min. Heparinase I activity was restored following hydroxylamine treatment. This, along with other experiments, strongly suggested that the inactivation of heparinase I by DEPC was specific for histidine residues. Chemical modification of the histidines under non-denaturing conditions, using non-radiolabeled and [^{14}C] DEPC indicated that between one and two histidine residues were modified. Heparin protection against inactivation suggested that the histidines lie in or near the active site of heparinase I. Site-directed mutagenesis of H129A, H165A and H339A did not affect enzyme activity, suggesting that these residues were not essential for heparinase I activity. However, H203A inactivated heparinase I while a H203D mutant had residual activity, indicating a role of this residue in catalysis. It is proposed that histidine-203, contained in the heparin binding site, is immediately adjacent to cysteine-135, and these residues together form the catalytic domain of heparinase I.

The role of positive charge in the heparin binding domain in heparinase I activity was investigated next through extensive site-directed mutagenesis studies. Results indicated lysines-198 and -199 were critical for heparinase I activity and altering the positive charge of these residues modulates the product profile. Altering glycine-213 (a critical conserved residue of the consensus calcium coordination motif) dramatically affected heparinase I activity, suggesting the importance of this motif. Further, lysine-132, close to active site cysteine-135 in the primary sequence, was also essential for enzymatic activity. Importantly, these findings are consistent with the hypothesis that apart from a linear heparin binding sequence a three-dimensional domain of basic residues around the catalytic site is involved in both, heparin binding through charge complimentarity, as well as providing the basic environment for cysteine-135. A model of the substrate binding cleft in heparinase I is proposed with the hypothesis that the substrate binding pocket contains the catalytic domain comprising cysteine-135 and histidine-203. Model predictions were tested through the effect of steric mutations on enzyme activity.

The individual roles of the catalytically critical residues in catalysis was investigated through systematic pH analysis and site-directed mutagenesis studies to identify the ionization states of these residues. Results from pH-rate analysis showed that cysteine-135 with a pKa of 5.5 is unprotonated and histidine-203 with a pKa of 8.2 is protonated for catalysis. The unusually low pKa for cysteine was attributed to the positive charge around it and this charged environment (basic residues from the heparin binding domain) perhaps stabilizes the anionic form of the thiol group. Site-directed mutagenesis studies on cysteine-135 and histidine-203 with residues with varying nucleophilicities suggested that cysteine-135 plays a nucleophilic role and histidine-203 acts as an acid/base catalyst. A mechanism for heparin degradation that is consistent with data, from both this research as well as previous work, is proposed.

Dedicated to

Amma

who created a dream for me .....

Appa

who stood by me and convinced me that it could come true.....

&

Raja

who made the dream come true.

Acknowledgments

Several people have contributed, either directly or indirectly, towards making this work possible and I take this opportunity to express my gratitude to all of them. I have made a sincere and humble effort to mention and thank everyone who helped me see this through—but I am positive, even before I begin, that I would have failed in my efforts! But atleast I tried...! Charlie Cooney has been a great advisor to work with—thank you for giving me the freedom to work on several different projects and for teaching me the concept of "maturing as a researcher"; it was a great learning experience overall. Alan Hatton, Bob Langer and Phil Robbins have been a great thesis committee and I thank you all for your support and encouragement. Special thanks to Phil Robbins, for making the Cancer Center facilities available.

Although I much wanted to avoid using the cliche' "friend, philosopher and guide", Ramnath Sasisekharan has played each of these roles throughout these past few years—as a friend, always there whenever I wanted him; also available for the several philosophical "LIAB" conversations; and as a guide and mentor during the entire project! He has been a constant source of motivation and encouragement and it would not be an exaggeration to say that but for him a lot of this would not be complete! I know that to *you*, I don't need to say thank you, Ramnath!

My lab colleagues—Guillermo, Matthew, Chris, Steffen, John Konz, and Stefan Winkler made working at MIT fun and enjoyable. I shall always remember Humberto, Charlie and several others in the Robbins lab and Margaret (Solomon lab) for their help and for letting me barge into their lab often! I can not forget to mention Mark Bulmer for helping me get my hands wet in the lab for the first time and John Augliera..thank you for all the screening and pouring those million sequencing gels! Several undergraduate students deserve credit for diligently carrying out their research tasks and enabling my work to proceed smoothly—Chia, Tamara, Sayee, Erika, Lisa, Janet, Juan Luis Quintero, and Sung. Sung..thanks for everything, including the computer..hope some day you learn the importance of "order and method"!! I am grateful to Elaine Aufiero-Peters, Pam Brown, Janet Fischer for making "non-research life at MIT" so efficient & easy! And special thanks to "mommy" Carol..for everything!

Rajaram aunty and uncle have always been very special people; thank you for making Bombay my second home! Deepa, aunty and uncle—I can never forget your support and affection during my most trying times. Sekhar, thanks for the great song collection for my homesick days! Some day I hope to beat you! I shall remain indebted to Nandy aunty for being responsible for the turning point in my career. Atul Kumar, your friendship I will always cherish—*mujhe shukriya kahne ki zaroorat nahin*!

My friends have been a great source of strength and motivation and I shall be ever grateful to them for their affection, help and support. Ratna and Asha—I enjoyed the *Burnside Avenue* days and it has influenced me in several ways. Rat, thanks for everything (including, making me a cleaning maniac!)—for giving me a different perspective...and for being a great friend! Ash, thank you for making me view things in different light (thanks to the zillion arguments!); for being around whenever I needed; for being a constant source of moral support..etc..fortunately, to you I know that all I need to say is...*you know*! Kala and Vijay have been great fun and New Years and Times Sq. will have an entirely new meaning in life, thanks to you guys! Kela, you have been very supportive and a wonderful, trusted friend—I shall always remember the hours and hours of phone conversations we shared! In short, *thanx lah, Kelah*! Asha Girimaji, Ben and Preeti, Chitra Dorai, Maniyah and Shampaji, Oscar and Ponni, Ramesh and Ash, Ski and Anu have been great friends and thank you all for teaching me that there is life outside MIT! Ganesh and Ramprasad—I will always remember you for making my transition into the US smooth and easy; thanks guys!

And finally, I don't have words to express my gratitude to my family. Chitta and chitti, you stood by me through the worst nightmare of my life and gave me the strength and courage to take up this challenge. Chitta, I can never forget your taking the burden and responsibility in making me come out of those difficult times. Sangeeta, its great to have you in the family and to be a family of four again. I shall always remember our endless conversations and your support during the "writing up" times—in venting mutual frustrations together!! Amma and Appa, you have been *more* than *ideal parents;* thank you for your selfless sacrifices; dedication; care; and affection..I don't know what I did to deserve all that. And one person for whom I have no words; for whom I *need* no words...Raja..to ***you*** all I have to say is.. but for you I would not be writing this and I owe the following pages to you.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

List of Tables

Chapter 1

Introduction

Enzymes play a significant role in the biotech industry with the market currently valued at over $ 1 billion. The food, pharmaceutical, detergent, pulp and paper, and specialty chemical industries use enzymes in several processes which often require enzymes to function under severe conditions such as extreme pH, temperature etc. This provides a strong motivation towards improving properties of existing enzymes so that they are well suited for the process. Protein engineering, which is a multi-disciplinary field, is a valuable tool towards achieving the above objective. This chapter describes the power of this technique citing several examples and applications of this technology. Most improvements in enzyme properties were achieved through modifications in the enzyme active site and substrate binding site. Hence the chapter ends with a description of the substrate binding and active sites and mechanisms of two well characterized enzymes—chymotrypsin [containing the classical catalytic triad (Ser; His; Asp) which is a common feature of several enzymes] and lysozyme, a well characterized polysaccharide degrading hydrolase.

1.1 Classical Catalyst Design and Engineering

Catalysts are substances that increase the rate of a chemical reaction without themselves being consumed or changed in the overall process. They have been used widely in the chemical engineering industry for a variety of products. A classical example is the automobile industry. Catalysts are used for refining crude oil into automotive fuels. Further, a wide variety of components used in automobiles are produced by different catalytic processes. These include polymers for structural components, paints and synthetic fibers in addition to various solvents, and other chemicals used in manufacturing processes. Finally, catalytic converters are used in automobiles for control of emissions due to fuel combustion (Hegedus, 1987).

There has been an increasing focus to improve properties such as selectivity, activity, life, and cost effectiveness of catalysts. With environmental constraints and need for improved fuel economy, the research efforts have aimed at improving catalysts for existing processes. Developing new or improved catalysts that could be tailored to achieve desired activities and selectivities became the goal of the catalyst industry.

One of the key steps to achieve the above goal is a thorough understanding of the factors that influence catalyst performance. This begins at the catalyst surface i. e. a better understanding of the relationships between structure and composition of the surface and reactivity is required. Modern surface science techniques such as solid state NMR, electron scattering, photon scattering, etc have been used effectively to elucidate molecular level information of catalytic systems (Duncan and Dybowski, 1981; Thomas and Klinowski, 1985). A molecular level understanding of the catalytic process coupled with reaction engineering tools can be used for optimal reactor design for the catalyst particles. In addition, once the correlation between kinetics (turnover rates) and selectivity (product profile) of a given catalyst system and its molecular structural properties are established, a new generation of catalysts can be readily developed (Somorjai, 1987).

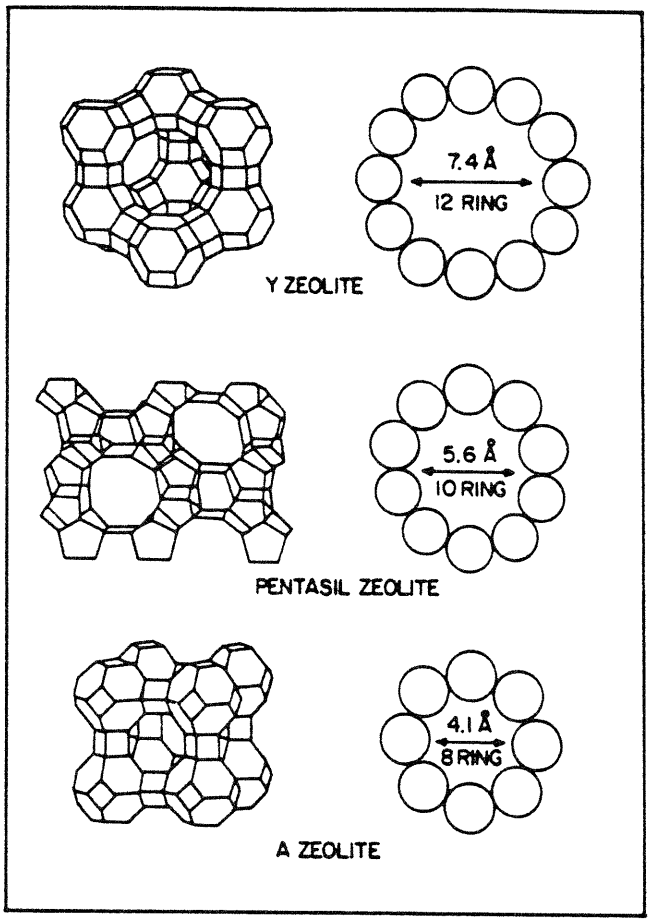
Good catalyst design involves building a catalyst that carries out a desired chemical reaction selectively and at high rates to maximize the conversion from reactants to

products. This has been achieved for several heterogeneous catalyst systems through an understanding of the atomic scale structure, composition and bonding properties. Zeolites are an example where molecular level structural information has been used effectively in designing catalysts with improved properties. Zeolites are aluminosilicates capable of promoting a large number of acid-catalyzed reactions, including polymerization, cracking, alkylation reactions, etc (Barrer, 1982; Breck, 1974; Haag and Chen, 1987). They are well suited for catalyst design since they have a well-defined pore structure and a large intracrystalline surface area (Barrer, 1982; Breck, 1974). Further, since the diameter of the pores is uniform, shape selectivity and molecular sieving became possible with zeolites (Csicsery, 1984).

Three main considerations for heterogeneous catalysis are activity, selectivity towards one or several products, and stability of the catalyst. These properties have been engineered successfully in a systematic way with zeolites (Haag and Chen, 1987). Molecular engineering of zeolites involves manipulating the crystal and pore shape and size; nature and number of active sites. Several methods are available to direct the synthesis towards a particular structure or crystal size. **Figure 1.1** shows the framework structure and pore openings of small, medium and large pore zeolites. Since they have a well defined structure with channels, it is possible to design shape selective catalysts based on either size or differences in diffusivities of participating molecules.

Specific acid sites located within crystalline cavities are responsible for the catalytic activity of zeolites. By changing the concentration of active sites, the catalytic activity of certain zeolites can be varied over wide ranges. The uniqueness of zeolites is that the activity can be varied over orders of magnitude in a predictive manner, independent of other properties such as catalyst size, surface area, etc. As a result this has, not only helped in gaining fundamental understanding of catalytic phenomena, but also has provided great flexibility in the design of novel catalysts (Chang et al., 1985; Dessau and Kerr, 1984; Shihabi et al., 1985).

Enzymes are biological catalysts that play important roles in regulating the chemistry of cells and organisms. The living cells employ enzymes for a variety of reactions which are highly controlled and regulated to modulate the production of different substances in response to cellular needs (Mathews, 1990).



[Haag and Chen., 1987.]

Figure 1.1: The framework and pore structure of small, medium and large pore zeolites. The well defined structure makes it possible to design shape selective catalysts based on either differences in diffusivities or size.

The bioprocess industry uses enzymes in various catalytic reactions. Typical examples include: food, pharmaceutical, detergent, textiles, wood pulp and paper industries. With the rapid growth of the industrial enzyme market (currently valued at over $1 billion, (1996)) there is an increasing need to develop novel enzymes and formulations. Further, several industrial processes require extreme conditions of pH, temperature, etc. Hence, it would be of great importance to be able to engineer enzymes with desired and improved properties such as enhanced activity; altered specificity (selectivity), improved stability, etc.

1.2 Enzyme Engineering in the Biotechnology Industry

As pointed out above, several industries use enzymes for various processes. For example, in the textile industry cellulases are used in fabric and garment processing and this has greatly helped the denim industry. The paper industry uses xylanases to bleach pulp and as a pretreatment to reduce the amount of bleaching agent. The advantages of enzymes in commercial processes over conventional chemical methods are their specificity and cost-benefits.

Designer enzymes can be made for specific applications, where currently available biocatalysts are inadequate. An available but unsuitable enzyme for a specific process can be modified, usually by making a minimal change in the catalytic active site of the enzyme (Wiseman, 1993). For example, as pointed out earlier, several industrial processes require conditions where it would be of great importance to be able to engineer existing enzymes with enhanced activity or stability. Recent examples include development of detergents with enzymes with increased thermal and oxidative stability at elevated pH (GEN, 1996).

Over recent years there has been a growing interest in biotech industries with the hope that new enzymes would be created. Enzymes have well-defined three-dimensional structures and bind their target ligands in specific conformations. Knowledge of three-dimensional structures of enzymes and defining its ligand binding surface would enable one to develop a rational approach towards engineering properties such as specificity, thermal stability, etc (Blundell, 1994). To enable this several interesting tools, such as screening of enzyme activities obtained from natural sources (selection), protein engineering, or chemical modification of

existing enzymes are available. A classical example of improvement through selection is the isolation of thermostable DNA polymerases from thermophilic bacteria such as *Thermococcus litoralis* (Belkin and Jannasch, 1985). The native organism is capable of growth upto 98°C and was isolated from a submarine thermal vent.

1.3 Protein Engineering—A powerful Tool to Alter Enzyme Properties

Protein engineering is a rapidly developing, highly multidisciplinary area and has several important implications for the modern biotechnology industry. This field involves developing rational approaches to the design of novel tailored molecules combining various techniques including gene cloning, site-directed mutagenesis, protein expression and structural and biochemical characterization. It is a technology that can be used to incorporate site specific changes at selected regions in proteins to improve properties such as stability, activity or potency, in order to adapt them for clinical and industrial applications (Blow et al., 1986; Blundell, 1994; Smith et al., 1988). While the methods to introduce mutations have become routine, the interpretation of the effects of the mutation on the structure and activity of a protein is less trivial (Rees, 1992).

A good example of applying protein engineering to improve an enzyme is the design of improved alkaline protease bacterial subtilisin used in enzyme washing powders, where stability to chemical oxidation was improved (Cunningham and Wells, 1987). As a result the enzyme in the detergent had enhanced stability at higher pH resulting in improved efficacy. Another important application that would have significant advantage is tailoring the pH dependence of enzyme catalysis to optimize activity in industrial processes. This was successfully accomplished using protein engineering, by varying the surface charge, which modified the pKa of the ionic catalytic groups (Thomas et al., 1985). A single charge alteration in the vicinity of the active site of subtilisin had a significant effect on the pH dependence of the enzyme. Fusion proteins that recognize specific ligands have been widely used in affinity chromatography to aid purification of recombinant proteins (Blundell, 1994). The specificity of enzymes has been varied through insertions and deletions (El Hawrani et al., 1994). In the pharmaceutical industry, tissue-type plasminogen activator enzyme (t-PA), which is used for improved

thrombolytic therapy, has been redesigned and improved using protein engineering techniques (Robinson and Browne, 1991).

However, engineering superoxide dismutase perhaps represents the best example where knowledge of the active site environment aided in designing a mutant enzyme with considerable enhancement in activity (Getzoff et al., 1992). This enzyme catalyzes the rapid detoxification of the harmful superoxide anion (O_2^-) which may be generated in the body during metabolism of some environmental chemicals (Halliwell and Gutteridge, 1989). The three-dimensional crystal structure revealed three amino acids (glutamic acid-132, glutamic acid-133 and lysine-136) that form a hydrogen bonded electrostatic environment for the negatively charged superoxide anion. Increasing the net positive charge in the active site using protein engineering (by converting glutamates to glutamines) increased the overall reaction velocity three-five fold. Such improved enzymes, although not designed for that purpose, could perhaps be used to identify superoxide releasing chemicals, before releasing them into the environment (Wiseman, 1993).

1.3.1 Application of protein engineering in enzyme immobilization

Immobilizing enzymes onto supports can have a significant impact on biotechnology and enzyme engineering with the development of bioreactors. The biggest advantage of using immobilized enzymes over free enzymes is their ease of recoverability and reusability (Messing, 1975). As a result, membrane based bioreactors can significantly lower the costs associated with the engineering processes by influencing the rate and extent of chemical reactions (Zhuang and Butterfield, 1992). During the past several years, synthetic membranes containing active enzymes have been studied leading to the development of various supports and techniques for immobilization (Imai et al., 1986; Kolisis and Thomas, 1987; Yokoyama et al., 1988). For supports, polymeric materials have been used extensively since they contain various functional groups which can be modified for covalent attachment, and for enzyme immobilization a number of coupling chemistries have been studied and applied (Scouten, 1987).

Protein engineering has played an important role in enabling effective enzyme immobilization onto supports. For example, the SH group of a cysteine is very

reactive and introducing a cysteine residue within the enzyme, without affecting the enzyme activity, would be a useful approach to immobilizing the enzyme (Stayton et al., 1992). This was successfully demonstrated in the case of dihydrofolate reductase (Iwakura and Kokubu, 1993). Using thiopropyl-Sepharose as a support, the engineered enzyme was effectively immobilized with retention of enzyme activity.

An interesting application of protein engineering in enzyme immobilization has been to engineer metal coordinating sites into the enzyme and using immobilized metal ion affinity chromatography (IMAC) for enzyme immobilization (Chaga et al., 1994). In the case of rat glutathione transferase, it was shown that two adjacent histidine residues on the surface, distant from the active site, were responsible for binding to Ni(II) affinity matrix. A homologous enzyme from the same family lacked one of the two histidine residues and as a result did not bind to the Ni-affinity column. Using protein engineering, a histidine residue was introduced at the position, adjacent to the existing histidine residue and this resulted in 75% of the mutant protein being bound to the column. Thus, this approach of introducing metal coordinating sites has significant potential in immobilization of proteins and other macromolecules, which in turn has applications in medicine and bioengineering.

A thorough understanding of the structure-function relationship of the protein to be immobilized can help in designing immobilization schemes, as was seen with papain (Zhuang and Butterfield, 1992). Papain is a proteolytic enzyme whose active site structure is well characterized (Brocklehurst et al., 1987; Kamphius et al., 1984). It was immobilized onto a support and detailed structural studies were carried out to examine the conformational change and the active site structure of the immobilized enzyme. It was found that properties of the bound enzyme were related to the structure and could be explained by the active site conformation of papain. Thus structural information and knowledge of active site conformation could aid in better immobilized enzyme system design with optimal properties such as different coupling chemistries (Zhuang and Butterfield, 1992).

1.3.2 *Application of protein engineering —for non-natural environments*

Most proteins can function in a certain environment and have not been optimized for *in vitro* applications under a different set of conditions. For example, an enzyme used in a biocatalytic process may have to withstand high temperatures for a considerable period of time to achieve optimal productivity. The ability to use proteins in nonnatural environments could greatly enhance their potential applications in biotechnology (Arnold, 1993). Further, modifying surface or ligand binding properties could simplify protein production and purification (Smith et al., 1988). Thus there is considerable scope for improvement in properties and altering the amino acid sequence to enhance or modify protein performance under extreme conditions of pH, salt, temperature etc is a useful approach (Bryan et al., 1986; Cunningham and Wells, 1987; Kong et al., 1993).

Two approaches using protein engineering can be used to achieve the above function. The rational design approach requires high level of structural information to effectively incorporate changes in the protein sequence. On the other hand, the random mutagenesis approach can be used when no information is available but this strategy requires efficient screening techniques for the property of interest (Arnold, 1993).

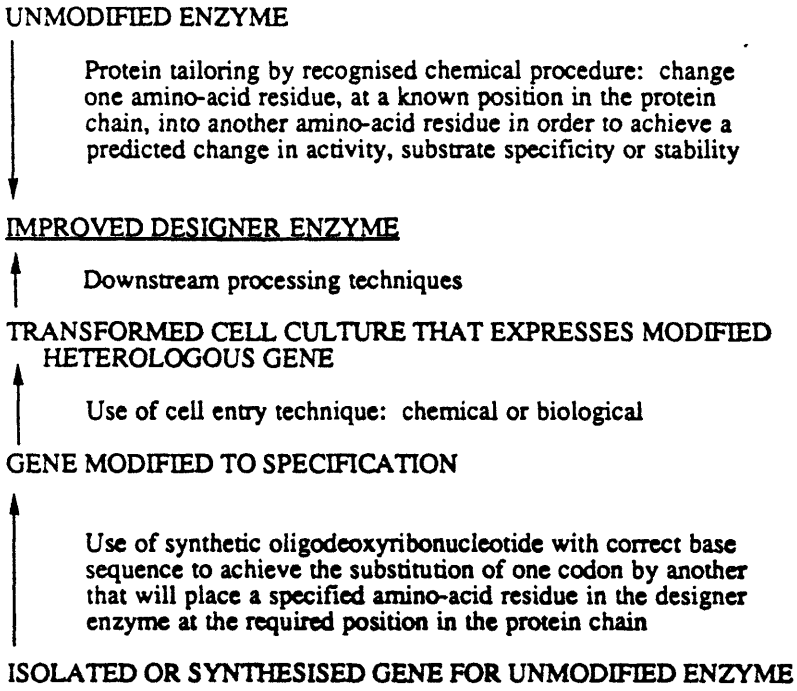
Using a rational design approach, the thermal and conformational stability of proteins have been improved. The strategy involves introducing metal chelating sites at specific positions on the protein which would bind to metal ions such as Cu^{+2} or Zn^{+2} and thereby stabilize the protein since the ions would bind preferentially to the folded state of the protein. Such an approach has been successfully applied for the stabilization of peptides (Ghadiri and Choi, 1990) and proteins (Pantoliano et al., 1988; Handel and DeGrado, 1990; Kellis et al., 1991; Umana et al., 1993).

However, properties such as catalytic activity and substrate specificity are more difficult to engineer using the above approach since they require detailed structural information and random mutagenesis could be used for improving such properties. The random mutagenesis approach has been successfully applied to engineer proteins with enhanced thermal stability (Bryan et al., 1986; Liao et al., 1986);

enhanced alkaline stability (Cunningham and Wells, 1987); and altered substrate specificity (Oliphant and Struhl, 1989). With this approach, sequential cycles of mutagenesis and screening can be carried out and progressively harsher conditions of screening can be applied at each stage.

Another significant application has been to enhance the catalytic activity of enzymes in polar organic solvents (Chen and Arnold, 1993). A significant observation from this study was that mutations that led to improved catalytic efficiency were limited to a specific region of the enzyme surrounding the active site and substrate binding pocket. As a result this random mutagenesis approach also provided suitable targets for further site-directed mutagenesis experiments. These examples clearly demonstrate the importance of understanding the structure and properties of the active site and substrate binding pockets of enzymes in improving their properties.

In summary, design of modified zeolites involve manipulation of crystal and pore shapes and sizes; nature of channel systems; and nature and number of active sites. Rational design of zeolite catalysts requires an awareness and preferably a quantitative knowledge of the influence of sorption, of reactant on the catalyst surface, on activity and selectivity. Another approach to modify activity is to manipulate the active site to adapt the zeolite catalyst to a particular process. Enzymes have far more complex structures comprising secondary, tertiary, and sometimes quarternary structures. Analogous to approaches to improve zeolite properties, strategies for modifying enzymes would hence include manipulation of active sites or substrate binding (i. e. sorption) sites. It is important to understand the nature of the active site environment and substrate binding pockets in enzymes and their effect on enzyme activity and specificity (selectivity). Further, understanding the enzyme's structure, surface properties and mechanism enables one to apply these tools to make modifications on a rational basis, as shown in **Figure 1.2**. Here again, protein engineering has been effectively used to elucidate structural and mechanistic information of several enzymes. The next section briefly discusses two significant examples.



[Wiseman, A., J. Chem Tech. Biotechnol, 56, 3, 1993.]

Figure 1.2: Approaches to protein engineering or protein tailoring to produce a designer enzyme for a particular application.

1.4 Structure-function and Mechanistic Studies of Enzymes

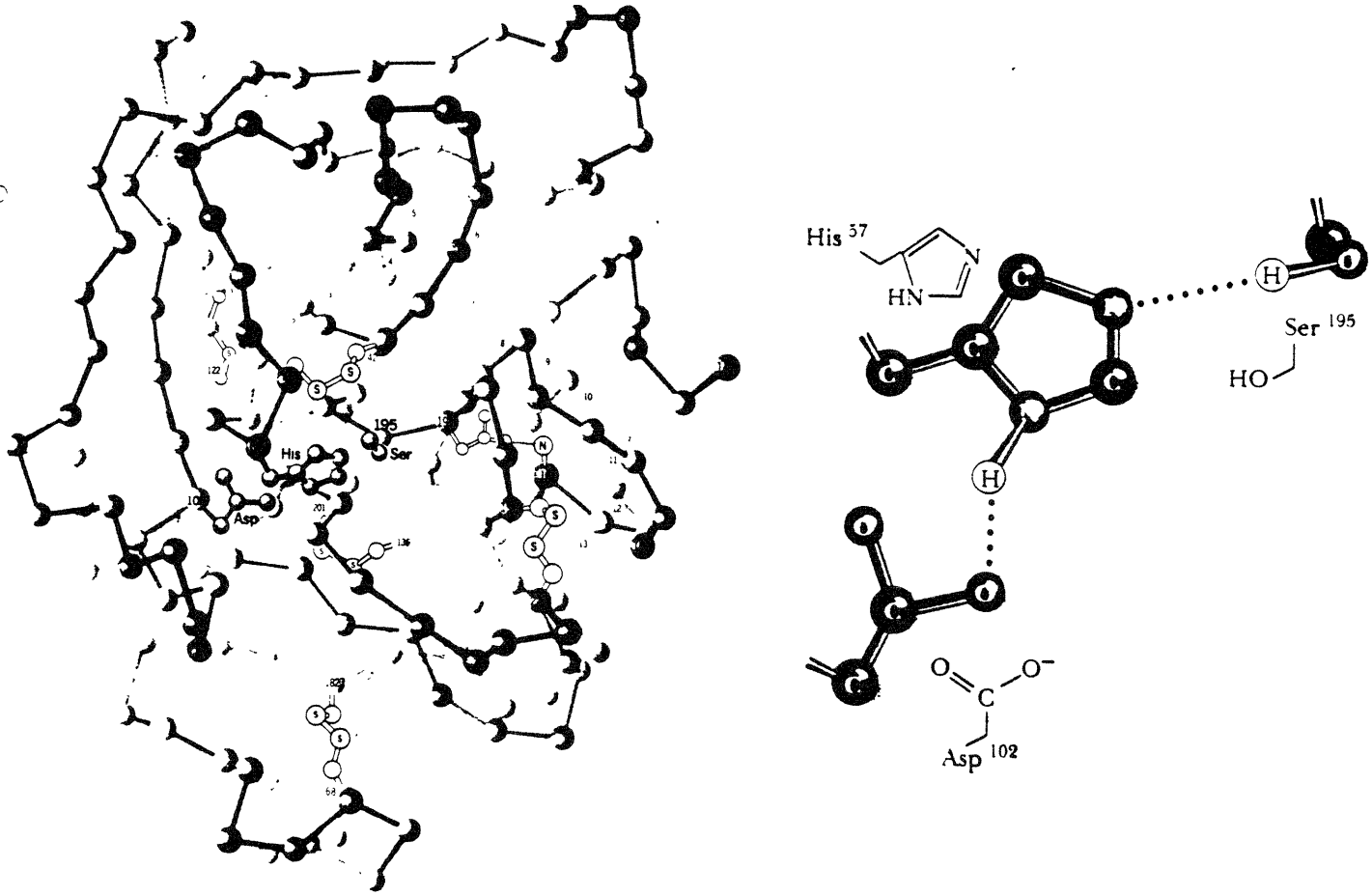
Enzymes have been formally classified by the International Commission on Enzymes, according to the reaction they catalyze, into six classes (Garrett and Grisham, 1995) . Each of these classes is further divided into subclasses. Hydrolases and lyases are two important classes of enzymes—hydrolases catalyze reactions which are accompanied by removal of a water molecule and lyases catalyze removal of a group from, or addition of a group to, a double bond. Hydrolases have been thoroughly studied and are well understood as a result of detailed mechanistic and kinetic studies. The three-dimensional crystal structures of hydrolases such as certain serine proteases (e. g. trypsin, chymotrypsin, etc.), and a polysaccharide degrading hydrolase, lysozyme, are known. The active site, substrate binding pocket and the catalytic mechanism of these enzymes have been studied extensively.

1.4.1 The serine proteases—catalytic mechanism of chymotrypsin

Serine proteases were among the first enzymes on which extensive structural and mechanistic studies were conducted (Neurath, 1985). Trypsin, chymotrypsin, subtilisin, thrombin, etc are some examples of members of this family. The three-dimensional crystal structures reveal that these molecules share a lot of structural and functional similarities (Blow, 1976; Kraut, 1977). The active site cleft contains a serine, histidine, and an aspartic acid which form a catalytic triad **(Figure 1.3)**.

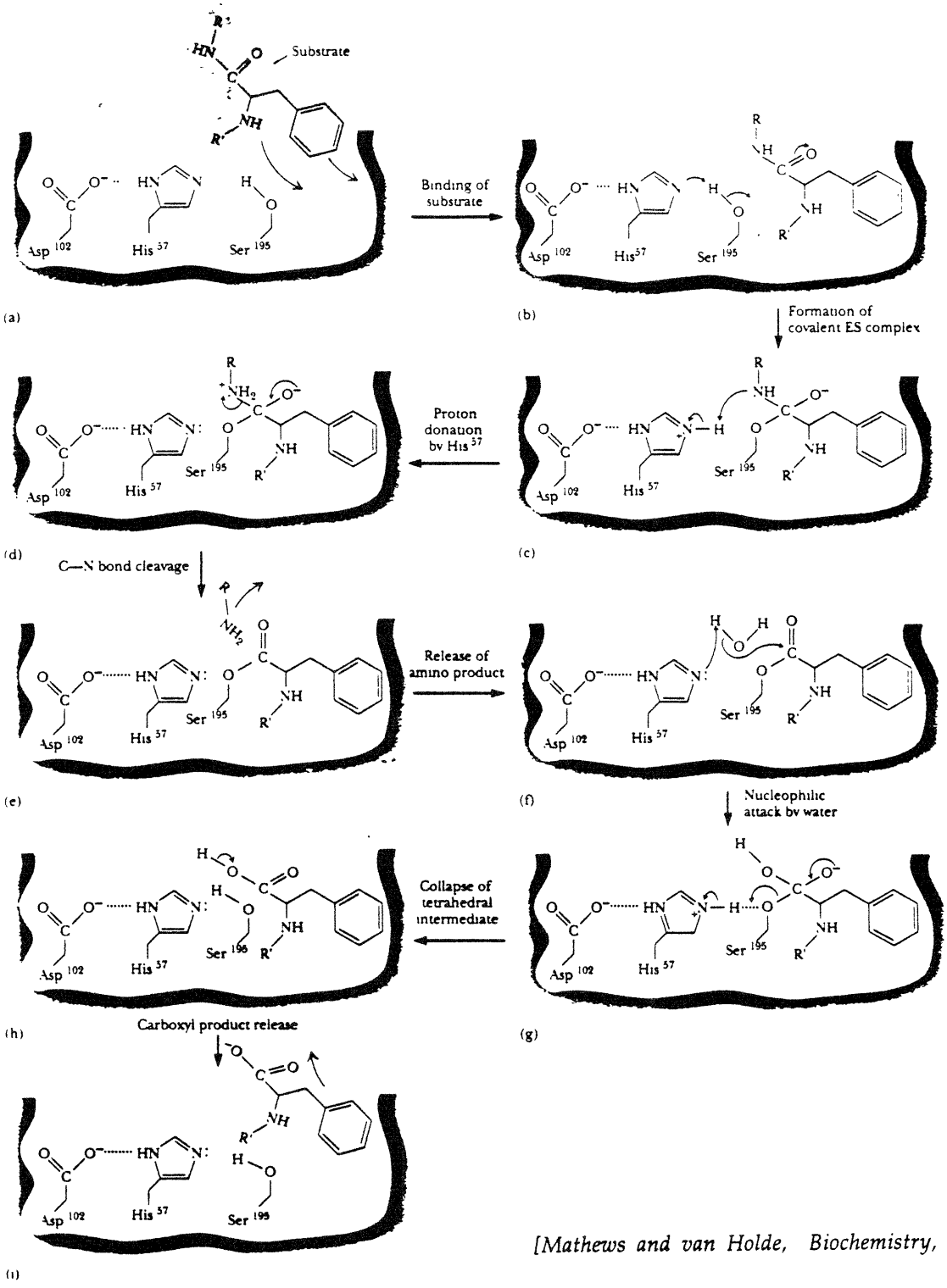
A detailed catalytic mechanism for peptide hydrolysis by chymotrypsin is shown in **Figure 1.4**. In the first step of the reaction, histidine acts as a general base by withdrawing a proton from the serine residue, thus facilitating a nucleophilic attack by serine on the carbonyl carbon of the peptide bond to be cleaved. This is followed by histidine donating a proton to the peptide C-terminal fragment, creating a protonated amine in the tetrahedral intermediate. This results in the cleavage of the CN bond and dissociation of the C-terminal peptide product.

In the next step, a water molecule performs a nucleophilic attack on the carbonyl carbon and generates another tetrahedral intermediate. The histidine in this step acts as a general base and accepts the proton from the water molecule. The intermediate subsequently collapses with the donation of a proton by the histidine



[Mathews and van Holde, Biochemistry, 1990.]

Figure 1.3: The structure of chymotrypsin. The active site contains the catalytic triad comprising Ser-195; Asp-102 and His-57.



[Mathews and van Holde, Biochemistry, 1990.]

Figure 1.4: Steps in the cleavage of a polypeptide by chymotrypsin. The shaded area represents the enzyme.

to the serine residue. The carboxyl group of the peptide gets deprotonated and is released from the enzyme to complete the reaction. The role of the aspartic acid of the triad remained unclear until recently, through site-directed mutagenesis studies (Craik et al., 1987) it was shown that this residue is essential for catalysis by stabilizing the histidine residue.

Over the recent years these enzymes have been the subjects for extensive mutational analysis aimed at understanding the substrate specificity (reviewed in (Perona and Craik, 1995)). For example trypsin and chymotrypsin have similar tertiary structures but different substrate specificities. Using site-directed mutagenesis it was shown that when two surface loops in trypsin are changed to analogous loops in chymotrypsin in conjunction with the substrate binding site, the substrate specificity of chymotrypsin is achieved (Hedstrom et al., 1992). Later, it was shown that arginine-172 in trypsin is an additional specificity determinant and changing that to a tryptophan (the corresponding residue in chymotrypsin) enhances the activity of the mutant enzyme containing the surface loops of chymotrypsin (Hedstrom et al., 1994).

In summary, the structural information has been used to design mutations in serine proteases altering substrate specificities, which in turn sheds more light on the enzyme action and mechanism. The charge relay mechanism involving a serine, histidine, and an aspartic acid catalytic triad has been recently identified in other enzymes as well, such as members of the lipase gene family including pancreatic lipase, lipoprotein lipase, and hepatic lipase (Brady et al., 1990; Winkler et al., 1990; Mooser, 1992).

1.4.2 The polysaccharide degrading hydrolase lysozyme-the first enzyme crystal structure solved

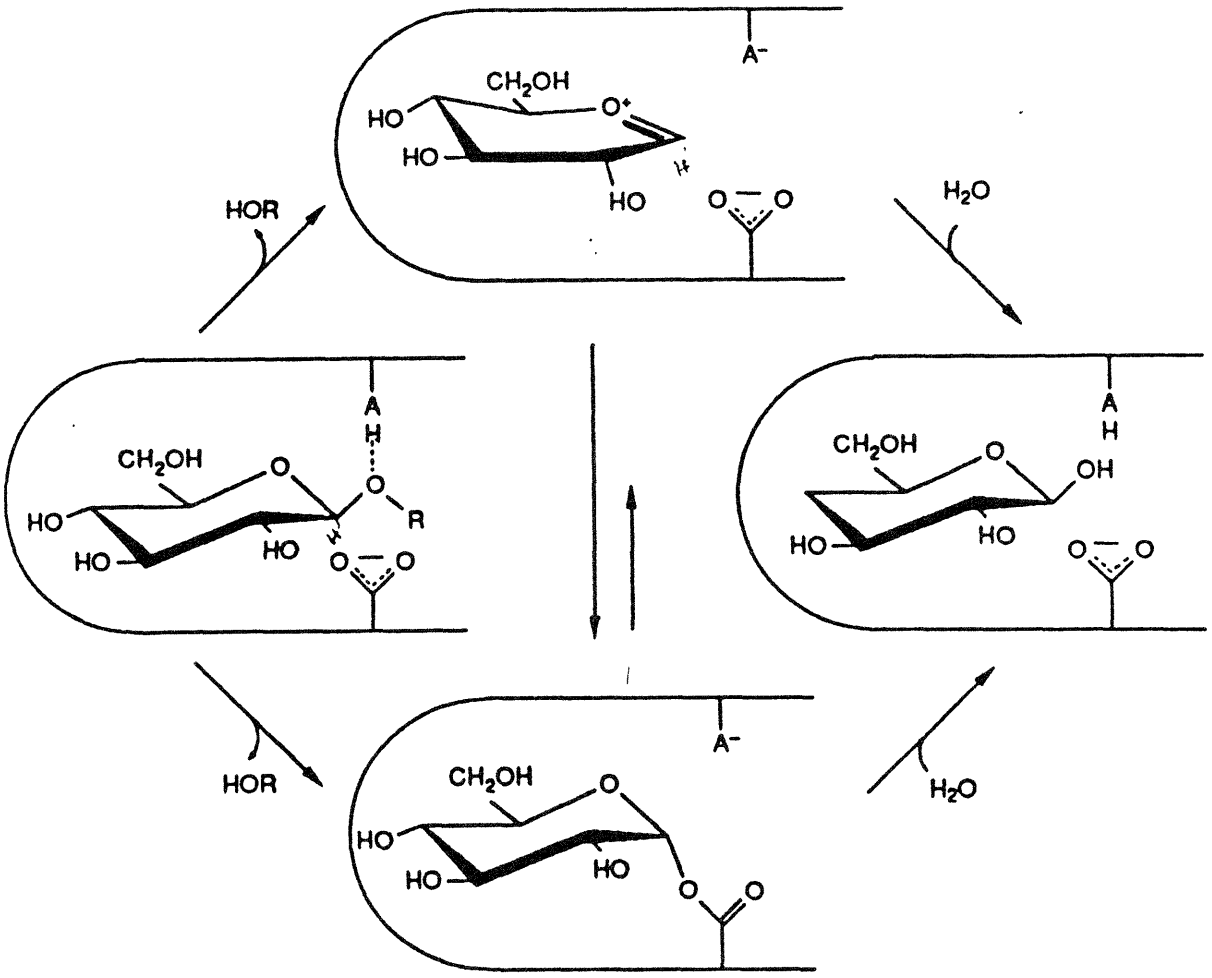
A large number of glycosidases (primarily catalyzing substrate hydrolysis) and glycosyltransferases (primarily catalyzing transfer to carbohydrate acceptors) have been reported but only a few have been studied in great detail (Mooser, 1992). Detailed studies on lysozyme, which is a glycosidase, have provided a lot of information on catalytic mechanisms of polysaccharide degrading hydrolases in general (Kirby, 1987). The enzyme reactions usually proceed through a discrete

glycosyl-enzyme intermediate. Two significant amino acid groups are generally involved: a carboxylate that stabilizes the glycosyl-enzyme intermediate and an amino acid that serves as a general acid to protonate the gylcoside oxygen.

Figure 1.5 shows a general mechanism for a hypothetical β-glucosidase that cleaves a β-glucoside bond liberating glucose. The carboxylate anion can participate in two ways—by electrostatic stabilization of the oxocarbonium ion forming (upper pathway) or by covalent nucleophilic attack of the reaction center. Both mechanisms are relevant and both transition states could exist in equilibrium (Mooser, 1992). The involvement of carboxylic acid residues in catalysis also has been shown in the recently solved crystal structure of the *Bacillus macerans* endo-1,3-1,4-β-glucanase (Hahn et al., 1990, 1995), where two glutamic acid residues (glutamic acids-103 and 107) are responsible for the cleavage of the β-1,4 glycosidic bond within the substrate molecule. However, the most well characterized polysaccharide degrading glycosidase is lysozyme.

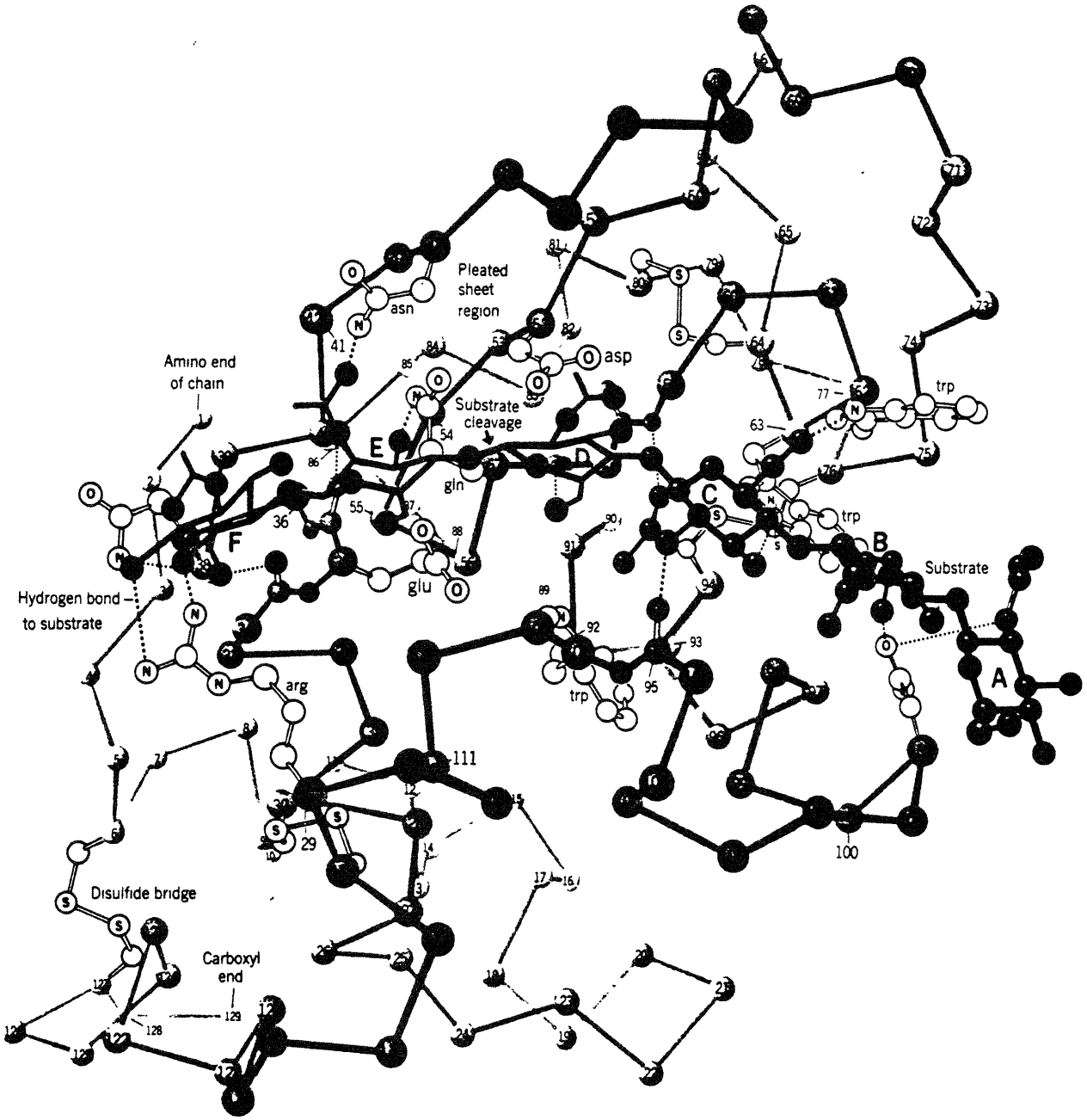
Lysozymes are N-acetylmuramidases that cleave the N-acetylmuramic-β-1,4-N-acetylglucosamine (MurNAc-GlcNac) linkages common to polysaccharides found in gram-positive bacterial cell walls. The three-dimensional structure of hen egg-white lysozyme was reported in mid-1960s (Blake et al., 1965; Johnson and Phillips, 1965; Phillips, 1966). Although X-ray crystal structures of proteins such as myoglobin and hemoglobin had been reported, this was the first enzyme structure to be solved **(Figure 1.6)**. Using substrate analogs and characterizing the bound complexes by X-ray crystallography the active site location was determined.

The primary structure of the hen egg-white lysozyme is composed of 129 amino acids that fold into a form with a deep groove separating two major protein lobes. The groove houses the active site, which is composed of six saccharide binding subsites (A-F) **(Figures 1.6 & 1.7)**. As shown in the figure, oligosaccharides bind with the nonreducing end extending towards subsite A and the scission occurring at the glycosidic bond between subsites D and E. Systematic model building and crystallographic studies on the enzyme complex with substrate analogs (tri and tetra saccharides) have provided valuable information on the interactions between the substrate and enzyme subsites A-C (Johnson et al., 1988); interaction at subsite D and at subsites E and F (Sarma and Bott, 1977). Extensive reviews on these structure-



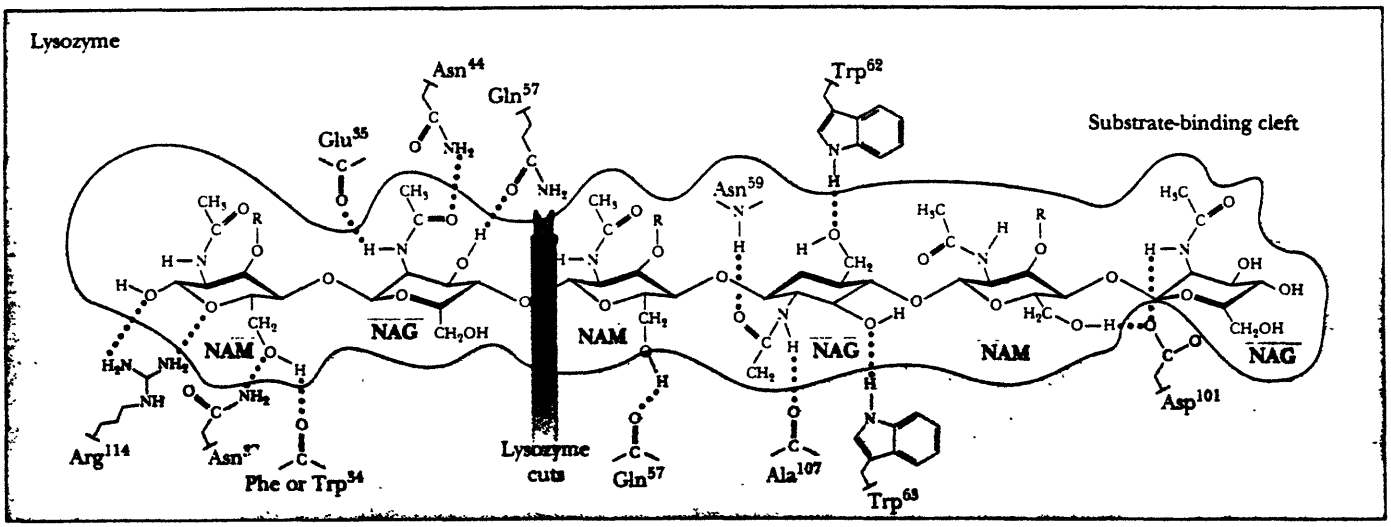
[Mooser, G., The Enzymes, 6, 1992.]

Figure 1.5: A general mechanism for a hypothetical β-glucosidase that cleaves a β-glucoside bond liberating glucose. The upper pathway passes through a oxocarbonium ion transition state and the lower pathway passes through a covalent α-glucosyl intermediate.



[Garrett and Grisham, Biochemistry, 1995.]

Figure 1.6: The crystal structure of lysozyme.



[Garrett and Grisham, Biochemistry, 1995.]

Figure 1.7: The substrate binding cleft of lysozyme.

function studies emphasize a structural perspective (Johnson et al., 1988) and a mechanism perspective (Kirby, 1987).

The catalytic mechanism in the case of hen egg-white lysozyme has been shown to proceed via a non covalent oxocarbonium transition state (Dalhquist et al., 1968; Rosenberg and Kirsch, 1981). Aspartic acid-52 is involved in the electrostatic stabilization of the developing carbonium ion and glutamic acid-35 acts as the general acid donating a proton to the glycoside oxygen. The significance of these residues has been clearly established from several different experimental directions. X-ray structure shows glutamic acid-35 and aspartic acid-52 in close proximity to the scissile bond. The surrounding environment of the two carboxylic acids differ: glutamic acid-35 is in a hydrophobic pocket, which elevates the pKa to ~6 and preserves the proton at catalytic pH until transfer to the glycoside oxygen; aspartic acid-52 being in a polar hydrophilic environment, maintains its normal pka of ~ 4.5 and is ionized at catalytic pH, consistent with its role of electrostatic stabilization (Kuramitsu et al., 1977; Parsons and Raftery, 1972). The significance of these amino acids is further supported by chemical modification and site-directed mutagenesis studies (Kuroki et al., 1986; Malcolm et al., 1990).

The catalytic mechanism proceeds as follows. Glutamic acid-35 acts as a general acid and protonates the oxygen atom of the glycosidic bond, thus forming a oxocarbonium ion transition state, which is stabilized by aspartic acid-52. Following bond cleavage, the product formed at the E site diffuses away and the carbonium ion intermediate can react with H_2O from the solution. Glutamic acid-35 acts as a general base now, accepting a proton from the attacking water molecule and the tetramer formed at sites A-D can now diffuse away and the enzyme is ready for the next round of catalysis.

An important feature of catalysis is substrate binding. In the case of lysozyme, the saccharide binding at subsites A-F involves both hydrogen bonding and nonpolar force interactions. The weakest interaction is at subsite D where the transition state develops, with the highest affinity at either sides of subsite D (subsites C and E) and becomes progressively weaker toward peripheral subsites A and F (Chipman and Sharon, 1969). The highest affinity subsite C, has an extensive hydrogen bonded network between the saccharide and the protein (Imoto et al., 1972). In addition to

the actual residues in the active site of the protein, adjacent to the scissile bond, that lead to substrate cleavage, other residues of the substrate binding pocket as well contribute indirectly to catalysis through an intricate network of hydrogen bonding with the substrate molecule at the active site. Electrostatic interactions may also play a role towards stabilizing either the active site or the transition state of the enzyme-substrate complex. The electrostatic interaction between the enzyme and the substrate becomes especially important when the substrate is a macromolecular polyelectrolyte (Muraki et al., 1988). This was evident when, varying the surface charge of lysozyme using protein engineering, altered the catalytic activity of the enzyme and the optimal pH and ionic strength for activity.

In conclusion, it is clear that hydrolases have been extensively studied and a lot of information is available through various crystallography, site-directed mutagenesis and other kinetic and mechanistic studies. In the case of the serine proteases such as trypsin, chymotrypsin, and α-lytic protease, this knowledge has been effectively used to design mutant enzymes with altered specificities using protein engineering techniques (reviewed in (Perona and Craik, 1995)). The polysaccharide degrading hydrolase, lysozyme, has also been thoroughly studied and protein engineering of lysozyme, as mentioned above, has resulted in enzymes with altered pH and ionic strength dependence.

There is another class of important molecules known as the polysaccharide degrading lyases (Linhardt et al., 1986) which have been studied with interest in the recent years and where there is not as much structure-function information available, as for example with lysozyme. Heparinase, which is the subject of this thesis, is a member of this enzyme class and the next chapter will discuss the salient features of these enzymes, focusing on any structural information available.

This chapter begins with a description of complex polysaccharides and GAGs. Heparin like molecules are acidic polysaccharides which play a significant role in modulating several important functions in the body. Heparin has a variety of biological activities and applications. Further, heparin interacts with several proteins through specific sequences on the protein as well as in heparin. Heparinases are valuable tools to study heparin-protein interactions since they cleave heparin with unique specificity and are a model system for this study to investigate GAG-protein interactions. The chapter ends with a discussion on recent background work relevant to heparinase I, which forms a theme for this work.

Chapter 2

Background— Polysaccharides and Polysaccharide Degrading Enzymes

Part 1 Polysaccharides

2.1 Proteoglycans and GAGs

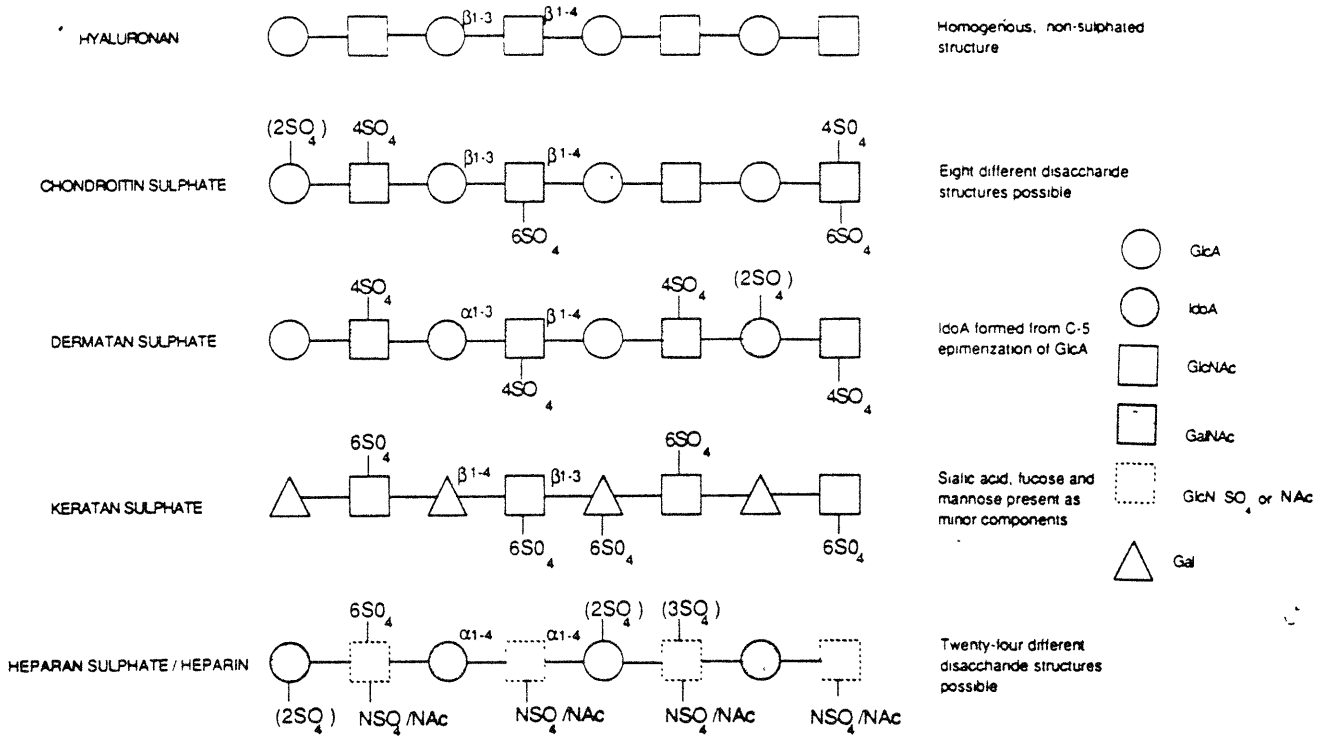
The extracellular matrix (ECM) (**Figure 2.1**) is made up of three components—the insoluble structural components such as collagen, laminin, and fibronectin; the soluble components called proteoglycans and the transient components such as growth factors. Proteoglycans are proteins that have one or more attached glycosaminoglycan (GAG) chains and are produced by most eukaryotic cells (Hardingham and Fosang, 1992). They interact with both insoluble and transient components and regulate various physiological processes such as cell adhesion and migration, in addition to providing hydration and architecture to the ECM.

GAG chains are large extended polysaccharide structures with highly charged sulfate and carboxylate groups and they dominate the physical properties of the protein to which they are attached. They are linear heteropolysaccharides usually comprising a typical characteristic disaccharide repeat unit of a hexosamine and uronic acid. The most common GAG structures are chondroitin sulfate, dermatan sulfate, heparan sulfate, keratan sulfate and hyaluronic acid **(Figure 2.2)**. These GAGs interact with an array of cytokines that regulate cell growth and differentiation such as fibroblast growth factor (Rapraeger et al., 1991) and vascular endothelial growth factor (Gitay-Goren et al., 1992) and play a central role in the ECM through modulation of signal transduction (Jackson et al., 1991; Kjellèn and Lindahl, 1991; Scott et al., 1992)

GAG regulation has been associated with various diseases such as angiogenesis, atherosclerosis, tumor growth and Alzheimer's disease (Gallagher et al., 1986; Celesia, 1991). They have numerous clinical applications as well. Heparin, for example, has been used as an anti-coagulant in surgery for several years. Heparin like molecules (heparin and heparan sulfate proteoglycans) are known to play a key role in the ECM through their interaction with several molecules.

[Venkataraman et al., (1996), PNAS, 93, 845.]

Figure 2.5: The crystal structure of bFGF.



[Hardingham and Fosang, FASEB J., 6, 861, 1992.]

Figure 2.2: *Repeating disaccharide structure of glycosaminoglycans.* Heparin is a more extensively epimerized and sulfated form of heparan sulfate. Heparan sulfate frequently contains some chain segments with little or no epimerization or sulfation.

2.2 The Heparin like Molecules—Heparin and Heparan Sulfate Proteoglycans

2.2.1 Heparin—biological activities and medical applications

Heparin is a highly charged polyanion under physiological pH. Within the past few years, several biological activities have been shown to be regulated by heparin. This led to an increasing necessity to understand heparin structure and elucidate its structure-activity relationship (Linhardt and Loganathan, 1990). Some significant medical applications and biological activities of heparin are mentioned below.

2.2.1.1 Anticoagulant activity—systemic and regional heparinization

Heparin has been used primarily as an anticoagulant in the clinic for over 50 years. Heparin is administered intravenously (systemic heparinization) during most extracorporeal procedures (where blood is removed from the body and passed through a device) such as kidney dialysis and membrane oxygenation used in heart by-pass procedures (Langer et al., 1982). Systemic heparinization is also used in treating deep vein thrombosis and a variety of other surgical procedures. Hemorragic complications are a major side-effect of using heparin due to its anti-coagulant activity.

Regional heparinization is an approach to solve the problem of heparin's associated side-effects. In an extra corporeal circuit, heparin is added just as the blood leaves the body and enters the device. Heparin is removed from the system as it leaves the device and re-enters the body using immobilized heparinase, which cleaves heparin and thus destroys its anticoagulant property (Langer et al., 1982). This will be described later in this chapter, in detail.

2.2.1.2 Antithrombotic activity

The blood coagulation cascade comprises a series of steps involving serine protease coagulation enzymes, finally leading to the formation of a fibrous clot. Thrombin is a key enzyme in the coagulation pathway which cleaves fibrinogen to form fibrin monomers that form the clot. The entire pathway is regulated by serine protease inhibitors. One important inhibitor is antithrombin III (ATIII) which binds most of

the coagulation enzymes and inactivates them by forming a covalent complex. Heparin binds to ATIII and thrombin in a ternary complex and accelerates the rate of thrombin inhibition by ATIII by 2000 fold (Rosenberg and Damus, 1973).

Heparin also activates platelets, which are an important component of thrombosis (Linhardt and Loganathan, 1990). Once a clot is formed, heparin can accelerate the dissolution of the clot, by a process known as fibrinolysis, and prevent reformation of the clot.

2.2.1.3 *Antiatherosclerotic activity*

When heparin is administered intravenously, lipoprotein lipase (LPL) is released from the endothelium as a consequence. As a result there could be increased triglyceride lipolysis and hence lowering of cholesterol-rich particles in contact with the arterial wall. However, the problem of using heparin as an antiatherosclerotic agent is its associated anticoagulant property which results in hemorrhagic complications.

Smooth muscle cell proliferation following endothelial cell damage is a part of atherogenesis. Both anticoagulant and non-anticoagulant heparin are able to inhibit smooth muscle cell proliferation (Catellot et al., 1984; Catellot et al., 1986). Heparin oligosaccharides that lacked the ATIII binding site also showed anti proliferative abilities. Hence it is likely that anti coagulant and anti proliferative activities are separable.

Heparin is associated with a variety of biological activities and plays a significant role in the body. Heparin binds, activates, and inhibits a number of enzymes. For example heparin inhibits leukocyte elastase (Jordan et al., 1987) and hence could be used as a therapeutic in treating rheumatoid arthritis where this enzyme plays a role. Heparin can activate protein kinases found in both liver and skeletal muscle (Erdodi et al., 1984). Section 2.3 discusses heparin-protein interactions in detail.

2.2.2 *Heparan sulfate proteoglycans—role in the ECM*

Heparan sulfate proteoglycans (HSPGs) are found widely distributed on cell surfaces. They participate in diverse processes, including blood coagulation, anchoring of enzymes in the vascular lumen, cell adhesion and growth (Hoogewerf et al., 1995). Endothelial cell-associated HSPGs bind and potentiate ATIII, contributing to a nonthrombogenic vascular surface (Lindahl et al., 1986). Vascular HSPGs also bind and modulate the functions of many enzymes, including lipoprotein lipase, elastase, and superoxide dismutase (Cheng et al., 1981; Hjalmarsson et al., 1987; Redini et al., 1988).

HSPGs also interact with other ECM proteins such as laminin and collagen and modulate cell-matrix adhesion (Lattera et al., 1980; Hedman et al., 1982). In addition, HSPGs have been shown to play a key role in presentation of heparin binding growth factors (FGF, VEGF, etc.) to their receptors on various cells and thereby influencing a variety of cellular processes ranging from cell adhesion, migration, growth, proliferation, differentiation, etc. (Rapraeger et al., 1991; Yayon et al., 1991; Ornitz et al., 1992; Rapraeger, 1993).

2.2.3 *Heparin like molecules—therapeutic potential*

The use of heparin as a therapeutic is limited due its accompanying anticoagulant property. A new class of drugs, low molecular weight (LMW) heparins is being developed to treat such disorders (Linhardt and Loganathan, 1990). These agents include synthetic oligosaccharides and polysaccharides and those derived chemically and enzymatically from heparin. LMW heparins and heparin-oligosaccharides are being studied as potential antiatherosclerotic agents, to separate the undesired anticoagulant activity of whole heparin. The use of such heparin oligosaccharides as therapeutic agents depend on improved bioavailability, reduced side-effects and higher specificity. Understanding the heparin structure activity relationship would result in development of more applications for these molecules (Linhardt and Loganathan, 1990).

2.2.4 *Heparin and heparan sulfate—chemical structure*

Heparin is made up of alternating units of glucosamine and iduronic acid. The heterogeneity in the molecule is introduced because of the varying degrees of sulfation (Comper, 1981). The iduronic acid could be sulfated at the 2 position and the glucosamine can be sulfated at the N, 3, and 6 positions. The N position can also be acetylated. 32 different possible disaccharides can be generated **(Figure 2.3)**.

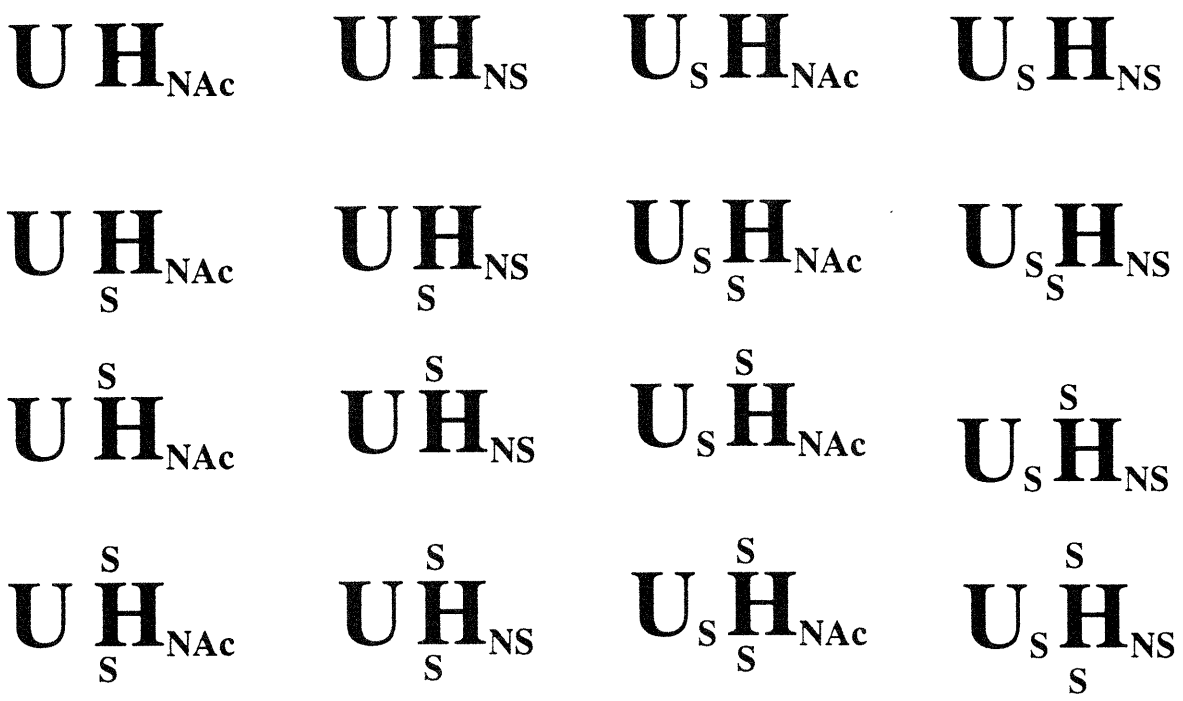
The structure of the major (85-90%) disaccharide repeating unit of heparin is shown in **Figure 2.4**. It consists of L-iduronic acid 2-sulfate which is α-1,4 linked to D-glucosamine 2,6 sulfate (IdoA-2S-Glc(N,6)SO_3). Other possible combinations of disaccharides are present in limited amounts. Commercial heparin has a molecular weight range of about 8,000-24,000 Da, the average molecular weight being 13,000 Da.

Heparan sulfate has several heparin like regions and has similar repeat units as heparin. The ratio of iduronic acid to glucuronic acid is higher for heparin than for heparan sulfate and this is the only significant difference between the two molecules. As a consequence, heparin is more negatively charged than heparan sulfate. Heparan sulfate is found widely distributed on the cell surfaces, while heparin is synthesized and secreted by mast cells.

2.3 Heparin-Protein Interactions

Being highly polyanionic, heparin binds various cationic proteins and controls certain physiological functions. Heparin has been used to purify proteins using heparin-sepharose chromatography (Shing et al., 1984). Over the past few years it has become increasingly evident that such interactions are governed by specific sequences on both the protein as well as heparin (Jackson et al., 1991; Lindahl et al., 1994).

On the protein side, Cardin and Weintraub identified two consensus sequences [XBBBXXBX or XBBXBX (B = basic residues; X = hydrophobic or other residues)] found in many heparin binding proteins (Cardin and Weintraub, 1989). Subsequently, site-directed mutagenesis and binding studies with synthetic or isolated peptides from several of these proteins have confirmed that this consensus



U = URONIC ACID

H= HEXOSAMINE

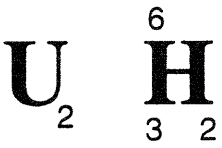


Figure 2.3: The different possible disaccharides of heparin based on variability in sulfation and acetylation.

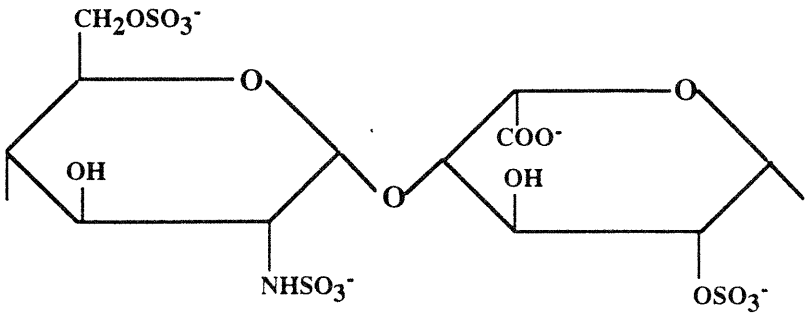


Figure 2.4: The structure of the major disaccharide repeating unit of heparin.

region often is involved in binding specifically to heparin (Baird et al., 1988; Bober Barkalow and Schwarzbauer, 1991; Thompson et al., 1994). On the polysaccharide side, oligosaccharides with certain sequences of modifications in the disaccharide repeat that confer binding specificity, have been isolated for AT III, bFGF, LPL and others (Lindahl et al., 1984; Maccarana et al., 1993; Parthasarathy et al., 1994).

Understanding the nature of heparin-protein interactions and specificity of such interactions is key to understanding various cellular events. This section reviews the current understanding of this interaction. These include: heparin binding proteins such as protease inhibitors (ATIII), growth factors (bFGF, aFGF); lipolytic enzymes (LPL); and extracellular matrix proteins (laminin, fibronectin).

2.3.1 Serine protease inhibitors

The role of the serine protease inhibitor, ATIII, in the blood coagulation cascade, was briefly discussed. Serine protease inhibitors (serpins) are a class of proteins which are a major component of blood plasma. They form covalent complexes with the active site of target proteases. Some serpins use GAGs as cofactors to enhance their activity as is seen with ATIII which uses heparin as a cofactor. Heparin binding to ATIII is thought to introduce a conformational change, which enhances ATIII binding to thrombin. This is a classical example of an important heparin-protein interaction, which has been studied in great detail, and sequences of both the protein as well as heparin that are involved in binding have been isolated.

2.3.1.1 The protein—ATIII sequence

Regions and residues of ATIII involved in heparin binding have been characterized using different approaches. The importance of basic amino acid residues in heparin binding was first demonstrated by Rosenberg and Damus (1973). Later, a heparin binding fragment from ATIII was isolated (Rosenfled and Danishefsky, 1986) using a CnBr digest. Smith and Knauer (1987) isolated a heparin binding fragment from protease digestion of ATIII. Based on these two studies, two regions of ATIII, a region between residues 123 and 145 and another in the vicinity of arginine-47, are involved in heparin binding. The role of arginines-47; -129 and -145 in heparin

binding has also been shown through labeling and mutation experiments (Sun and Chang, 1990).

2.3.1.2 *The heparin sequence*

The heparin end of the interaction has been extensively studied. A region on the heparin chain has been correlated to ATIII binding and shown to be responsible for the anticoagulant activity. This region is a pentasaccharide with the sequence $\mathbf{H}_{NS/Ac,\ 6S}$-**G**-$\mathbf{H}_{NS,(6S),3S}$-$\mathbf{I}_{2S}$-$\mathbf{H}_{NS,6S}$ (Lindahl et al., 1984; Atha et al., 1985; Yamada et al., 1993). However, this pentasaccharide needs to be present in a sequence atleast 12-16 saccharides long for anticoagulant activity (Marcum and Rosenberg, 1989a).

Commercial heparin chains contain about 4 to 7% of the pentasaccharide by weight and this corresponds to about one domain per 2-3 heparin chains (average molecular weight 14 kDa) (Linhardt et al., 1988; Yamada et al., 1993). Thus the ATIII binding pentasaccharide seems to be uniformly distributed within the chain. Recently subfractions of heparan sulfate, with a high content of the disaccharide **G**-$\mathbf{H}_{NS,6S,3S}$, have also been shown to possess anticoagulant activity, indicating that the binding sequence is highly specific and similar to that of heparin (Pejler et al., 1987; Marcum and Rosenberg, 1989b; Kojima et al., 1992).

2.3.2 *Growth Factors*

Growth factors are a class of molecules that regulate cell growth and differentiation. Various cellular processes are controlled by the molecular interaction of growth factors with their cell surface receptors and subsequent signal transduction mechanisms. This interaction is highly specific and regulated.

In the past few years, several peptide growth factors have been isolated using heparin-sepharose chromatography (Shing et al., 1984) and are termed as heparin binding growth factors. Acidic and basic fibroblast growth factors (bFGF) are the best characterized among this family of heparin binding growth factors. bFGF is found widely distributed in various tissues. bFGF interaction with the cell surface receptor is mediated through its binding to heparan sulfate proteoglycans on the cell surface.

This interaction has been studied in great detail and heparin binding domains on bFGF and the heparin sequence responsible for binding bFGF have been determined.

2.3.2.1 *The protein—bFGF sequences*

A region between amino acid residues 18-22 in bFGF was shown to be involved in heparin binding (Gospodarowicz et al., 1987). Later, using a synthetic peptide approach, Baird et al., (1988) determined the regions of bFGF involved in heparin binding to be between residues 24-68 and 106-120. The crystal structure of bFGF **(Figure 2.5)** shows that three regions fold appropriately to generate the heparin binding domain. The heparin binding sequences are basic clusters that conform to the Cardin and Weintraub consensus sequence described earlier. It is the three dimensional heparin binding domain, and not just the linear sequences alone, that is responsible for heparin binding.

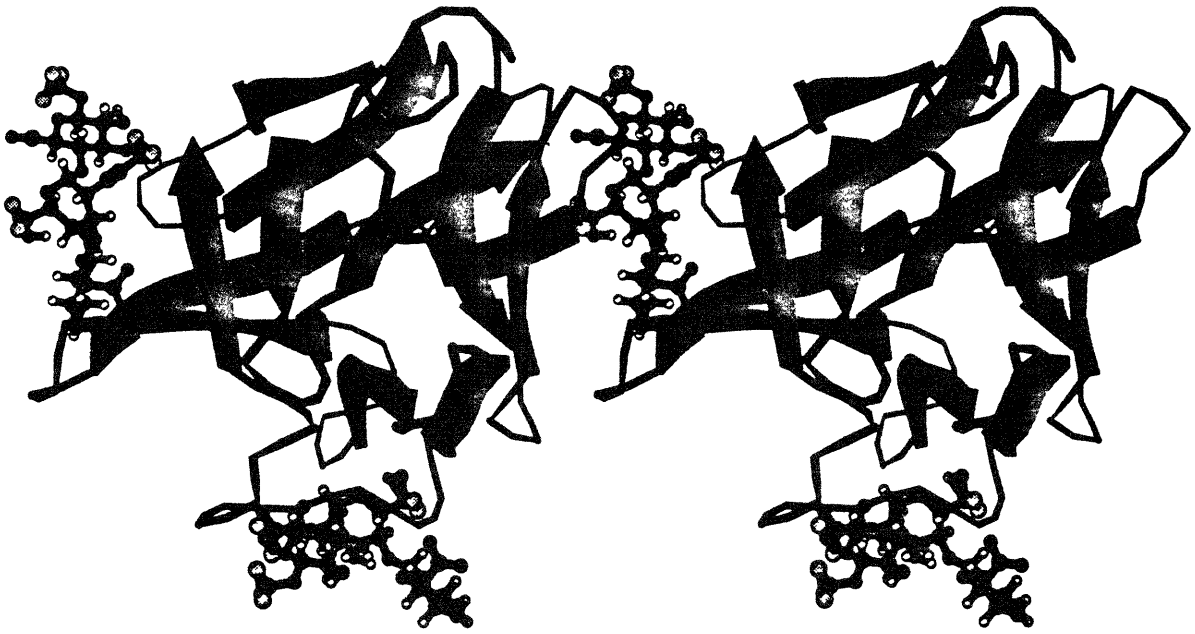
2.3.2.2 *The heparin sequence*

The minimal structure in heparin or heparan sulfate required for binding to bFGF has been determined (Maccarana et al., 1993). Using a combination of different oligosaccharides with varying degrees of sulfation, the minimal sequence was found to be a pentasaccharide of the sequence **U-H_{NS}-U-H_{NS}-I_{2S}**. While the 6-O sulfation of the hexosamine does not seem to affect binding, 2-O sulfation of the iduronate seems to be important for binding bFGF (Maccarana et al., 1993).

Although penta- and hexasaccharides are sufficient for binding bFGF, longer polysaccharide sequences are required to induce a mitogenic response in cultured cells. These additional sequences should contain 6-O and 2-O sulfation to promote bFGF mediated receptor signaling (Isihara et al., 1993; Guimond et al., 1993).

2.3.3 *Lipoprotein lipase*

Lipoprotein lipase (LPL) is a key enzyme in lipoprotein metabolism hydrolyzing triglycerides in plasma lipoproteins. This enzyme is synthesized and transported to the luminal surface of the vascular endothelium, where it binds to the HSPGs and regulates triglyceride hydrolysis (Parthasarathy et al., 1994). Addition of heparin



*[Venkataraman et al., (1996), PNAS, **93**, 845.]*

Figure 2.5: The crystal structure of bFGF.

releases the bound LPL from the HSPG and LPL binding is reduced by treatment with heparinases. These studies provide evidence of the role of HS chains in LPL binding to the cell surface HSPG.

The interaction of heparin sequences and LPL has been studied in some detail in the recent years. Specific basic clusters within LPL have been shown to be involved in heparin binding through detailed site-directed mutagenesis studies (Davis et al., 1992). In addition, the sequence of heparin has also been determined.

2.3.3.1 *The protein—LPL sequences*

Three positively charged clusters were identified in LPL which conform to the consensus sequence. The positive charges in all three clusters were altered systematically using site-directed mutagenesis and the mutant enzymes were subjected to affinity chromatography using heparin-sepharose to identify the structural determinants of heparin binding (Hata et al., 1993). Five basic residues in two of the three clusters were found to be responsible for heparin binding. The two clusters together form a pocket flanked by five basic residues for interaction with the heparin polyanion (Hata et al., 1993). Here again it is important to note that two of the three consensus segments combine to form a domain in the three dimensional structure indicating that it is the clustering of basic residues and not just a linear sequence that is important for heparin binding.

2.3.3.2 *The heparin sequence*

It was first reported (Bengtsson et al., 1980) that partial N-desulfated heparin was not able to displace LPL from immobilized heparin; however re-N-sulfation restored the ability to displace LPL. This suggested the importance of N-sulfation in commercial heparin for LPL binding but did not provide information on the nature of the interaction of LPL with oligosaccharide sequences in endothelial HS.

Using a hydrazinolysis/deamination procedure heparan sulfate chains which were rich in N-sulfated residues were generated (Parthasarathy et al., 1994). A decasaccharide was isolated that bound LPL with strong affinity. The decasaccharide consists of five repeating disaccharide units with the structure $\mathbf{I}_{2S}$-$\mathbf{H}_{NS(6S)}$. Binding

of the above mentioned positively charged clusters of LPL to cell surface HS might require an oligosaccharide conformation that greatly facilitates the interaction of negatively charged sulfate groups of the oligosaccharide. It should be pointed out that the LPL binding decasaccharide is different from the HS sequence involved in bFGF binding and does not contain the characteristic 3-O sulfated glucosamine, which is part of the ATIII binding sequence of heparin/HS.

Although GAGs contain several repeating units of hexosamine and uronic acid residues, the binding specificity to a number of important molecules seems to be confined to a small region comprising a few disaccharide units (Parthasarathy et al., 1994). The binding sequences comprise only a small fraction of the total oligosaccharides, indicating that although present in low proportions, the location of the binding sites on the cell surface might be important for LPL binding.

2.3.4 *Hepatocyte growth factor*

Hepatocyte growth factor (HGF) is an important molecule in embryogenesis, wound healing and tumor invasion. HGF displays a strong affinity for heparin. The precise identity of the sequence within HGF responsible for heparin binding is unknown. However, the HGF binding region within heparan sulfate has been characterized (Lyon et al., 1994). The sequence was determined to be of the type $\mathbf{G}\text{-}\mathbf{H}_{NS}\text{-}[\mathbf{I}\text{-}\mathbf{H}_{NS(\pm 6S)}]_{3\text{-}5}\text{-}\mathbf{I}\text{-}\mathbf{H}_{NAc}$. Two of the glucosamine residues are actually sulfated at the C6 position. The precise location of the $\mathbf{H}_{NS(6S)}$ residues within the sequence needs to be determined.

When comparing the above sequence to the bFGF binding sequence, it is interesting to note that in case of the bFGF sequence, both 2-O sulfation of uronic acid and N-sulfation of hexosamine are present and required for interaction. 6-O sulfation is absent in the sequence. Contrary to this, the HGF binding sequence has a high content of 6-O sulfation and relatively lower amounts of 2-O and N-sulfation (Lyon et al., 1994). The two growth factors which have different cellular specificities and biological activities interact strongly with heparan sulfate but have very different structural requirements for their interaction. These differences in HS recognition suggest that it is the variety and complexity of the HS structure that is able to sustain

the large number of interactions with various proteins in the ECM (Lyon et al., 1994).

2.3.5 *ECM proteins*

ECM is the neighborhood of every cell and plays a significant role in modulating cell adhesion, proliferation and migration. GAGs interact with the structural components of the ECM such as collagen, fibronectin, laminin, etc. The matrix HSPG binds fibronectin with high affinity. Based on the heparin binding consensus sequence, and other site-directed mutagenesis studies (Bober Barkalow and Schwarzbauer, 1991) two regions were determined that are responsible for heparin binding. However, the sequence of the heparin chain involved in binding is not yet known.

In conclusion, heparin interacts with a variety of proteins in a highly specific manner. Specific domains in both proteins as well as heparin influence this interaction. For proteins, it was observed that basic clusters forming heparin binding consensus sequences usually folded appropriately in the three-dimensional structure of the protein, thus forming a heparin binding domain. Although there are slight variations in the consensus sequence, this has been increasingly observed in a number of heparin binding proteins. Only a few heparin sequences involved in binding have been determined so far and they all are confined to small oligosaccharide sequences. This provides ample evidence to the fact that the nature of heparin-protein interactions is not simply non-specific electrostatic interactions but governed by very specific structural requirements on both the macromolecules.

Part 2 GAG Degrading Enzymes

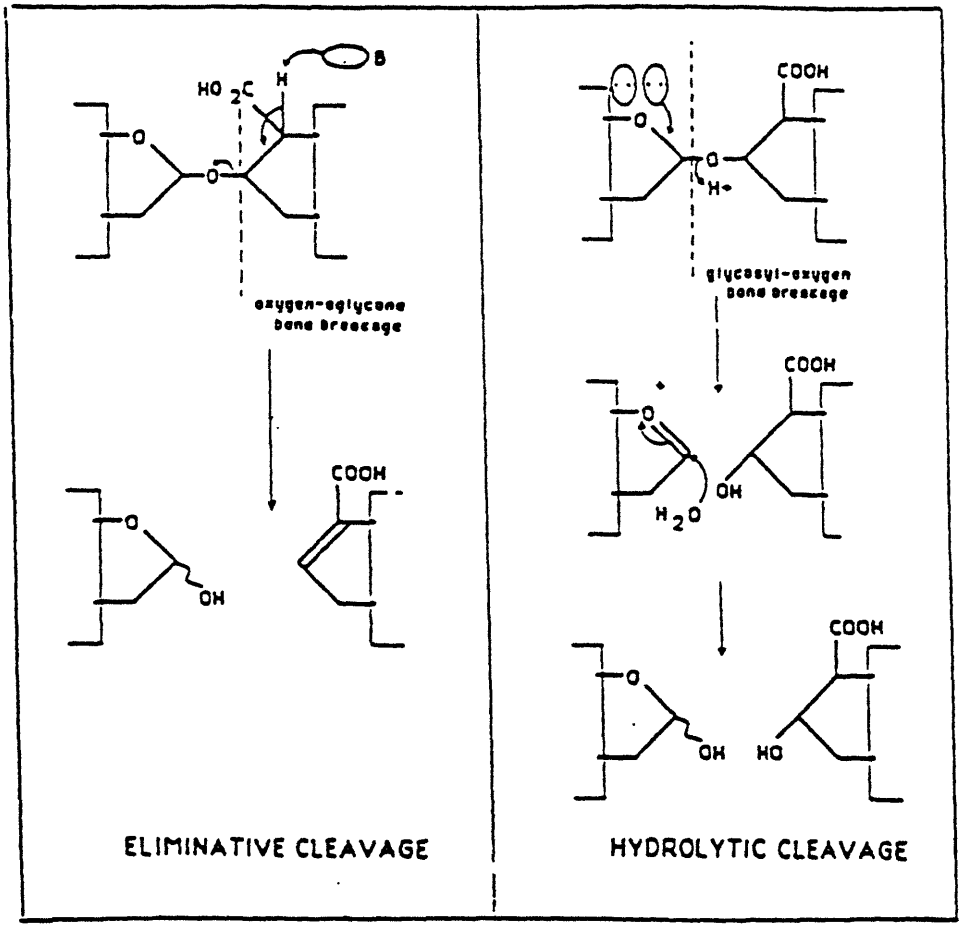
2.4 Polysaccharide Degrading Lyases

GAGs can be degraded either chemically or by enzymes. Enzymatic degradation of GAGs has several advantages (such as better yield, milder conditions and high specificity) over chemical methods of depolymerization. GAG degrading enzymes have proven useful in studying structure of GAGs, and mechanisms of GAG activity (Ernst et al., 1995). These enzymes, for example heparinases, are often specific for certain sequences in the GAG chain (Desai et al., 1993) and are hence interesting models to study GAG-protein interactions. Further, they also have shown potential for pharmaceutical and diagnostic applications (Sawaguchi et al., 199; Tejidor et al., 1993; Sasisekharan et al., 1994).

There are three categories of glycosaminoglycan (GAG) degrading enzymes of bacterial origin—heparinases, chondroitinases, and hyaluronidases; and alginases, pectinases, and xanthan lyases (reviewed Ernst et al., 1995). The general chemistry in a lyase mechanism is compared with that of a hydrolytic mechanism in **Figure 2.6**. In the hydrolytic cleavage mechanism a proton is donated to the glycosidic bond creating an oxocarbonium ion intermediate which is neutralized by a water molecule. In the elimination mechanism in lyases, the C5 hydrogen of the uronate is abstracted resulting in a C4-C5 unsaturation (Linhardt et al., 1986). Prior to work in our laboratory (Sasisekharan et al., 1995; Sasisekharan et al., 1996; chapters 4-6), there was no information regarding specific residues involved in the catalytic mechanism of these polysaccharide degrading lyases.

2.4.1 Chondroitinases

Chondroitin sulfate degrading lyases, chondroitinases, have been isolated from several different bacterial sources (reviewed, Ernst et al., 1995). Until recently, chondroitinase AC from *A. aurescens* was the only chondroitinase for which the amino acid composition had been determined (Takegawa et al., 1991). The enzyme activity was inhibited by sulfhydryl reactive agents such as PCMB, IAA and NEM, but enhanced by DTT and β-mercaptoethanol. This suggested that a reduced



[Linhardt et al., Appl. Biochem. Biotech., 12, 135, 1986]

Figure 2.6: General chemistry in a lyase mechanism compared to a hydrolytic cleavage mechanism.

cysteine could possibly be important for enzyme activity.

Chondroitinases show high substrate specificity and this has been characterized but there is no sequence or structure-function information available with respect to characterization of active site residues and mechanism. Kinetic studies describing the substrate specificities of chondroitinase AC from *A. aurescens* and *F. heparinum* have revealed that the *A. aurescens* chondroitinase AC produces disaccharides immediately suggesting an exolytic cleavage pattern, whereas the *F. heparinum* chondroitinase AC does not produce disaccharides until late in the reaction, suggesting a random endolytic cleavage pattern (Takegawa et al., 1991; Hiyama et al., 1975a,b; Hiyama et al., 1976; Hiyama et al., 1977). In a recent report, Gu et al., (1995) describe the purification, kinetic analysis and the N-terminal sequences for chondroitinase AC and chondroitinase B from *F. heparinum.*

2.4.2 *Hyaluronidases*

Hyaluronidase has been isolated from bacterial and animal sources. The enzyme activity is not a strong function of pH and they exhibit a broad slightly acidic pH optimum (Tam and Chan, 1985; Hamai et al., 1989). This observation suggests a fundamental difference in cleavage mechanisms since most other lyases exhibit slightly basic pH optima. A hyaluronidase from *Staphylococcus aureus* was inhibited by 10 mM IAA but not by cysteine, NEM, or EDTA (Vesterberg et al., 1968). Inhibition by IAA indicates that a cysteine residue might be involved in the catalytic mechanism.

While studying a hyaluronidase from a *Streptococcus* Group A bacterium, it was found (Greiling et al., 1975) that photo oxidation eliminated enzymatic activity and destroyed all histidine residues and some methionine and lysine residues as well. Incubation with histidine specific reagents (TLCK and TPCK) resulted in reduced activity suggesting the importance of histidine residues in enzyme activity. A model was proposed where histidine could play a nucleophilic role by abstracting a proton from the uronic acid and thus eliminating the glycosidic bond.

2.4.3 *Pectate lyases*

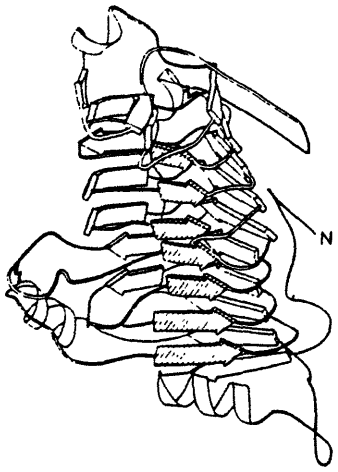
The crystal structure of the pelC pectate lyase from *E. chrysanthemi* has been solved by Yoder et al. (1993a). The structure revealed a unique secondary motif, a "β-helix", which is a coil of 22 residues per turn and 0.22 A° rise per residue (Yoder, 1993a; Cohen, 1993) **(Figure 2.7)** which comprises several parallel β-sheets which interact through hydrogen bonds. The interior of the spiral contains a series of asparagine residues which form a ladder. Calcium is required for pectate lyase activity and the crystal structure provides a possible calcium binding site, which is in a groove on the surface. The immediate vicinity of the calcium site consists of charged residues aspartic acid-170, lysine-172 and arginine-218, which are conserved among species (Hinton et al., 1989) and could be possible candidates for the active site.

The crystal structures of pectate lyase E from *E. chrysanthemi* (Yoder et al., 1993b) and pectate lyase from *Bacillus subtilis* (Pickersgill et al., 1994) were later solved and had similar three-dimensional folds. It is possible that such a unique fold could be conserved among other GAG lyases. Calcium is required for enzymatic activity for heparinases I and III as well and this is discussed in detail later in the subsequent chapters.

2.4.4 *Heparinases —a model system for protein engineering*

It is evident that there is very limited structure-function information available in the case of polysaccharide degrading lyases in general. The only available three-dimensional structure is for the pectate lyases. *F. heparinum* makes three heparinases—heparinases I, II and III which have been purified to homogeneity (Lohse and Linhardt, 1992). These enzymes are useful tools in elucidating heparin structure and have proven to be an ideal subject for study because of their unique substrate specificity.

Heparinase I from *F. heparinum* has been used in numerous applications including structural determination of heparin (Silva et al., 1976), preparation of low molecular weight heparin anticoagulants (Linhardt et al., 1982), antitumor agents (Folkman et al., 1983), and anti-atheroscelorotic therapeutics (Merchant et al., 1986). A very important application is the development of an immobilized enzyme filter for



*[Yoder, et al., Science, **260**, 1503, 1993.]*

Figure 2.7: The crystal structure of pel C pectate lyase from *E. chrysanthemi*.

blood deheparinization (Langer et al., 1982). It was showr that heparinases I and III are potent inhibitors of angiogenesis (Sasisekharan et al., 1994).

Hence it is important to understand the heparin-heparinase interactions at a molecular level. This interaction is a complex phenomenon involving specific domains of the enzyme (heparinase) binding to specific regions of the polysaccharide substrate (heparin) and subsequent cleavage of the heparin molecule which results from the interaction of catalytic residues. There is a high degree of specificity involved, which again could be dictated by structural requirements (i. e. specific sequences) in both heparinase as well as heparin. Until recently there has been very limited structural and mechanistic information concerning this enzyme. These studies will form a theme for this work and will be described in detail in the following chapters.

Part 3 Heparinases—Sources, Properties, and Applications

2.5 The Heparinases

Heparinases have been isolated from different sources; however, the only source of enzymes that have been actively studied and characterized are the heparinases I, II, and III from *Flavobacterium heparinum* (Linker and Hovingh, 1972; Dietrich et al., 1973; Silva et al., 1976).

2.5.1 Heparin lyases—sources and properties

Heparin lyases that cleave either heparin or heparan sulfate have been isolated from four different bacteria viz. *F. heparinum, Bacteroides heparinolyticus, Bacillus* sp., and an unclassified source. Several other bacteria are known to produce heparinases based on a plate assay but they have not been isolated and characterized (Salyers et al., 1977; Steyn et al., 1992). Heparinases have been classified as (McLean et al., 1985) heparinase I primarily acting on heparin; heparinase II acting on both heparin and heparan sulfate; and heparinase III acting primarily on heparan sulfate. In all, six heparinases have been isolated from the above mentioned four sources, with *F. heparinum* being the only one to make all the three heparinases.

Most of the heparinase producing bacteria were isolated from soil, although it is not known why a soil bacteria would produce an acidic polysaccharide degrading lyase. It is possible that these bacteria could use these enzymes to degrade the GAGs from carcasses. The flavobacteria species are capable of growing on several unnatural substrates such as nylon or oil (Sasisekharan, 1991). *F. heparinum* is classified under a gram negative group of bacteria, and is primarily found in soil and water. It produces several lyases such as heparinases, chondroitinases, and hyaluronidases. It uses heparin as a sole source of carbon, nitrogen and oxygen. In addition, sulfur plays a role in regulating the growth of the organism (Cerbelaud et al., 1986); under high sulfur conditions in the media, heparinase production is repressed.

Most of the heparinases are positively charged at neutral pH (isoelectric points in the range of 8.5-10), except for the *Bacillus* sp heparinase (pI of 6.6). One would assume that enzymes that degrade highly polyanionic substrates would be basic in nature. Hence it is possible that the *Bacillus* heparinase, which is relatively more acidic, cleaves via a different mechanism from the *F. heparinum* heparinases. Antibody assays and southern blotting between heparinases from different bacteria showed no cross-reactivity (Bellamy and Horikoshi, 1992; Sasisekharan et al., 1993), suggesting a low degree of sequence and structural homology among heparinases from different species.

2.5.2 *Purification of heparinases I, II, III from F. heparinum*

Heparin degrading activity was first observed (Linker and Hovingh, 1965) in a crude enzyme mixture obtained from induced *Flavobacteria*. Later they fractionated the crude enzyme activities into a heparinase (heparinase I) and a heparitinase (heparinase III) (Hovingh and Linker, 1970) and showed that both enzymes behaved as eliminases rather than hydrolases. Both these enzyme activities were partially purified but no physical characterization of these enzymes were performed. However, useful information on the substrate specificity was reported; heparinase fractions required the presence of both O- and N-sulfation and heparitinase fractions required the absence of O-sulfation and either N-acetylation or N-sulfation.

Subsequently, three lyases, a heparinase (heparinase I) and two heparitinases (heparinases II and III) were isolated from *F. heparinum* (Dietrich et al., 1973;

Michelacci and Dietrich, 1974; Ototani et al., 1981). A report on the substrate specificity of the two purified heparitinases was presented, but the physical properties of these enzymes were not characterized (Nader et al., 1990). Heparitinase I (heparinase II) degraded both N-sulfated and N-acetylated heparan sulfate, while heparitinase II (heparinase III) degraded primarily N-sulfated heparan sulfate. The specificity of a partially purified heparinase II was also studied and shown to possess a broad specificity on various substrates(Moffat et al., 1991a).

Heparinase I was purified to a single band on an SDS-PAGE gel (Linhardt et al., 1984). However, the first purification process for isolation of large scale, catalytically pure heparinase I was reported by Yang et al., (1985). By adapting a batch process of ion-exchange chromatography including hydroxyapatite and QAE-sephadex, a quick large scale purification was accomplished. Subsequently a quick two-step purification scheme was developed for large scale isolation of pure heparinase I, free of all contaminants such as chondroitinases, sulfatases, hyaluronidases, etc. that are present in *F. heparinum*.

A single purification method to purify all the three heparinases was recently developed and used to purify these enzymes to homogeneity (Lohse and Linhardt, 1992). The physical and kinetic characteristics were studied thoroughly and optimal reaction conditions were established and these are summarized in **Table 2.1.**

Table 2.1 Properties of heparinases I, II, and III from *F. heparinum*.

Enzyme	M W (kDa)	pH opt	Temp opt (°C)	k_{cat} (s^{-1})	K_m μM	Metal inhibitor	Metal activator
heparinase I	42	7.15	35	140	17.8	Zn, Hg, Cu	Ca
heparinase II	84	6.9	40	27-51	57.7[a] 11.2[b]	Ca, Hg, Li, Zn, Cu, Li	
heparinase III	70	7.6	45	165	29.4	Zn, Hg, Cu	Ca

[a] Substrate heparin

[b] Substrate heparan sulfate

From Table 2.1 it is clear that all three enzymes (pI 8.5-9.2) are basic and positively charged at their pH optima. Calcium activates heparinases I and III but inhibits heparinase II. Calcium binds heparin and causes a structural change (Hunter et al., 1988). Considering this observation, the differential effect of calcium on heparinase activity is interesting and could possibly be a factor in controlling specificity of these enzymes. Heparinases I, II, and III cleave the heparin backbone with a high degree of specificity. They recognize unique and specific sequences within the heparin chain as discussed in the following section.

2.5.3 Substrate specificities of heparinases I, II, and III

Based on the disaccharide sequence to be cleaved, the substrate specificity is defined as follows. Heparinase I cleaves linkages of the type $H_{NS,\ 6X}$-I_{2S}, heparinase II cleaves linkages $H_{NY,6X}$-U_{2X}, and heparinase III recognizes the linkages H_{NAc}-I and $H_{NY,6X}$-G, where Y is either sulfated or acetylated and X is sulfated or unsubstituted.

2.5.3.1 Substrate specificity of heparinase I

Heparinase I primarily degrades heparin, although heparan sulfate is degraded also to some extent. The degradation products have been separated using strong anion-exchange (SAX) HPLC by Linhardt and coworkers and extensively characterized (Rice and Linhardt, 1989; Linhardt et al., 1990; Nader et al., 1990; Linhardt et al., 1992b) . **Table 2.2** lists all the oligosaccharide fragments characterized thus far as they elute from SAX-HPLC column. These fragments were characterized by various NMR techniques.

Table 2.2 Substrate specificities of heparinase I

Peak number	Sequence	Mole %
1	U_{2S}-H_{NS}	2.5
2	U-$H_{NS,6S}$	0.5
3	U_{2S}-$H_{NS,6S}$	71.0
3a	U_{2S}-H_{NS}-I_{2S}-H_{NS}	0.8
3b	U_{2S}-H_{NS}-G_{2S}-$H_{NS,6S}$	0.6
4	U_{2S}-H_{NS}-I_{2S}-$H_{NS,6S}$	3.1
4a	U_{2S}-$H_{NS,6S}$-I-$H_{NS,6S}$	0.5
5	U_{2S}-$H_{NS,6S}$-G-$H_{NS,6S}$	6.9
6	U_{2S}-$H_{NS,6S}$-I_{2S}-$H_{NS,6S}$	3.2
6a	U_{2S}-$H_{NS,6S}$-I-$H_{NAc,6S}$-G-$H_{NS,6S}$	0.8
7	U_{2S}-$H_{NS,6S}$-G-$H_{NS,6S,3S}$	0.6
8	U_{2S}-$H_{NS,6S}$-I-$H_{NAc,6S}$-G-$H_{NS,6S,3S}$	2.2

adapted from (Linhardt et al., 1992)

The trisulfated disaccharide (peak 3) is the major product and corresponds to the most frequently occurring linkage in heparin. Although the tetrasaccharides (peaks 4 and 6) contain a linkage that is recognized by heparinase I, they are not cleaved indicating that they are poorer substrates compared to full length heparin. Indeed, when heparinase I was incubated with the tetrasaccharides they were cleaved at a much slower rate than heparin (Rice and Linhardt, 1989; Linhardt et al., 1990). The tetrasaccharides containing glucuronic acid do not get cleaved suggesting that iduronic acid is an essential requirement for catalysis. On the hexosamine side, 6 sulfation does not seem to have an effect but N-acetylated residues seem to be less susceptible to cleavage (Yang et al., 1985; Desai et al., 1993a). In summary, it can be said that heparinase I is specific for linkages of the type $H_{NS,6X}$-I_{2S}, with the additional constraint being the size of the fragment—hexa- or higher order fragments.

2.5.3.2 *Substrate specificity of heparinase II*

Heparinase II cleaves both heparin as well as heparan sulfate type linkages, with the heparan sulfate activity being approximately twice as high as heparin activity (Linhardt et al., 1990; Nader et al., 1990; Lohse and Linhardt, 1992). Studies with chemically modified heparin showed that N-acetylated, 2-O or 6-O sulfated heparin were all cleaved by heparinase II (Moffat et al., 1991b; Desai et al., 1993), thus indicating a broad specificity. The major tetra-and hexasaccharide products from heparinase I degradation, that contained glucuronic acid, were degraded by heparinase II. It is intriguing that both the isomers of uronic acid are recognized by this enzyme. This suggests that the enzyme's active site is able to accommodate both the epimers, indicating the versatility of the catalytic domain. However nothing is known about the active site structure of this enzyme. As mentioned previously, the susceptible linkages for heparinase II are of the type $H_{NY,6X}$-U_{2X}, where X can be either sulfated or unsubstituted; N can be acetylated or sulfated; and U can be either iduronic or glucuronic acid.

2.5.3.3 *Substrate specificity of heparinase III*

Heparinase III is very specific for heparan sulfate like sequences. When heparinase III was incubated with heparin, no products could be detected (Linhardt et al., 1990; Desai et al., 1993). The products of heparan sulfate degradation by heparinase III have very low sulfation. Larger fragments are better substrates compared to shorter oligosaccharides. Heparinase III cleaves the heparinase I tetrasaccharide containing glucuronic acid but is inactive against oligosaccharides with linkages containing iduronic acid (Silverberg et al., 1985; Rice and Linhardt, 1989; Linhardt et al., 1990). However, a recent study reported that heparin that was O- and N-desulfated and N-acetylated was cleaved by heparinase III, suggesting that unsulfated iduronic acid containing linkages can also be cleaved (Desai et al., 1993). Hexosamines containing either N sulfation or acetylation or 6-sulfation were all cleaved, indicating that N- and 6-sulfations do not interfere with cleavage. Thus results seem to indicate that the general linkage recognized by heparinase III is of the form H_{NAc}-I or $H_{NY,6X}$-G.

Recent studies with tetrasaccharides derived from the ATIII binding site in heparin showed that this was resistant to cleavage by all the three heparinases (Rice and

Linhardt, 1989; Yamada et al., 1993). The generic structure of the tetrasaccharide was determined to be U-H_{NAc}-G-$H_{NS,6X,3S}$ (Yamada et al., 1993). Thus in addition to the constraints within the H-U linkage, the H-U-H linkage also seems to be important. The uronic acid with a hexosamine having 3-O sulfation on the reducing side, does not get cleaved at the non-reducing end too. The specificity of the enzyme and the substrate involves a chemical recognition of atleast three saccharide units.

2.5.4 Heparinases I, II, III—clinical and diagnostic applications

The importance of HSPGs in the ECM and their role as active biological modulators was described earlier. Their degradation would be expected to have significant regulatory consequences. HSPG catabolism is observed in inflammation, wound healing, diabetes, and cancer metastasis, suggesting that these enzymes play important roles in pathological and physiological processes (Hoogewerf et al., 1995). Heparanase activity has been described in activated immune system cells and cancer cells; however due to lack of biologic tools, the potential cause of heparanase in diseased states is not understood (Hoogewerf et al., 1995).

Heparinases have proven to be useful tools for studying GAG structure since these enzymes are specific for certain sequences of the heparin biopolymer. In addition, heparinases by themselves have significant therapeutic and clinical applications. The use of heparinase I *in vitro*, for clinical monitoring of heparin concentration has been approved by the FDA (Tejidor et al., 1993). Heparinase I is in phase I clinical trials for systemic heparin neutralization *in vivo* (Gen Eng. News, 1995).

However, the most significant therapeutic applications of the heparinases are as agents for heparin neutralization after extracorporeal therapies and as potent inhibitors of angiogenesis. These two applications have an enormous impact on the development of new medical technologies and provide a strong motivation towards understanding the mechanisms governing the heparin-heparinase interactions. These use of heparinases in these two therapies is described in greater detail in the following sections.

2.5.4.1 *Immobilized heparinase I—a route to heparin neutralization*

Every year thousands of operations involving artificial kidneys, open-heart operations employing cardio-pulmonary bypass procedures are performed (Langer et al., 1982). In such procedures heparin is administered intravenously since systemic levels of heparin, which fully anti-coagulate the patient, are required **(Figure 2.8)**.

Blood re-enters the body after passing through the device and the presence of heparin would lead to severe hemorrhagic and bleeding complications. Hence, it is important to deheparinize the blood. This is usually done by administering a cationic polypeptide known as protamine sulfate, which binds non-specifically to heparin. However, protamine has its own associated side effects such as causing hyper-tension and its anti-heparin effect is highly variable. In addition, the protamine-heparin complex has been linked to immune related loss of platelets (Linhardt and Loganathan, 1990).

An approach to solve this problem of systemic heparinization was proposed (Langer et al., 1982). It makes use of the ability of heparinase I to degrade heparin to smaller oligosaccharide fragments which do not possess the anticoagulant property. Heparinase I was immobilized onto a filter and placed at the effluent of an extracorporeal device **(Figure 2.8)**. This effectively cleaved heparin to smaller fragments which subsequently get discarded from the body. The high degree of specificity of heparinase I for heparin makes this approach very effective in neutralizing the anti-coagulant activity of heparin since it gets catalytically cleaved.

The major advantages of this approach include: a) obviating the need of using protamine which is a potentially toxic neutralizing substance and b) opportunity of heparinizing the extra-corporeal system without simultaneous heparinization of the patient. Feasibility studies were conducted using filters of immobilized heparinase with human blood and in dogs (Langer et al., 1982). Heparinase I was purified and immobilized to sepharose. The reactor used is shown in **Figure 2.9**. In *in vitro* experiments, when heparinized human blood was passed through the filter, 60% of heparin activity was destroyed in 2 min (one pass through the filter) and almost all activity was gone in 6 min (three passes). Similar results were obtained in *in vivo* experiments using dogs. The enzymatically degraded fragments were tested

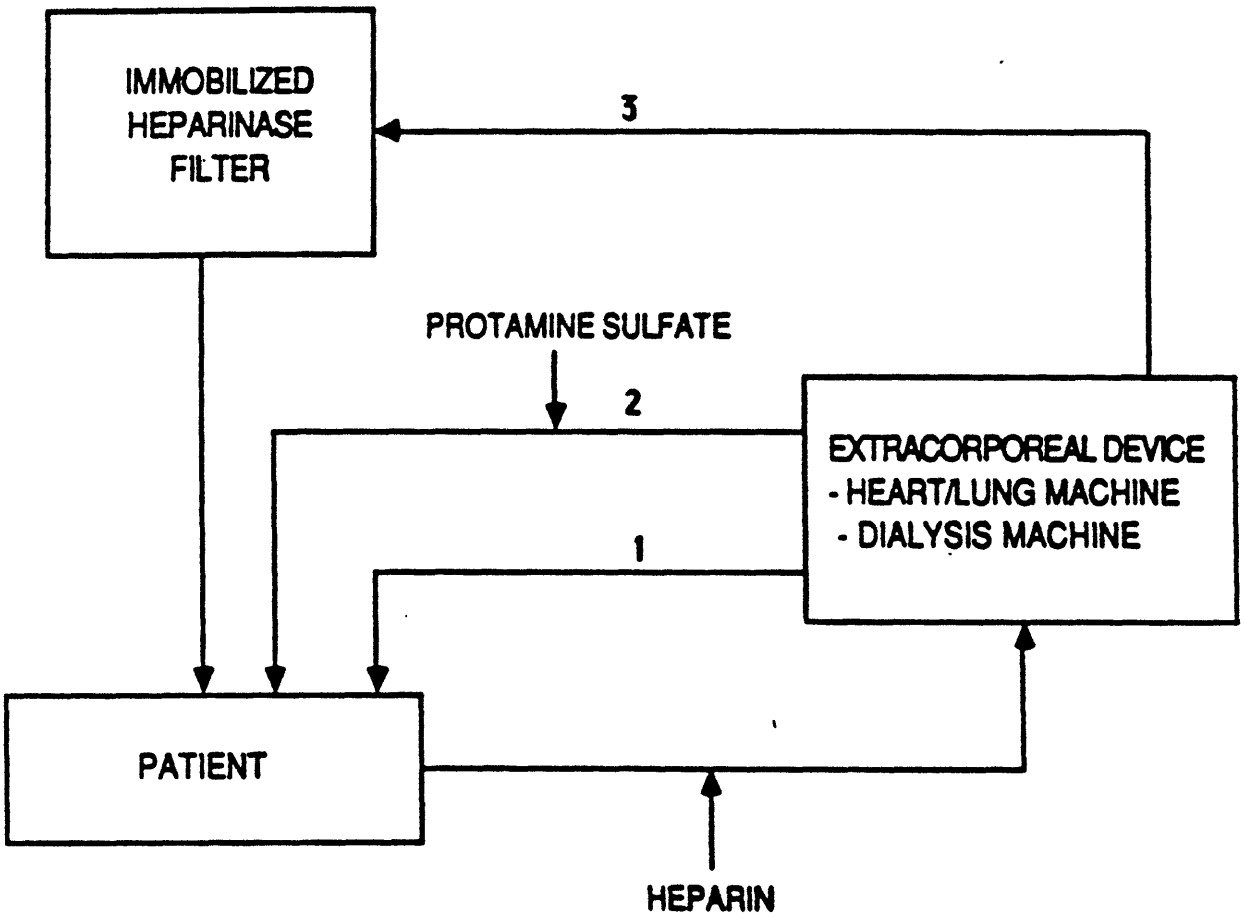
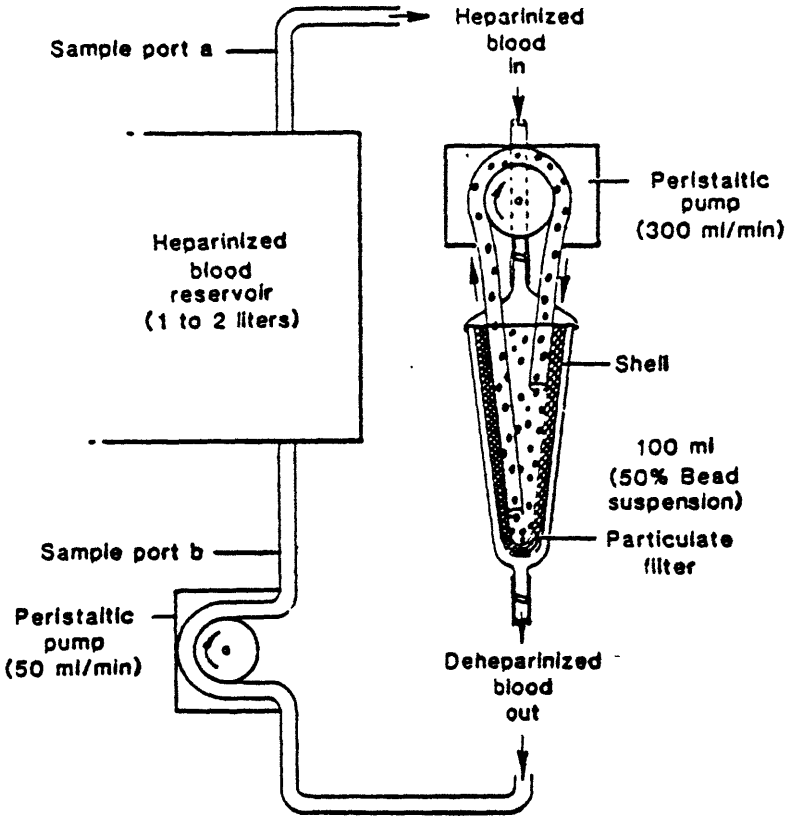


Figure 2.8: Immobilized heparinase I in the extra-corporeal circuit.



[Langer et al., Science., 217, 261, 1982]

Figure 2.9: Immobilized heparinase reactor system.

for toxicity effects and no effects were observed. This approach presented a novel and effective way of solving the hemorrhagic complications arising due to systemic heparinization. Currently, ongoing efforts are aimed at optimizing conversion, efficiency and enzyme stability for applications involving prolonged use such as hemodialysis (Ameer, 1995).

2.5.4.2 *Heparinases as inhibitors of angiogenesis*

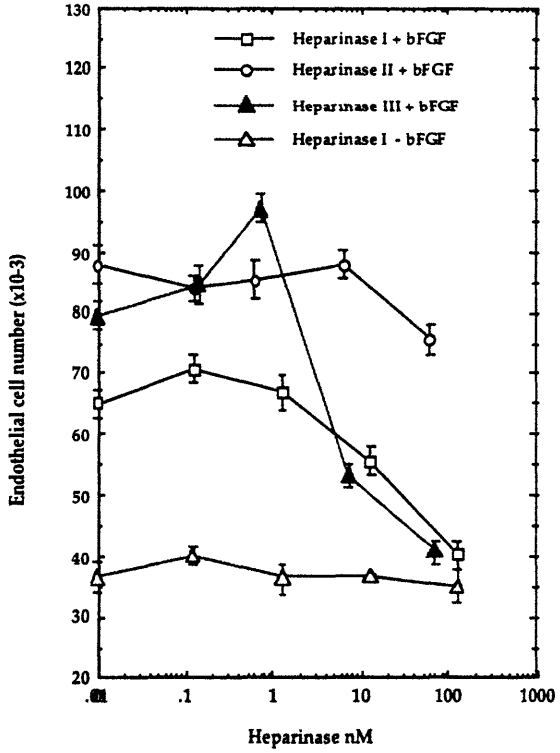
Angiogenesis or neovascularization is the process of formation of new blood vessels. It is a highly regulated phenomenon and involved in both normal and pathological events (Klagsburn and D'Amore, 1991). Heparin like molecules are found distributed on various cell surfaces and bind several cytokines. These cytokines are angiogenic factors, and heparin modulates their function by either stabilizing them or controlling their bioavailability (Folkman and Shing, 1992). Since angiogenic stimulants such as bFGF bind heparin, the role of heparinases I, II, and III in modulating angiogenesis was investigated (Sasisekharan et al., 1994). Inhibition of neovascularization *in vivo* was studied using the chorioallantoic membrane (CAM) assay **(Figure 2.10)**. When heparinases I, and III were used in the assay, approximately 4 μg of the enzymes were sufficient to observe avascular zones or 100% inhibition of embryonic neovascularization. However heparinase II did not have any effect in controlling neovascularization.

bFGF is a potent mitogen for bovine capillary endothelial cells (BCEs) and is considered to be an important mediator of neovascularization (Nugent and Edelman, 1992). Hence the effect of the three heparinases on BCE proliferation *in vitro* in the presence and absence of bFGF was investigated (Sasisekharan et al., 1994). While heparinases I and III significantly inhibited bFGF mediated BCE proliferation, heparinase II showed no effect on BCE proliferation **(Figure 2.11)**. Binding of ^{125}I-bFGF to heparan sulfate and receptors suggested that the mechanism of inhibition of BCE proliferation was associated with impaired bFGF binding to its receptor.

Heparinases I and III are among the potent anti-angiogenic inhibitors reported to date. BCE proliferation is a result of a tri-molecular interaction of a growth factor (such as bFGF), a high affinity receptor, and a low-affinity receptor (heparan sulfate)

*[Sasisekharan et al., P. N. A. S., **91**, 1524, 1994]*

Figure 2.10: Inhibition of neovascularization using the CAM assay.
a) Normal CAM containing an empty methylcellulose disk
b) CAM, with heparinase I (4 μg) containing disk, with avascular zone.



*[Sasisekharan et al., P. N. A. S., **91**, 1524, 1994]*

Figure 2.11: Effect of heparinases on bovine capillary endothelial cell proliferation.

and this subsequently leads to angiogenesis (Folkman and Shing, 1992). Heparinases I and III inhibit neovascularization by cleaving heparin-like and heparan sulfate-like regions of the cell surface HSPG respectively and thus depleting the low affinity receptors required for bFGF-mediated BCE proliferation. A better understanding of the underlying mechanisms could offer opportunities for therapeutic intervention in the treatment of pathologies ranging from abnormal wound healing to tumor angiogenesis (Sasisekharan et al., 1994).

In summary, it is clear that the heparinases have enormous clinical and therapeutic potential. Mechanisms governing these phenomena are poorly understood. An understanding of the heparin-heparinase interaction and the catalytic mechanism is lacking. Hence, efforts were directed towards understanding the structure-function relationships of the enzyme. Heparinase I, in particular, has been the focus of this research and has been extensively studied and characterized. Prior to this study preliminary work was done (Sasisekharan, 1991) towards understanding heparinase I structure-activity relationships and the following sections describe salient features of this work.

2.6 Recent Background on Heparinase I

Two recent studies laid the foundation for this work. The first study involved cloning of the heparinase I gene followed by biochemical characterization of an active site residue of the enzyme (Sasisekharan, 1991). The other study (Venkataraman, 1992) involved the identification and mapping of a heparin binding site in heparinase I. This work is, not only aimed at investigating both these themes in greater detail, but also weaves the interplay between binding and catalysis for this polysaccharide degrading lyase.

2.6.1 Cloning of heparinase I and recombinant expression in E.coli

The gene for heparinase I was cloned; sequenced and expressed in *E. coli* . The gene had an open reading frame (ORF) of 1152 bps corresponding to 384 amino acids **(Figure 2.12)**. The N-terminal amino acid sequence revealed 17 amino acid peptide with a typical prokaryotic Ala-xxx-Ala signal sequence (Sasisekharan, 1993).

Figure 2.12: Heparinase I amino acid sequence.

Three cysteine residues were present in the protein sequence, with one of them in the signal sequence. Interestingly, the sequence also revealed the presence of a heparin binding consensus sequence (Cardin and Weintraub, 1989). Recombinant expression of heparinase I with the signal sequence was unsuccessful (Sasisekharan et al., 1993). The construct without the signal (leader) sequence (-L pET) was expressed in *E. coli*, as an active protein. This indicated that the cysteine residue in the leader sequence is not essential for enzyme activity. However, the molecular weight of this protein was about 1 kDa less than that of the native *F. heparinum* heparinase I. Recently it was shown that heparinase I is glycosylated with the molecular weight of the sugar being about 1 kDa (Huang, 1995). This could account for the observed difference in molecular weight.

Heparinase I expression was optimized using other expression systems such as pET-15b and pET-28a (Novagen, WI) (Ernst et al., 1996). Heparinase I was expressed as a soluble protein in pET-15b with optimal yields. High levels of expression were obtained using pET-28a, where the protein was expressed as inclusion bodies and refolded. The principle underlying the purification scheme is described in the Materials and Methods chapter.

2.6.2 The cysteines of heparinase I

Early studies on heparinase I immobilization provided some information on important amino acid residues in heparinase I. Modification by amine-reactive reagents and immobilization of heparinase I on amine-reactive supports resulted in extensive activity losses, suggesting the importance of lysines (Bernstein et al., 1988; Comfort et al., 1989; Leckband and Langer, 1991). The effects of sulfhydryl modifications demonstrated that cysteine reactive probes resulted in loss of enzymatic activity and the presence of a free cysteine (Leckband and Langer, 1991). These observations provided only a qualitative description of the active site or its surrounding environment.

The number of cysteine residues in heparinase I remained unclear. Earlier biochemical studies reported that heparinase I contained three to four cysteines (Yang et al., 1985), and the presence of a disulfide bond (Comfort et al., 1989). However, the heparinase I primary sequence revealed three cysteines, one of them

being in the leader sequence. Hence Sasisekharan (1991) addressed the number, the states and the role of cysteines in heparinase I activity. The approach consisted of a combination of chemical modification studies and tryptic mapping experiments. The results provided valuable insight into the active site of heparinase I and hence are briefly summarized below.

2.6.2.1 *Pyridylethylation of cysteine residues*

To determine the number of cysteine residues through amino acid analysis (AAA) it is important to modify the cysteine residue during Edman degradation. 4-vinyl pyridine (4-VP), a hydrophobic group was used to modify cysteine residues since this modification has high sensitivity and easy detection (Sasisekharan, 1991). AAA results on the 4-VP modified enzyme indicated two cysteine residues. Tryptic mapping confirmed these to be cysteine-135 and cysteine-297. Hence it remained to be seen if one or both of these was in the active site of heparinase I.

2.6.2.2 *Chemical modification of heparinase I*

Thiol modification with IAA

In the presence of 2 mM IAA 95 ± 5% of heparinase I was inactivated within 10 min. The inactivation rate was concentration dependent and pre-treatment of the enzyme for 4 h at 4°C with DTT under nitrogen had no effect on the modification. The IAA reactive form of cysteine is the mercaptide anion, and the reaction rate increases with increasing pH (Torchinsky, 1981). In particular, the relative free cysteine alkylation rates at pH 5.6, pH 7.02, and pH 8.36 are 0.14, 1, and 2.1, respectively. When the IAA inactivation was studied as a function of pH, the pseudo first order heparinase I inactivation rate constants were independent of pH in the range studied (6.5-8.5) (Sasisekharan, 1991). The small variation in the rate of the inactivation suggested that the cysteine was activated by the presence of nearby basic amino acids (Rabin and Watts, 1960).

Inactivation with PCMB

PCMB, a reversible, sulfhydryl specific anion, was used to confirm the sulfhydryl selective inactivation of heparinase I. Heparinase I treatment with PCMB at 2.5-100

μM and 4°C resulted in a reversible loss of 95 ± 5% of enzyme activity, verifying the sulfhydryl specificity of the reaction (Sasisekharan, 1991). In the presence of 0.5 mg/ml heparin (~5 x K_m), the pseudo-first order rate of inactivation was significantly decreased. These results suggested the presence of a cysteine residue in or near the active site or the heparin binding site.

Radio-labeling with [^{3}H]iodoacetic acid

Carboxymethylating heparinase I cysteines with [^{3}H]IAA was used to confirm the quantification of the number of cysteines obtained from pyridylethylation. Results were consistent with two residues being labeled under denatured conditions. Further, prior treatment of heparinase I with PCMB reduced [^{3}H]IAA labeling by 80% suggesting that the IAA binding site and the PCMB binding site were identical.

Tryptic mapping of the cysteines of heparinase I

In order to map the PCMB reactive cysteine, the PCMB-labeled heparinase I was isolated, denatured and then reacted with iodoacetamide to block the other cysteine. Following this, the enzyme was treated with DTT to remove the bound PCMB, and then labeled with [^{3}H]IAA. Modified heparinase I was digested with trypsin and the tryptic peptides were separated. Only one cysteine, cysteine-135, was selectively labeled by [^{3}H]IAA. In another experiment, heparinase I was first labeled at the reactive cysteine with PCMB. The enzyme was denatured, labeled with [^{3}H]IAA, and re-chromatographed to remove the excess radiolabel. Following this, the enzyme was digested with trypsin and the tryptic peptides were separated by RPHPLC. Cysteine-297 was selectively [^{3}H] labeled, while cysteine-135 was not. These results confirmed that cysteine-135 was the PCMB-labeled or the active site cysteine (Sasisekharan, 1991).

Thiol reactivity to negatively charged compounds

Treatment of heparinase I with 1 mM NEM (neutral compound) at pH 7.0, showed little change in activity over an 8 h period (Sasisekharan, 1991). Further, heparinase I treated with 1 mM NEM overnight at 4°C, resulted in an activity loss of about only 15%. In addition no significant levels of [^{3}H]NEM bound heparinase I even under denaturing conditions.

The effect of reagent charge on cysteine reactivity was also investigated by using IAA's neutral analog IAM. Heparinase I incubated with 2, 5, 10, and 120 mM IAM in PBS at pH 7.0 and 4°C for up to 24 h exhibited little change in activity. Further, the labeling of IAM was not different under denaturing conditions. IAM, therefore, does not significantly modify heparinase I. The selective reactivity of the thiol group to negatively charged compounds strongly suggested the presence of a positively charged environment i. e. basic amino acid residues in proximity.

2.6.2.3 Preliminary work on active site characterization

Some interesting and significant questions that arose from the above study (Sasisekharan, 1991) were the following : a) What is the mechanistic role of cysteine-135 in the enzymatic pathway? b) Is there a positively charged environment around cysteine-135? If so, what is the role of this positive charge? These questions were addressed first, as part of this research towards understanding the role of cysteine-135 in heparinase I activity and the surface properties around this residue.

One possible interpretation of the sulfhydryl labeling experiments is that the chemical modification of cysteine-135 may affect the enzymatic activity by steric impedance of heparin access to the active site. Another possible explanation is that heparin bound to heparinase I could alter the charge characteristics near cysteine-135 and thereby reduce the labeling kinetics. While chemical modification experiments give important information about the surface properties around the residue, they fail to identify a mechanistic role for a residue in enzyme activity. Hence, in order to ascertain a mechanistic role of the surface-accessible cysteine in the enzymatic pathway, site-directed mutagenesis was performed. This is a very powerful technique to examine the effects of point mutations; multiple mutations within a domain; and deletions and insertions of specific domains. A single residue can be altered and the effect on the properties of the resulting enzyme can be studied.

Results from preliminary work

A) Site-directed mutagenesis of heparinase I

Site-directed mutations were performed to confirm the role of cysteine-135 in heparinase I activity. The experimental details are described in the next chapter. Alanine was chosen as a residue for substitution due to its small size and neutral charge. Cysteines-135 and -297 were individually altered to alanires (C135A and C297A). To gain insight into the nucleophilic role of the cysteines, two additional mutant recombinant heparinases were designed (C135S and C297S) since serine is also a potential nucleophilic residue.

All mutant enzymes were expressed in BL21(DE3) (Novagen, WI; see Materials and Methods) host. *r*-heparinase I construct devoid of the putative signal sequence (-L *r*-heparinase I) was expressed as a control. The level of protein expression for all the recombinant heparinases was identical in the BL21(DE3) host. While -L *r*-heparinase I control was expressed as a soluble protein in *E. coli* with an activity of ~ 5.2 U/mg of *E. coli* crude extract, the C135A *r*-heparinase I was expressed in BL21(DE3), with no enzymatic activity. The C135S *r*-heparinase I, however, was expressed in BL21(DE3) with an activity of ~ 0.06 U/mg of *E. coli* crude extract. Importantly, the mutations at cysteine-297 (C297S and C297A) were both expressed in the same host with no change in their enzymatic activity compared -L *r*-heparinase I control. All recombinant enzymes were purified using Ni-affinity chromatography (see chapter 3). **Table 2.3** lists the kinetic parameters for the recombinant and mutant heparinases. The purity of *r*-heparinases as determined by SDS-PAGE was estimated to be 80-90% after the first step and greater than 98% with silver stain after the second step **(Figure 2.13)**.

Table 2.3 Kinetic constants of recombinant heparinase I, and mutant enzymes

Enzyme	*kcat sec^{-1}*	*K_m μM*
recombinant wild type (-L)	92	10.2
C135S	2	4.2
C135A	ND[a]	ND

ND[a]= not determined since the enzyme was inactive.

The kinetic constants reported in this work for wild type enzyme are consistent with previous reports (Lohse & Linhardt, 1992) for the native enzyme.

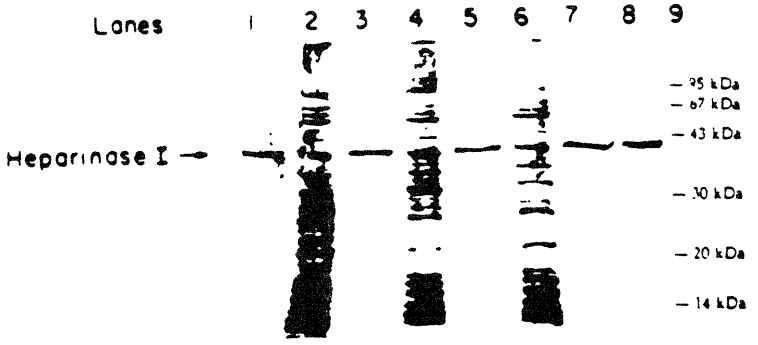


Figure 2.13: *SDS-PAGE of r-heparinase and mutant heparinase I:* A 12% SDS-PAGE of heparinase I. Lanes are described from left to right. **Lane 1:** purified *F.heparinum* heparinase I. **Lane 2:** crude extract of BL21(DE3) with C135A mutant heparinase I. **Lane 3:** Ni column purified C135A mutant heparinase I. **Lane 4:** crude extract of BL21(DE3) with C135S mutant heparinase I. **Lane 5:** Ni column purified C135S mutant heparinase I. **Lane 6:** crude extract of BL21(DE3) with recombinant heparinase I. **Lane 7:** Ni column purified recombinant heparinase I. **Lane 8:** purified *F.heparinum* heparinase I. **Lane 9:** Molecular weight standards.

Replacement of cysteine-135 with alanine inactivated the enzyme, and demonstrated the importance of this amino acid in heparinase I activity. Replacement of cysteine-135 with serine drastically reduced heparinase I activity and suggested a nucleophilic role of this residue in the endolytic cleavage of heparin (Linhardt et al., 1986). The above results taken together show that cysteine-135 is important for heparinase I activity and is part of the active site. Further, altering cysteine-297 does not alter heparinase I activity, suggesting that this residue is not important for heparinase I activity.

B) Effect of charge and electrostatic interactions

The preferential reactivity of cysteine-135 towards negatively charged compounds such as PCMB and IAA over neutral compounds such as NEM and IAM (which failed to inactivate heparinase I) suggested the presence of a positively charged environment around cysteine-135. In order to show that the positively charged environment around the reactive cysteine influences the labeling of negatively charged PCMB, cysteine labeling by PCMB was performed under different salt concentrations. **Figure 2.14** shows the time course of inactivation of heparinase I by PCMB with increasing salt concentrations. Rate of heparinase I inactivation by PCMB was significantly reduced with increasing salt concentration. This result is consistent with the observation that the environment around the reactive cysteine-135 is positively charged. It should be pointed out here that the salt concentration range used in the study was based on the previously studied conditions for both enzyme stability as well as activity (Lohse and Linhardt, 1992).

Discussion—Active site of heparinase I

It has been proposed that nucleophilic amino acids play an important role in the catalytic activity of polysaccharide lyases in general (Linhardt et al., 1986). These residues abstract the C5 proton on the uronate of the disaccharide repeat unit of the acidic polysaccharides, and initiate the elimination based depolymerization reaction. The preliminary results from this study in conjunction with the previous work (Sasisekharan, 1991) strongly suggest that cysteine-135, surrounded by a positively charged environment, is an important nucleophile in the elimination reaction of

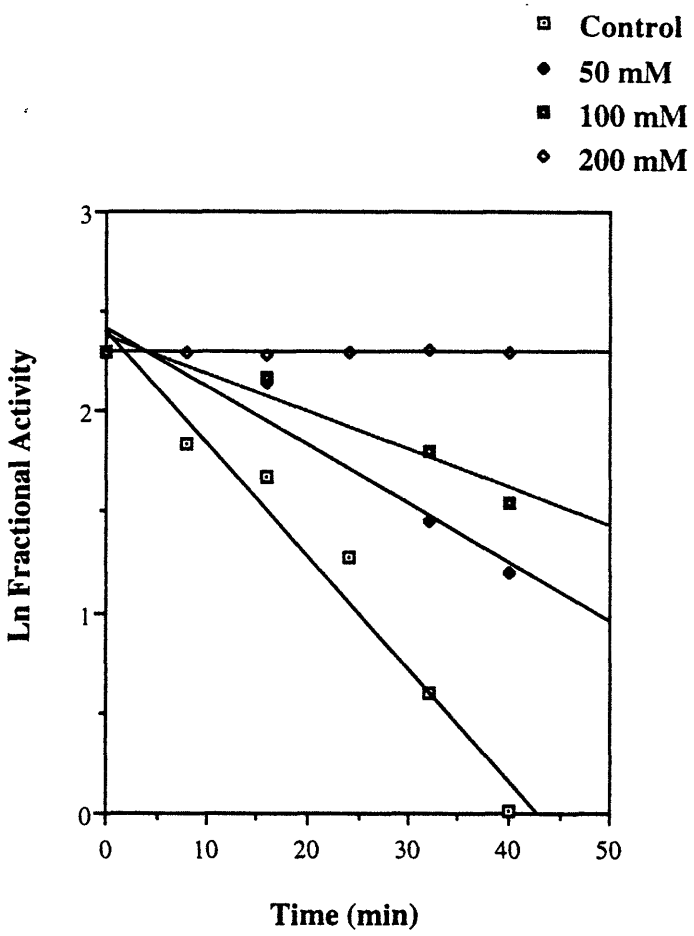


Figure 2.14: *PCMB inactivation of heparinase I in varying salt concentrations.* Heparinase I was incubated with 2.5 μM PCMB in 10 mM MOPS (pH 7.0) at 4 °C. (a) control containing no salt; (b) containing 50 mM NaCl; (c) containing 100 mM NaCl; (d) 200 mM NaCl. The salt concentration range used in this study are based on the previously studied conditions for both the enzyme stability as well as activity (Yang et al., 1984; Lohse and Linhardt, 1992). The values are averages of triplicate determinations.

heparinase I.

The thiol group of cysteines has a pKa of 8.35 in free solution (Fersht, 1985), indicating that this residue will be fully protonated at pH 7.0 - the pH optimum for heparinase I (Yang et al., 1985). It is possible, however, that a positively charged environment will tend to keep the thiol group negatively charged (i.e. lowering its pKa) so that it can act as a base for proton abstraction. Similar behavior was observed in other systems in which the cysteine reactivity was due to a highly charged or a non-polar environment (Riordan and Vallee, 1972). More specifically, cysteine reactivity was enhanced by the proximity of basic amino acids (Rabin and Watts, 1960; Torchinsky, 1981), accounting for the high reactivity of cysteine-135 at pH 6.5. Further, the replacement of cysteine-135 with a weaker nucleophile (serine) using site-directed mutagenesis significantly reduced heparinase I activity, consistent with a possible proton abstraction role.

This model of a positively charged active site environment is expected since the substrate heparin is a highly sulfated, negatively charged polymer. Further, heparinase I, has a pKa of 9.1 and it also contains a heparin binding consensus motif made up of a cluster of basic residues, and exhibits high degree of specificity for heparin. Chemical modification studies of heparin binding proteins such as ATIII and bFGF implicate lysines in heparin binding (Sasisekharan et al., 1995). It is likely that heparin binds to heparinase I via charge complementarity (with basic residues), and the active site environment perhaps plays a key role in biasing the active site reactivity.

2.6.3 The Heparin Binding Site of Heparinase I

In section 2.3 heparin-protein interactions were discussed and the isolation of specific sequences in proteins that bound heparin was described with examples of ATIII, bFGF, and LPL. The discussion also included the polysaccharide side, where oligosaccharides with certain sequences of modifications in the disaccharide repeat that confer binding specificity, have been isolated for AT III, bFGF, LPL and others. This is parallel to the substrate specificities of heparinases, and suggests that the mechanism of substrate binding to heparinases may be involved in generating specificity. Hence this motivated the identification of a primary heparin binding site in heparinase I. The approach and key results are summarized below.

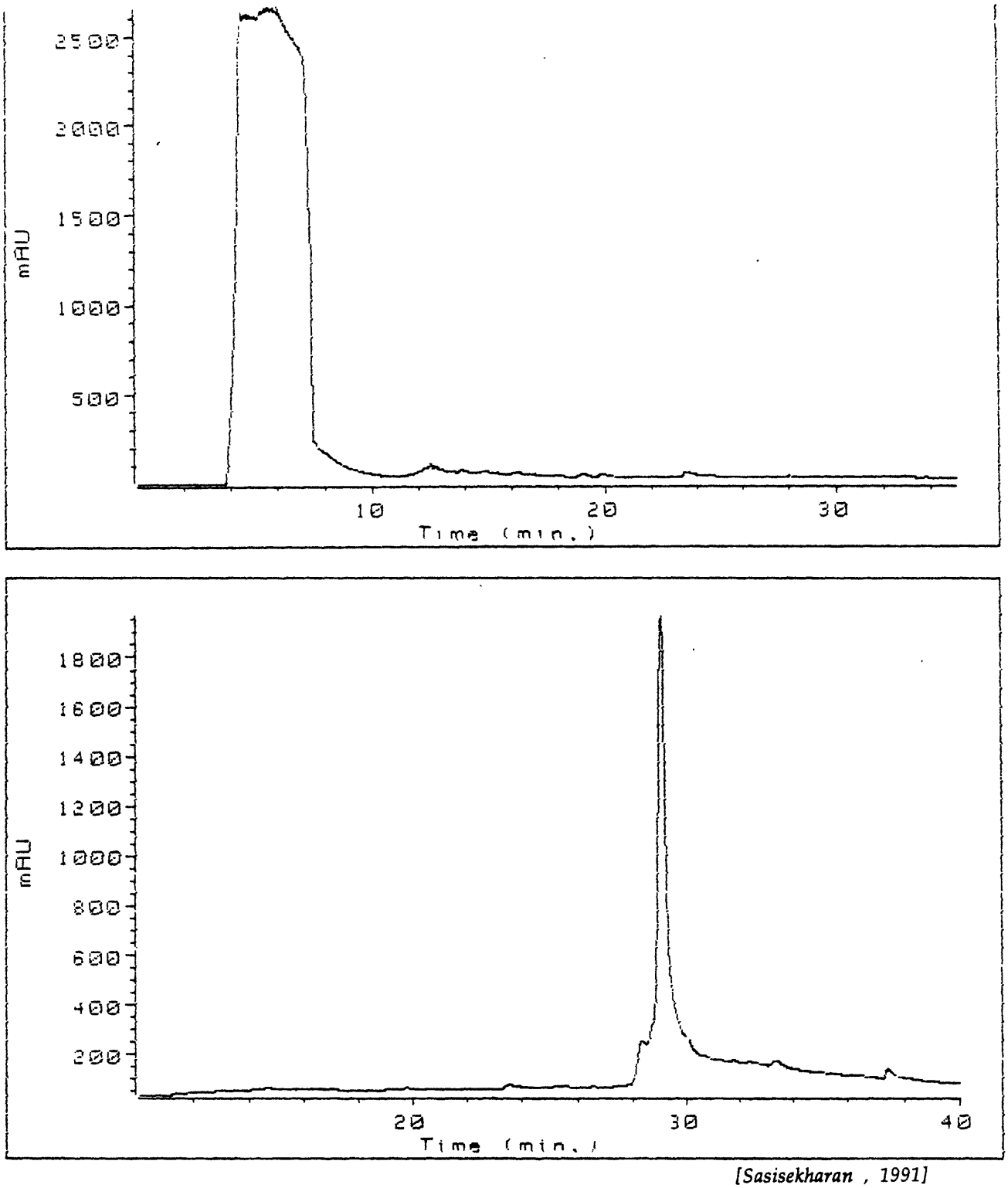
2.6.3.1 *Heparin blotting of CnBr digests of heparinase I*

Heparinase I was digested by cyanogen bromide (CnBr), which cleaves at methionine residues, separated by SDS-PAGE, and the resulting peptide fragments were transferred onto a nitrocellulose membrane (Venkataraman, 1992). The membrane was hybridized with ^{125}I labeled heparin, and the fragments were counted for ^{125}I incorporation. Results indicated that the region 195-270 contained the primary heparin binding site. The N-terminal region was lysine rich , containing two potential Cardin-Weintraub heparin binding consensus sequences and a calcium binding loop of the 'EF-hand' structural domain (Kretsinger, 1980). To further narrow down the heparin-binding region, tryptic digests were performed, when heparinase I would be cleaved to smaller fragments.

2.6.3.2 *Competition experiments with tryptic digests of heparinase I*

Even though heparinase I is a highly basic protein (having a pI of 9.1), it binds well to a hydrophobic surface as its elutes at a relatively high acetonitrile concentration in RP-HPLC (Sasisekharan, 1991) **(Figure 2.15 a)**. Interestingly, it was found that heparin, but not chondroitin sulfate, prevented heparinase I binding to a reverse-phase column in a concentration dependent manner **(Figure 2.15 b)**. This prompted further investigation of heparin's effect on binding of heparinase I or heparinase I fragments to a reverse-phase column.

Tryptic digestion was used to map the heparin binding site of fibroblast growth factor, by showing that heparin protects this site from trypsin cleavage (Coltrini et al., 1993). The ability of heparin in protecting the heparin binding domain of heparinase I from trypsin cleavage was tested (Venkataraman, 1992). Under the conditions tested, heparin was ineffective in protecting the heparin binding domain; however, it was able to compete specifically with the binding of some heparinase I tryptic peptides to the reverse phase column. Peaks that shifted significantly in their elution time, or disappeared (presumably eluting in the void volume), represented tryptic peptides that bound to heparin. Chondroitin sulfate was used as a control to account for non-specific ionic effects of heparin on elution of the peptides.



[Sasisekharan , 1991]

Figure 2.15: Heparinase I incubated with heparin and analyzed on a reverse-phase column. Top chromatogram is control heparinase I and bottom chromatogram is of heparinase I in the presence of heparin.

Results from these experiments indicated that tryptic peptides td45 (aa 215-221) and td4 (aa 132-141) bound specifically to heparin; consistent with td45 being a part of CnBr-8 peptide, and td4 being a part of the active site of heparinase I (Sasisekharan et al., 1995). The combined heparin binding results from experiments with CnBr and tryptic digests of heparinase I pointed to the region of residues 195-221 as being directly involved in heparin binding. The region from 195-215 contained multiple lysines and hence would have been degraded to very short peptides (di- and tripeptides) by trypsin. It was not expected to show up on the tryptic digest chromatogram, and its potential heparin binding properties might therefore have gone undetected.

To investigate the heparin-binding properties of these residues, a peptide corresponding to the region 196-213 was synthesized as a peptide (HBP I) containing the sequence IFKKNIAHDKVEKKDKDG. Interestingly, HBP-I affected the product profile of heparinase I degradation of heparin (Venkataraman, 1992). Heparinase I depolymerization of heparin results in two disaccharides, three tetrasaccharides (**1-3**), and a hexasaccharide (Rice and Linhardt, 1989). In a concentration dependent manner, the addition of HBP-I to the reaction mixture caused the peak corresponding to tetrasaccharide **3** (ΔU_{2S} $H_{NS,6S}$ I_{2S} $H_{NS,6S}$) to disappear. A control peptide (maxadilan) with similar charge properties (and at the concentration ranges tested above) had no effect on the enzyme activity or on the oligosaccharide product profile. This demonstrates that HBP-I affects the selectivity of heparin degradation by heparinase I: tetrasaccharide **3**, but not tetrasaccharides **1** or **2**, is degraded to a large extent in the presence of HBP-I. It must be pointed out that HBP-I does not degrade heparin or heparin oligosaccharides. The region 196-213 was determined to be the primary heparin binding site of heparinase I.

2.6.3.3 Role of calcium: in heparinase I activity and heparin binding

A region of the heparin binding site of heparinase I (residues 206-213) was found to be homologous to the calcium binding loop of the 'EF-hand' structural domain. Hence, the role of calcium in heparinase I activity was investigated. Heparinase I samples were extensively desalted using Centricon P-30 microconcentrator to remove residual calcium from the hydroxyapatite step during the enzyme purification (Venkataraman, 1992). The relative heparinase I activity as a function of

added calcium concentration was studied and the enzyme activity increases with the addition of calcium upto 5 mM. The results suggested a possible co-factor role for calcium in heparinase I activity.

In order to further investigate the calcium based heparinase I activation, the role of calcium in heparin binding to heparinase I was studied using heparin affinity chromatography (Sasisekharan et al., 1996). Heparin affinity chromatography of heparinase I was carried out in the presence and absence of calcium. Heparinase I bound to heparin-POROS in the absence of calcium, and the bound enzyme could be eluted at a salt concentration of about 200 mM. When the affinity separation was carried out in the presence of calcium, heparinase I eluted in the void volume since the enzyme cleaved heparin to which it bound; this was confirmed by the appearance of oligosaccharide products in the void volume.

2.6.3.4 Heparin binding constant of heparinase I

The ability of heparinase I to bind heparin, in the absence of calcium enabled the determination of a heparin binding constant for heparinase I, using affinity co-electrophoresis (ACE) (Sasisekharan, 1991). The technique measures the extent of binding based on the retardation of heparin when electrophoresed in the presence of heparinase I embedded in an agarose gel (Lee and Lander, 1991). ACE was carried out in the presence or absence of IAA to determine the importance of the active site cysteine-135 in the binding of heparin to heparinase I (Sasisekharan, 1995), and in the absence of calcium to prevent heparin degradation. There was no difference in the retardation of heparin for the IAA modified heparinase I when compared to the unmodified heparinase I, indicating that blocking the active site cysteine does not alter heparin binding. The binding constant for heparinase I binding to heparin (calculated from a Scatchard plot) was 60 nM.

2.6.3.5 Preliminary work on the heparin binding site

The above study (Venkataraman, 1992) identified the heparin binding site as the region between amino acids 196-213. As pointed out earlier this region contains two basic clusters satisfying the heparin binding consensus sequence. Further, as discussed in section 2.6, cysteine-135 of heparinase I, in a positively charged

environment, is catalytically active and it was hypothesized that a heparin binding domain is in close proximity to cysteine-135. Some interesting questions that remained unanswered and that were addressed as part of this research were:

a) Is the heparin binding site in close proximity to the active site cysteine-135?
b) If so, do the basic residues in the heparin binding site provide the positively charged environment to cysteine-135 and activate it for catalysis?

Results from preliminary work

Proximity of heparin binding site to active site

To address the above questions, the following experiment was performed. Heparinase I derivitization by sulfhydryl specific reagent PCMB inactivated the enzyme due to selective modification of the active site cysteine or cysteine-135 (Sasisekharan et al., 1995). Tryptic digest of PCMB modified heparinase I was carried out to determine if PCMB (being a bulky negatively charged molecule) could cause steric impedance to trypsin in cleaving the basic residues (contained in the heparin binding domain) in the close vicinity of cysteine-135. Such a protection of the basic residues from proteolytic digestion could argue for the close proximity of these residues to cysteine-135.

Figure 2.16, compares the tryptic maps of control heparinase I and that of PCMB modified heparinase I. The PCMB-heparinase I tryptic map has the appearance of a new peak marked hbp. The sequence of the peptide was determined to be: (K)NIAHDKVEKK. The results suggest that PCMB labeled cysteine-135 protects the lysine rich peptide (contained in the heparin binding site) from trypsin cleavage, when compared to a control digest where this peptide is not observed. This result supports the notion that a heparin binding site is in close proximity to the active site cysteine-135. Further, it raises the question as to whether, the basic residues (of the heparin binding sequence) provide for the observed positively charged environment around cysteine-135. This was addressed later and confirmed through extensive site-directed mutagenesis experiments (chapter 5).

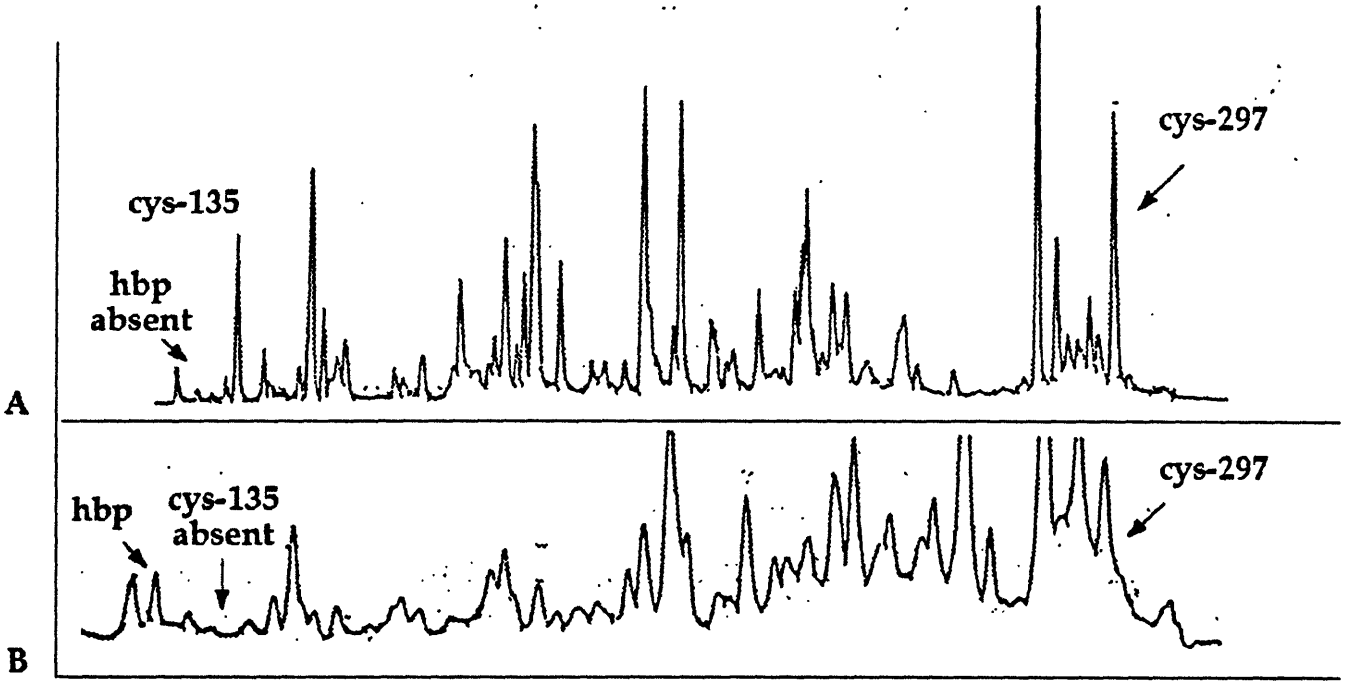


Figure 2.16: a) Tryptic map of heparinase I (control).
b) Tryptic map of PCMB labeled heparinase I. This has the appearance of a new peak marked hbp.

Discussion-the heparin binding site of heparinase I

For enzymatic reactions the interplay between binding and catalysis determines both, specificity, and rate of reaction. This interplay is especially interesting for eliminases/lyases that degrade complex acidic polysaccharides like pectin and heparin which are highly heterogeneous polymers. The results from the previous studies (Sasisekharan, 1991; Venkataraman, 1992) and the preliminary results from this research have served to identify a region of heparinase I involved in catalysis (cysteine-135) and the heparin binding site of heparinase I (amino acid residues 196 to 221), which has been shown to be in proximity to the active site cysteine-135. The primary heparin binding site in heparinase I belongs to a basic region of heparinase I, with two Cardin-Weintraub heparin binding consensus sequences and a calcium coordination motif.

Another interesting result from the previous study (Venkataraman, 1992) is that heparin binds to td4 (containing cysteine-135). It is possible that in native heparinase I, protein folding perhaps brings the lysine and/or arginine residues contained in td4 close to the heparin binding sequence, and all these residues together constitute a heparin binding domain in heparinase I. The presence of such a domain is supported by the PCMB labeling data presented in this study, where PCMB labeled heparinase I protects the lysine residues in the heparin binding sequence from tryptic cleavage.

Role of calcium

It was concluded (Venkataraman, 1992) that calcium is not required for heparin binding to heparinase I, but is important for catalysis. From equilibrium dialysis, it was found that calcium binds to heparin at a stoichiometry of about 1 calcium ion per tetrasaccharide (Sasisekharan et al., 1996), and fiber diffraction studies show that calcium causes a significant change in the helical structure of heparin, heparan sulfate, and other glycosaminoglycans (Nieduszynski, 1985). Acidic polysaccharides from plants and algae (polyguluronate and polygalacturonate) bind calcium effectively, which cause conformational changes in the polysaccharides into a more compact form (Perlin, 1975; Rees et al., 1982). This suggests that calcium activates heparinase I by binding to specific regions of the substrate, altering the substrate conformation perhaps by orienting the appropriate functional group(s) of the sugar

unit to the active site of the enzyme and making it more accessible to cleavage (Sasisekharan et al., 1996).

Interestingly pectate lyase, an enzyme from the same family of acidic polysaccharide eliminases as heparinase, is activated by calcium and the recently solved crystal structures of three pectate lyases showed that the substrate binding cleft was embedded with one or more calcium ions (at the calcium binding loop of EF-hands), indicating a direct interaction between calcium and the enzyme (Sasisekharan et al., 1996).

In heparinase I the calcium binding sequence is contained in the heparin binding site, which appears to be close to the active site cysteine-135 in the correctly folded active protein. Thus, the effect of calcium on heparinase I and other acidic polysaccharide lyases, is possibly mediated through a conformational change of the substrate upon calcium binding, which is further stabilized by direct interaction between enzyme and calcium at the active site.

The heparin binding domain of heparinase I

Specific interactions between the heparin binding domain of heparinase I and unique heparin sequences might determine the catalytic specificity of heparinases. The three heparinases from *F. heparinum* are specific for different linkages in the heparin chain, and heparinase I is much less active in cleaving tetrasaccharides than cleaving full heparin (Rice and Linhardt, 1989). This argues that heparin binding to heparinase I involves at least a pentasaccharide or longer region of heparin, similar to the heparin binding regions of bFGF and AT-III which have rather large (20Å or more) binding regions to accommodate close contacts (Sasisekharan et al., 1996).

In summary, from the preliminary results described in this work and from the earlier studies (Sasisekharan, 1991; Venkataraman, 1992), the following picture is emerging: cysteine-135 is catalytically active, but is not a determinant for heparin binding, since chemically derivatizing it did not affect heparinase I binding to heparin. The heparin binding site (residue 196-221) drives enzymatic selectivity in terms of substrate size. This site also contains the calcium co-ordination site which bridges heparin to heparinase I through calcium, and perhaps orients the functional group(s) of the

uronate to the active site region. The heparin binding domain in heparinase I perhaps provides the necessary charge complementarity for very specific heparin binding on the one hand, while on the other it serves for the active site environment which plays a key role in biasing the active site reactivity. These aspects become clear in the subsequent chapters.

2.7 Motivation, Mission and Objectives

Motivation

Immobilized enzymes have been used in various applications such as industrial processing, chemical purification, therapeutic waste processing, and pharmaceutical and waste processing (Bernstein et al., 1987). In most cases the costs of the enzymes are high, and it is essential to optimize immobilization efficiency and minimize enzyme waste. Enzyme immobilization efficiency depends on the amount of enzyme initially added that is bound to the support material and the fraction of the bound enzyme that is catalytically active.

The amount of enzyme bound could be enhanced by improving coupling chemistries. In addition, selecting chemistries that involve coupling of amino acid residues that form the catalytic domain of the enzyme would result in poor immobilization efficiencies. It becomes very important to select a proper coupling chemistry that would enable enhanced immobilization by improved activity retained.

Hence understanding the "business domain" of the enzyme, comprising the catalytic and substrate binding domains, the mechanism of catalysis and the surface properties of the enzyme would greatly aid the immobilization design process. Improved understanding of the heparin-heparinase interaction would enable one to design improved enzyme systems for better immobilization and also rationally engineer improved properties such as stability, altered specificities etc. In addition, modified heparinases could be a useful tool in elucidating heparin structure information and in heparin sequencing, as well as in generating heparin fragments with potential therapeutic value.

Overall Mission

From the above discussion it is clear that heparinases have significant therapeutic applications—as agents for heparin neutralization after extra-corporeal therapies and as potent inhibitors of angiogenesis or new blood vessel formation. These two applications have an enormous impact on the development of several new medical technologies. One such example is the development of an immobilized heparinase reactor to be used for heparin neutralization during acute hemodialysis. The overall mission of our group has been to understand, develop and extend these applications to enable heparinase to be an important component of the clinical armamentarium.

Objectives

The above mission provides a strong motivation towards understanding the chemistry governing heparin-heparinase interactions. Identifying domains of the enzyme involved in catalysis and heparin binding, the mechanism of catalysis and surface properties of the enzyme would greatly aid several issues relating to heparinase applications including immobilization design processes as well as enhancing enzyme stability etc. The overall objectives of this research are :

1) To develop
a) An experimental approach that enables one to rapidly engineer mutations in heparinase through protein engineering
b) An assay system to monitor both qualitatively and quantitatively, the effect of these changes.

2) Use the above experimental approach to understand the enzyme's catalytic mechanism and relationship between the surface properties (structure) and function (activity). Understanding the structure-activity relationships of heparinase I would eventually provide for a framework towards the development of an approach in creating "designer enzymes" for specific applications.

More specifically, the above studies raised several interesting questions such as those mentioned below and this work addresses those questions specifically.

- What is the role of the positively charged environment around cysteine-135?
- Are other residues involved in enzyme activity?
- Is the heparin binding site in close proximity to active site cysteine-135?
- What is the role of heparin binding site in heparinase I activity?
- What is the role of the positive charge in the heparin binding site, in heparinase I activity?
- What is the catalytic mechanism of heparinase I?
- How can this information be used to improve some of the properties of heparinase I viz. stability, which would greatly enhance heparinase efficiency during deheparinization procedures such as hemodialysis, and catalytic efficiency?

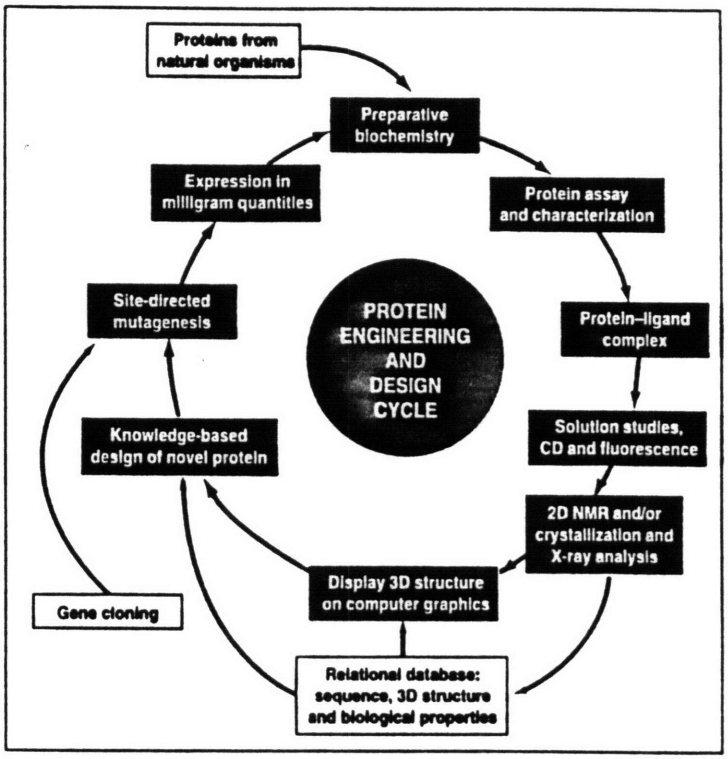
These were some of the questions addressed as part of this research and are described in detail in chapters 4, 5, and 6. The next section describes the overall experimental strategy developed to address the above objectives.

2.8 Experimental Strategy—A Design Cycle

Rational approaches to engineering proteins are usually organized as design cycles. The ability to modify proteins using site-directed mutagenesis has generated interest in the development of 'protein engineering design cycles' **(Figure 2.17)**. One could enter the cycle by using the protein from a natural organism. This protein is then extensively characterized biologically and biochemically. The three-dimensional structure is defined by x-ray crystallography or NMR analysis and a computer graphics display of the structure is usually obtained.

The next step in the cycle is to test the design hypothesis by site-directed mutagenesis, followed by expression and characterization of the mutant protein. The development of design cycles is highly multi-disciplinary involving molecular biologists, biochemists, crystallographers etc. and these various disciplines need to be integrated in a systematic manner (Blundell, 1994).

To address the objectives in this research, a slightly modified version of the above design cycle was developed **(Figure 2.18)**. The cycle was entered by gene cloning and using the recombinant protein. Site-directed mutagenesis was used to not only test



[Blundell, T. I. B Tech., 12, 145, 1994]

Figure 2.17: The Protein Engineering design cycle

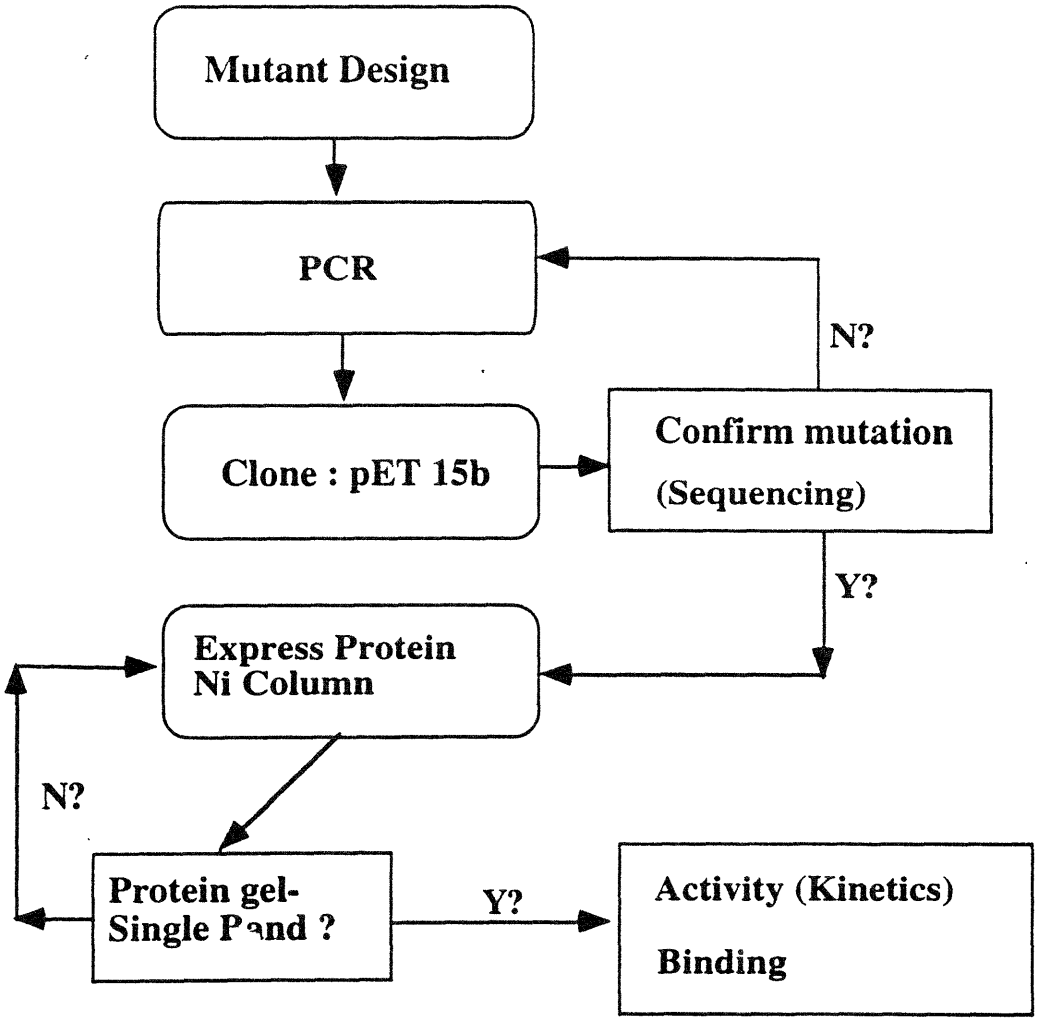


Figure 2.18: Overall methodology — the design cycle.

hypotheses based on other approaches (such as chemical modification of the native protein) but also to generate structural information. The mutant protein was expressed and purified using a one step Ni-affinity chromatography procedure (see chapter 3) and extensively characterized. This information was used in developing a structural model. The model was subsequently tested (through steric and functional mutations) using site-directed mutagenesis techniques.

Chapter 3

Materials & Methods

This chapter describes the experimental protocols and procedures for all the different experiments. The sources of various chemicals and reagents are listed.

Chemicals and Materials

Heparin, from porcine intestinal mucosa with an average molecular weight of 12 kDa and activity of 157 USP units/mg was from Hepar (Franklin, OH). Urea, dithiothreitol[1] (DTT), IAM, TFA, PCMB, and acetonitrile were from Allied Chemicals (Deerfield, IL). IAA, DEPC, [^{14}C]DEPC (specific activity of 2.6 mCi/mmol), hydroxylamine hydrochloride, chondroitin sulfate, heparan sulfate, guanidine hydrochloride, ammonium bicarbonate, guanidine thiocyanate, MOPS, Tris were from Sigma Chemicals (St. Louis, MO). PCMB, and glacial acetic acid were from Aldrich, WI. Other chemicals were from Mallinckrodit Inc. (Chesterfield, MO). [^{3}H]iodoacetic acid was from Amersham, IL. Trypsin was from Boehringer Mannheim, CT. Molecular weight standards were obtained from GIBCO BRL/Life Technologies, MD. *E. coli* BL21(DE3) host was from Novagen, WI. Molecular biology reagents and their sources are listed in the appropriate sections below.

3.1 Biochemical experimental techniques

3.1.1 Chemical Modification with DEPC

Inactivation with DEPC

Heparinase I (25 mg/ml), in 100 mM MOPS buffer (pH 7.0), was incubated with DEPC, a histidine specific reagent, at DEPC concentrations of 0.01 mM, 0.02 mM, 0.05 mM, and 0.1 mM at room temperature. DEPC solutions were freshly diluted in ethanol just before use, from a 6M stock solution. The amount of ethanol in the reaction mixture did not exceed 2 % (v/v) and the control reaction mixtures contained the same amount of ethanol (instead of DEPC). At fixed time intervals, 25-ml aliquots were used for activity assays. The time course of inactivation was obtained by determination of the enzymatic activity retained after successive incubation intervals. Pseudo-first order rate constants k, for inactivation by DEPC, were determined by calculating the slopes from a plot of log % fractional activity retained versus time.

[1] Abbreviations: Dithiothreitol (DTT), iodoacetamide (IAM), iodoacetic acid (IAA), diethylpyro carbonate (DEPC), 3-(N-morpholino)-propanesulfonic acid (MOPS), tris-(hydroxymethyl)-aminomethane (Tris), p-chloromercuribenzoate (PCMB), trifluroacetic acid (TFA), bovine serum albumin (BSA), cyanogen bromide (CnBr), ethylene di-amine tetra acetic acid (EDTA), sodiumdodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE), reverse-phase high pressure liquid chromatography (RPHPLC), recombinant heparinase I (*r*-heparinase I),

Reactivation of DEPC modified enzyme with hydroxylamine

Heparinase I (25 mg/ml) was incubated with 7.5 mM DEPC for 1 min and an aliquot was withdrawn for measuring the fractional activity retained. Hydroxylamine (in 100 mM MOPS buffer, pH 7.0) was then immediately added to the reaction mixture to a final concentration of 300 mM, and the mixture was incubated at 4° C for about 4 hr. Aliquots were taken every hr and assayed for activity. A control reaction was also performed without DEPC but with the same amount of hydroxylamine to account for the effect of hydroxylamine on the activity of heparinase I.

Substrate protection of heparinase I against modification by DEPC

To investigate the ability of heparin to protect heparinase I against modification by DEPC, heparinase I (25 mg/ml) was incubated with 2 mg/ml heparin (final concentration 0.5 mg/ml; with and without 5 mM calcium acetate) for a period of 10 mins and 1 hr. The concentration of DEPC was fixed at 0.05 mM in all the experiments. 25-ml aliquots of the sample were withdrawn every 3 min, for a total time of about 15 min, for enzyme activity determination.

Quantification of modified histidine residues

Quantification of DEPC modified spectra was determined by difference spectra. Two cuvettes contained heparinase I (0.02 mg/ml) in 100 mM MOPS buffer, pH 7.0, incubated at room temperature. At time zero, DEPC was added to the sample cuvette to a final concentration of 0.01 mM and an equal volume of ethanol was added to the reference cuvette. The change in the absorbance at 240 nm was monitored every 5 min for 30 min. Enzyme activities of the sample and control were also measured every 5 min.

Labeling with [^{14}C] DEPC

Heparinase I was modified using [^{14}C] DEPC. About 40 mg (1 nanomole) of heparinase I was radiolabeled with 0.15 mM [^{14}C] DEPC for about 4 min. The reaction was stopped by adding 40 mM imidazole (pH 7.0) and immediately loaded onto a RPHPLC column to get rid of the excess radiolabel. The labeled heparinase I peak was collected and concentrated using a speed vac to get rid of acetonitrile. The sample was counted for radioactive incorporation and the number of moles of

modified residues per mole of enzyme was calculated from the specific activity. The concentrated, labeled enzyme was used for a tryptic digest.

Inactivation with DEPC as a function of pH

Heparinase I (25 mg/ml) was incubated with 0.025 mM DEPC at room temperature and in buffers with different pH values. 50 mM succinic acid buffer was used for studies in the pH range of 3.5-6.5 and 50 mM bis-tris propane buffer was used for the assays done in the pH range of 6.5-9.0. As before, the amount of ethanol in the reaction mixture did not exceed 2 % (v/v) and the control reaction mixtures contained the same amount of ethanol (instead of DEPC). At fixed time intervals, 25-ml aliquots were used for activity assays. The time course of inactivation was obtained by determination of the enzymatic activity retained after successive incubation intervals. Pseudo-first order rate constants k, for inactivation by DEPC, were determined by determining the slopes from a plot of log % fractional activity retained versus time. These rate constants were then plotted as a function of pH and the data were fitted to the appropriate equation (chapter 6) to determine the pKa of the modified residue.

3.1.2 Chemical Modification with IAA

Enzyme inactivation as a function of pH

Heparinase I was alkylated with IAA by incubation of the enzyme (0.1 mg/ml) with 0.1-10 mM IAA, a thiol modifying reagent, at room temperature and in buffers of different pH. 50 mM succinic acid buffer was used for studies in the pH range of 3.5-6.5. Aliquots were withdrawn for activity assays and pseudo-first order rate constants for inactivation by IAA were calculated as described above. The pKa of the cysteine residue was then calculated by plotting the rate constants as a function of pH and fitting the data to the appropriate equation (chapter 6).

PCMB inactivation of heparinase I

PCMB was prepared and used as described (Sasisekharan et al., 1995). Briefly, to 0.01 mg/ml enzyme in 0.1 M MOPS, 2.5-100 μM PCMB was added. The mixture was incubated at 4°C for 4 hr. The time course of inactivation was obtained by determination of the enzymatic activity retained after successive incubation intervals. The reversibility of the reaction was established by incubation of the

PCMB treated enzyme with 50 mM DTT. PCMB labeling experiments determining heparinase I inactivation were also carried out in the presence of different salt concentrations (50 mM, 100 mM and 200 mM NaCl in 0.1 M MOPS solution).

3.1.3 Tryptic Digest and Protein Sequence Analyses

Tryptic digest of the samples were performed as described previously (Sasisekharan et al., 1993). In the case of [^{14}C]DEPC labeling experiments, one nanomole of [^{14}C]DEPC labeled heparinase I was digested by trypsin. Tryptic peptides were separated by RPHPLC and monitored at 210 and 277 nm. The peaks were collected and counted for [^{14}C] incorporation on a Beckman scintillation counter. The [^{14}C] incorporated peptide peaks were sequenced. Peptides were sequenced using an Applied Biosystems Sequencer Model 477, with an on-line Model 120 PTH Amino Acid Analyzer (Biopolymers Laboratory, MIT).

3.2 Mutagenesis and protein purification techniques

3.2.1 Site-directed Mutagenesis Techniques

The mutations were introduced using the PCR method of Higuchi (Higuchi, 1990). Two primers in both directions, that incorporated the mutation, were designed for each mutation . **Table 3.1** lists all the different primers that were synthesized for the various mutations.

Table 3.1 Primer sequences for mutations

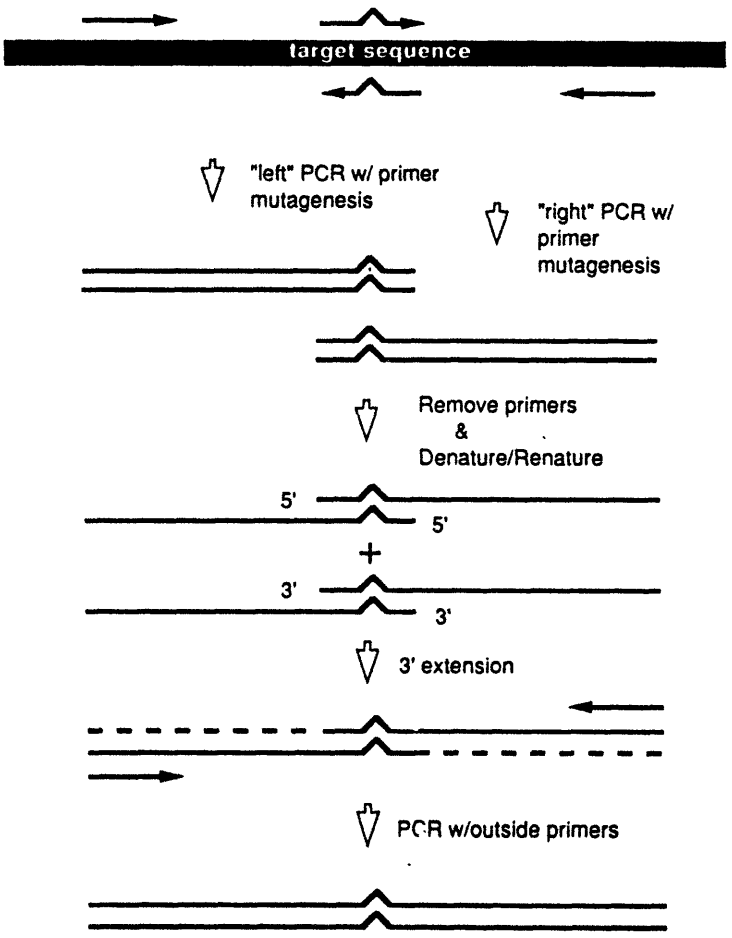
Mutation	5′ Primer Sequence (5′--->3′)	3′ Primer Sequence (5′--->3′)
C135A	AAA GGG ATT GCT GAA CAG GGG	CCC CTG TTC AGC AAT CCC TTT
C135S	AAA GGG ATT TCT GAA CAG GGG	CCC CTG TTC AGA AAT CCC TTT
C135D	AAA GGG ATT GAT GAA CAG GGG	CCC CTG TTC ATC AAT CCC TTT
C135E	AAA GGG ATT GAA GAA CAG GGG	CCC CTG TTC AAC AAT CCC TTT
C135H	AAA GGG ATT CAT GAA CAG GGG	CCC CTG TTC ATG AAT CCC TTT
H203A	AAT ATC GCC GCT GAT AAA GTT	AAC TTT ATC AGC GGC GAT ATT
H203S	AAT ATC GCC TCT GAT AAA GTT	AAC TTT ATC AGA GGC GAT ATT
H203D	AAT ATC GCC GAT GAT AAA GTT	AAC TTT ATC ATC GGC GAT ATT
H203C	AAT ATC GCC TGT GAT AAA GTT	AAC TTT ATC ACA GGC GAT ATT
K198AK199A	ATC TTC GCA GCA AAT ATC GCC	GGC GAT ATT TGC TGC GAA GAT
K198DK199D	ATC TTC GAC GAC AAT ATC GCC	GGC GAT ATT GTC GTC GAA GAT
K198RK199R	ATC TTC AGA AGA AAT ATC GCC	GGC GAT ATT TCT TCT GAA GAT
K208AK209A	GTT GAA GCA GCA GAT AAG GAC	GTC CTT ATC TGC TGC TTC AAC
K208DK209D	GTT GAA GAC GAC GAT AAG GAC	GTC CTT ATC GTC GTC TTC AAC
K208RK209R	GTT GAA AGA AGA GAT AAG GAC	GTC CTT ATC TCT TCT TTC AAC
N200A	TTC AAA AAA GCT ATC GCC CAT	ATG GGC GAT AGC TTT TTT GAA
N200K	TTC AAA AAA AAA ATC GCC CAT	ATG GGC GAT TTT TTT TTT GAA
K205A	GCC CAT GAT GCA GTT GAA AAA	TTT TTC AAC TGC ATC ATG GGC
K205Y	GCC CAT GAT TAC GTT GAA AAA	TTT TTC AAC GTA ATC ATG GGC
E207A	GAT AAA GTT GCA AAA AAA GAT	ATC TTT TTT TGC AAC TTT ATC
D210A	GAA AAA AAA GCT AAG GAC GGA	TCC GTC CTT AGC TTT TTT TTC
D212A	AAA GAT AAG GCC GGA AAA ATT	AAT TTT TCC GGC CTT ATC TTT
G213A	GAT AAG GAC GCA AAA ATT ACT	AGT AAT TTT TGC GTC CTT ATC
K214A	AAG GAC GGA GCA ATT ACT TAT	ATA AGT AAT TGC TCC GTC CTT
K333AK334A	AAG AAT GCG GCA CCA CAA AAA	TTT TTG TGG TGC CGC ATT CTT
H129A	ACC GTT TAT GCT TAC GGC AAA	TTT GCC GTA AGC ATA AAC GGT
H165A	GCC CAA TGG GCT GGT GCA CCC	GGG TGC ACC AGC CCA TTG GGC
H339A	CAA AAA GCG GCT ATC GTA AAC	GTT TAC GAT AGC CGC TTT TTG
DEL HBD II	ACT TAT ACC ATC GTA AAC CAG	CTG GTT TAC GAT GGT ATA AGT
K132A	CAT TAC GGC GCA GGG ATT TGT	ACA AAT CCC TGC GCC GTA ATC

K141A	GGG AGC TCC GCC AGC TAT ACC	GGT ATA GCT GGC GGA GCT CCC
C297A	CCC TAA AGA TGC CTG GAT TAC	GTA ATC CAG GCA TCT TTA GGG
C297S	CCC TAA AGA TCC CTG GAT TAC	GTA ATC CAG GGA TCT TTA GGG

The method of introducing the mutation involved the following steps. Two primary PCR reactions (12 cycles) were set up using the -L/pET 15b construct as a template and the following combinations of primers: T7 promoter primer (from Novagen, WI) and the 3' mutant primer; T7 terminator primer and the 5' mutant primer. Three samples of each reaction were pooled and run on a 0.7% low melting point (LMP) agarose gel. A lower number of PCR cycles was used to reduce the possibility of random mutations.

The primary PCR products were isolated from the gel, melted at 65° C and 2 ml of each of the molten fragments were used as a template for the next (secondary) PCR reaction, using T7 promoter and T7 terminator primers. This amplified the entire gene product with the incorporated mutation **(Figure 3.1)**. The PCR fragment was again isolated from a LMP gel and used directly for a ligation reaction. The PCR product was directly ligated into T-vector, which was prepared as follows. *Taq* polymerase adds overhanging base 'A' to amplified PCR fragments. This is taken advantage of in ligation reactions with T-vector. pBluescript vector (Stratagene, CA) is linearized using the restriction enzyme (Eco RV) and isolated from a LMP gel. The fragment is melted and separated from the agarose using gelase (Epicenter technologies). This is incubated with dTTP at 72° C to add overhanging 'T' ends as described (Marchuk et al., 1991).

The ligation reactions (overnight at 16° C) were transformed into *E. coli* DH_5a competent cells and plated on Xgal/IPTG plates to screen for blue/white colonies. The colonies were screened using PCR and colony screening techniques. The colony screening involved picking a part of the colony into a lysis mix (containing lysozyme, RNAse, and bromophenol blue dye) for 5 min and extracting the DNA using phenol/chloroform. The mixture was spun down and the aqueous layer loaded onto an agarose gel directly. The presence of an insert was determined by the relative migration of the colony with respect to a control plasmid without the insert. The positive clones with the insert were further screened by PCR and the plasmid



[Higuchi, 1990.]

Figure 3.1: *PCR based mutagenesis strategy.* Combining two PCR products with overlapping sequence into one longer product. The two overlapping primers are shown containing a mismatch to incorporate the mutation.

was isolated. This was then sequenced to verify the mutation. The insert was digested from the plasmid using *Nde* I and *Bam* HI restriction enzymes and isolated from a LMP gel. This was cloned into pET-15b, which is the expression vector and transformed into *E. coli* DH_5a competent cells. The plasmid was again isolated and used for expression and purification whenever desired.

3.2.2 Expression, Isolation and Purification of Mutant r-heparinase I in *E. coli*

The constructs were transformed in BL21(DE3) (Novagen) and 5 ml overnight cultures were grown for inoculation. The overnight cultures containing the *r*-heparinase I (-L and mutant) genes in pET-15b were diluted and grown until an OD_{600} of about 0.5 and induced with 5 mM IPTG and harvested. The cell pellet was resuspended in 1/20 th volume binding buffer (20 mM Tris, 500 mM NaCl, 5 mM Imidazole). The resuspended culture is placed in an ice bath, sonicated for 2 min using a Branson sonicator (Model no. 450, power 3, 50% pulse; Branson, CT) and centrifuged at 4°C and 15,000 g for 30 min. The supernatant was assayed for activity and purified by Ni^{2+} affinity chromatography using sepharose 6B Fast Flow resin covalently linked to nitrilotriacetic acid (Novagen, WI). The resin was charged with 5 column volumes 200 mM $NiSO_4$ and equilibrated with 3 column volumes binding buffer. Then, 6-10 ml sample was applied followed by 10 column volumes binding buffer, 6 column volumes 15% elution buffer (20 mM Tris, 500 mM NaCl, 200 mM Imidazole) and 4-5 column volumes 100% elution buffer. The eluted enzyme was then used for kinetic analysis and protein characterization. Purity of the *r*-heparinase was determined by SDS-PAGE analysis of Laemmli (Laemmli, 1970) with a Mini Protean II electrophoresis apparatus (BioRad, CA). Visualization of proteins in gels (12% gels) was accomplished with 0.1% Coomassie Blue stain or using Silver Stain (BioRad, CA).

To cleave the poly-histidine tag, which however did not affect the enzyme activity, heparinase I recovered in 4-5 ml of the 100% elution step, was desalted on two PD10 columns (BioRad, CA) and incubated overnight at 4°C with 0.5 units thrombin (Novagen, WI). After stripping (20 mM Tris, 500 mM NaCl, 100 mM EDTA) and charging the metal chelate column, the cleaved heparinase I was applied and collected in the flow through fraction.

3.2.3 Heparinase I Activity Assays

UV 232 nm assay

For measuring heparinase I activity during purification, a known amount of the purified enzyme was added to 2 mg/ml heparin solution in 100 mM MOPS buffer at pH 7.0, containing 5 mM calcium acetate. The formation of double bonds was monitored with a corresponding increase in the 232 nm absorption spectra and the slope of the line was calculated using zero order kinetics, since the substrate is in saturating amounts.

Kinetic analysis

For determining the kinetic parameters of the recombinant heparinase I and mutant enzymes, equivalent amounts of proteins were used in activity assay. The UV 232 nm online assay was conducted at different substrate concentrations (0.1 mg/ml heparin to 2 mg/ml heparin, conditions as mentioned above). The reaction velocity was determined from the first 2-3 min of the reaction and the kinetic parameters (k_{cat} and K_m) were determined for each of the mutants assuming Michaelis-Menten kinetics.

pH profile

To determine the pH profile the UV 232 nm assay was performed using approximately 0.35 mg heparinase I for each assay in the pH range 4-9.5. 50 mM succinic acid buffers were used for the pH range 4-6.5 and 50 mM Bis-tris propane buffers were used for the pH range 6.5-9.5. The data were fitted to the following equation using non-linear least square analysis.

$$k_{cat} = C/(1+H/Ka)(1+Kb/H)$$

where C is the pH independent value of k_{cat} obtained experimentally, Ka and Kb are the ionization constants of the two residues.

HPLC of heparin oligosaccharides

Heparin (2 mg/ml) was incubated with heparinase I; *r*-heparinase I and mutant enzymes in 5 mM calcium acetate, 100 mM MOPS buffer, pH 7.0, for 18 h. The reaction was subjected to anion-exchange HPLC (POROS column on a BIOCAD^{TM}

workstation, PerSeptive Biosystems, Cambridge, MA) to resolve the oligosaccharide products, as described (Sasisekharan et al., 1993).

3.3 *Cloning of heparinases II and III*

3.3.1 Heparinase Purification and Characterization

Heparinases II and III were purified as described (Lohse and Linhardt, 1992; Sasisekharan et al., 1993). An additional purification step was carried out by RPHPLC (in a HP 1090 from Hewlett Packard, CA, with diode array detection) using a Vydac C_{18} reverse-phase column, with a gradient of 0 to 80% acetonitrile in 0.1% TFA for 60 min. Protein was monitored at 210 and 277 nm. The separated heparinase peaks were collected and lyophilized (Virtis Freeze Mobil Model 12, Virtis Inc., NY). Protein concentration was determined using Micro BCA reagent (Pierce Inc., IL) relative to a BSA standard. The single peak was used in the tryptic digest. Mass spectrometry was performed on the preparation to determine the purity, homogeneity and molecular masses of heparinases II and III (Sasisekharan et al., 1994). Amino acid composition analysis was performed on an amino acid analyzer (Model 420, Applied Biosystems, CA) in Biopolymers Laboratory, Center for Cancer Research, MIT.

3.3.2 Tryptic Digest and Protein Sequence Analyses

One nanomole of purified heparinases II and III were denatured in 50 µl (8 M urea, 0.4 M ammonium carbonate) solution and reduced with 5 mM DTT at 50 °C, cooled to room temperature, and alkylated with 10 mM iodoacetamide for 15 min in the dark. The reaction was quenched with water by bringing the total reaction volume to 200 µl. To the above reaction, 4% w/w of trypsin was added and the digestion was carried out at 37 °C for 24 hr. The proteolytic reaction was terminated by heating the sample at 65 °C for 2 min. The digest was separated using a gradient reverse-phase HPLC (0 to 80% acetonitrile in 0.1% TFA for 120 min). Tryptic peptides were monitored at 210 and 277 nm and collected. Based on the homogeneity of the peptide peaks, a few peaks were sequenced using an Applied Biosystems Sequencer Model 477, with an on-line Model 120 PTH Amino Acid Analyzer (Biopolymers

Laboratory, MIT). N-terminal sequence analysis was performed using undigested wild-type heparinases II, and III.

3.3.3 Genomic DNA Isolation, Library preparation and Southern Blotting

The *F. heparinum* genomic DNA was isolated as described earlier (Sasisekharan et al., 1993). A custom λ ZAP II library was made (Stratagene, La Jolla, CA) comprising 7 x 10^6 independent clones. The insert size range was 6-10 kb and the estimated amplified titer was about 1 x 10^6 pfu/ml. Genomic DNA (10 μg) was prepared for Southern blotting by digestion with *Eco*RI, *Bam*HI, Pst I, Bgl II and *Hin*dIII, individually or in combination, for 2 hr, and separated on a 0.8% agarose gel for 16 hr (60 V). The gel was transferred onto a nylon membrane (Hybond-ECL, Amersham, IL) by capillary action. The probe was prepared using the ECL Labeling kit as per the manufacturer's recommendations (Amersham, IL). The membrane was blocked, hybridized, probed with the PCR products, and detected using the ECL kit as per the manufacturer's recommendations (Amersham, IL).

3.3.4 Amplification of PCR Probe

PCR was performed using heparinase II and III primers in a 25 mL reaction volume containing 1 mg *F. heparinum* genomic DNA as the template, PCR buffer (Perkin Elmer, CA), four dNTPS at 200 mM, 0.5 mM primers, 2.5 units Taq polymerase (Cetus), and 25 ml mineral oil. The samples were amplified on a thermal cycler (Perkin Elmer, CA) essentially as described (Sasisekharan et al., 1993). The PCR product was isolated from a 0.7% low-melting agarose gel and used directly for ligation.

3.3.5 Screening a λ ZAP II *F. heparinum* Library

The amplified library stock was diluted about 1000 fold in SM buffer (Sambrook, 1989) and the phage were titered to give about 20-30,000 plaques on a single 130 mm plate. In all about 1 x 10^6 primary plaques were screened. The plates were grown for 6-8 hr at 37°C and then incubated at 4°C for an hr. The phage were then transferred on to nylon membranes in duplicates using standard protocols (Sambrook, 1989). The membrane was UV crosslinked and placed overnight in hybridization solution

(Sambrook, 1989) at 42°C with shaking. The PCR product obtained above was concentrated using centricon tubes (Amicon, MA) with a molecular weight cutoff of 100 kD, to purify the product. This was then labeled using the random primed DNA labeling kit (US Biochemicals, MD) and [a-^{32} P] dCTP (3000 Ci/mmol; 1 Ci = 37 GBq; Amersham, IL). Following labeling of the probe, free nucleotides were removed with a Sephadex G-50 column (Nick column, Pharmacia, NJ) and the probe was collected in a final volume of 100 ml in water . This was subsequently added to the hybridization solution at a probe concentration of about 10 ^{6} cpm/ml of hybridization solution. About 10-12 membranes were probed in 20 mls of solution. Each of the positive primary plaques were purified through secondary and tertiary rounds of screening. The single plaque isolated from tertiary screening was stored in SM buffer and used for *in vivo* excision process to rescue the pBluescript phagemid from the l Zap II vector as per the manufacturer's protocol (Stratagene, CA). A single colony was picked to isolate the plasmid for sequencing.

3.3.6 DNA Sequencing

DNA sequencing, in both orientations, was performed using the dideoxy-chain termination method with a S^{35} labeled dATP (NEN, MA) and Sequanase™ (US Biochemical, MD) as described by the manufacturer. The primers used in generating the PCR product, were used to determine the initial sequence. Using the DNA sequence generated, specific primers were designed and synthesized (Biopolymers laboratory, MIT). Sequence data were analyzed using the Mac Vector software (IBI, CA), BLAST and FASTA programs in the DNA sequence database.

3.3.7 Expression, Isolation and Purification of Recombinant Heparinases II and III in *E. coli*

The heparinase III gene was cloned into the *Nde* I/*Bam*H I sites of plasmid pET-15b (Novagen, WI), as described earlier for heparinase I (Sasisekharan et al., 1995) while the heparinase II gene was cloned into the *Nde* I/*Sac*I sites of pET-28a. PCR amplification of the heparinase III leader and coding sequences from the original clone in pBluescript (and genomic DNA in the case of heparinase II) used 5′ and 3′ primers (Gibco·BRL) that contained mismatches to introduce the restriction sites, respectively. The gene is located immediately 3′ to a His_6 tag to facilitate purification

by Ni^{2+}-affinity chromatography. The plasmid containing the heparinase genes were isolated and used to transform the host cell BL21(DE3) (Novagen, WI).

For protein purification, a 100 mL LB culture containing 100 mg/mL ampicillin (kanamycin for pET-28a) was grown at 37° C with vigorous aeration to OD_{600} = 1.0, induced with 2 mM IPTG, and grown for an additional 6 hr to yield 0.4 g of wet cell paste. All subsequent steps were done on ice or at 4° C. The cell pellet was suspended in 50 mM Tris·HCl, 2 mM EDTA (pH 8.0) and the cells were lysed by sonication. The lysate was centrifuged at 5000x*g* (4° C, 20 min.) and, after saving the supernatant which represented the soluble fraction, the pellet was washed with PBS. After solubilizing the washed pellet in 5 mL binding buffer (20 mM Tris·HCl/500 mM NaCl/6 M urea, pH 7.9) containing 5 mM imidazole, centrifugation was repeated to remove cellular debris. The crude, solubilized protein fraction was purified on a 1 mL column packed with His·Bind resin (Novagen) as follows. After loading the sample, the column was washed with 10 mL of binding buffer containing 5 mM imidazole followed by 6 mL of binding buffer containing 20 mM imidazole. Elution was accomplished by washing the column with 6 mL of binding buffer containing 1 M imidazole. Refolding of enzymes was performed by rapid 1:10 to 1:20 dilution into 20 mM Tris·HCl/500 mM NaCl/5 mM imidazole (pH 7.9), incubation on ice for 60 min, and concentration in Centriprep-30 cells (Amicon). Purity of the recombinant heparinases were determined by SDS-PAGE analysis. Enzymatic activity assays of the recombinant enzymes were carried out for all the preparations.

3.3.8 Gel Electrophoresis

Sodium dodecylsulfate polyacrylamide (12%) gel electrophoresis (SDS-PAGE) of heparinases II and III was carried out according to the method of Laemmli (1970) with a Mini Protean II electrophoresis apparatus (BioRad, CA). Visualization of proteins in gels was accomplished with 0.1% Coomassie Brilliant Blue stain (BioRad, CA) or using Silver Stain (BioRad, CA).

3.3.9 Heparinase Activity Assays

UV 232 nm assay

For measuring heparinase III activity during purification, 10-100 ml of sample was added to a cuvette with 1 ml of 1 mg/ml heparan sulfate in 100 mM MOPS, 5 mM CaOAc at pH 7.6 and 35 °C and the increase in absorbance at 232 nm was directly measured as a function of time. For heparinase II the substrate was 2 mg/ml heparin in 100 mM MOPS buffer (pH 7.3) and assays were done at 30° C.

HPLC of heparin oligosaccharides

Heparin (2 mg/ml) and heparan sulfate (1 mg/ml) in their respective buffers were degraded by native and recombinant heparinases II and III for 18 hr. The reaction was stopped by injecting the solution into a POROS Q/M (4.6 mm x 100 mm) anion-exchange column (PerSeptive BioSystems, MA) connected to a BIOCAD system (PerSeptive BioSystems, MA). Each run was for 10 min using a salt gradient of 0-2 M NaCl (in 10 min) in 5 mM Tris-HCl, pH 7.0, monitored at 232 nm. Anion-exchange HPLC resolved the oligosaccharide products.

Histidine residues play an important role in the activity of several enzymes. This chapter describes the mapping of a critical histidine residue (histidine-203) through extensive chemical modification and site-directed mutagenesis experiments. The mutant enzyme H203A was inactive, while alanine mutagenesis of the other histidines of heparinase I had no effect on activity. An important observation from this work is the finding that not only is histidine-203 contained in the heparin binding site, but also it is critical for catalysis. The determination of a catalytic role for histidine-203 provides the most compelling evidence for heparin binding domain being in close proximity to the active site cysteine-135. Thus, it is proposed that histidine-203 and cysteine-135 are adjacent to the scissile HI linkage during catalysis and form the catalytic domain of heparinase I.

Chapter 4

The Catalytic Domain of Heparinase I

4.1 Motivation

Chapter 1 discussed the need for investigating and understanding structure-function relationships of enzyme systems and the power of protein engineering in elucidating such information. Chapter 2 described the role of heparin like molecules and the significant implications of heparin-protein interactions in the extra-cellular matrix. The clinical applications heparinases and their use as tools in understanding heparin structure were also discussed. An interesting feature of the discussion was the unique substrate specificity of the three heparinases from *F. heparinum* and their ability to distinguish subtle differences in the chemical composition of heparin. However, the molecular features, of the enzyme, as well as the polysaccharide, governing this specificity are not well understood.

Heparinase I is a lyase that cleaves heparin via the elimination mechanism. There is very limited structure-function information or studies on enzyme systems in literature, involving enzymatic depolymerization through elimination (Linhardt et al., 1986; Ernst et al., 1995). Heparinase I is a model enzyme system for studying elimination reactions and hence it is important to determine the catalytic residues of heparinase I in order to understand the mechanism of heparin depolymerization and substrate specificity.

Chapter 2 described some preliminary structure-function work (a combination of chemical modification and site directed mutagenesis experiments) which demonstrated that a surface accessible cysteine (cysteine -135), contained in a positively charged environment, is essential for heparinase I activity (Sasisekharan et al., 1995). Further, the observation of a decrease in the rate of heparinase I inactivation by chemical modification of cysteine-135 in the presence of heparin, along with other experiments, led to the hypothesis that a heparin binding domain is in close proximity to cysteine-135. In addition, the primary heparin binding site of heparinase I, containing two heparin binding consensus sequences (Cardin and Weintraub, 1989) and a calcium coordination consensus motif (Kretsinger, 1980), was identified and mapped (Sasisekharan et al., 1996). Further, as pointed out in chapter 2, sulfhydryl selective labeling of cysteine-135 in heparinase I protected the

lysines of the heparin binding sequence from proteolytic cleavage, suggesting the close proximity of the heparin binding site to the active site.

To identify the catalytic domain in heparinase I it was necessary to investigate other amino acid residues in heparinase I which are essential for enzyme activity. It was proposed that polysaccharide lyases proceed via a base catalyzed mechanism, where one amino acid acts as a general base by abstracting a C5 proton from the uronate of the disaccharide repeat unit of the acidic polysaccharide and a different amino acid acts a proton donor to protonate the leaving group (Gacesa, 1987). Further, it was suggested that a histidine residue is involved in the activity of a hyaluronate lyase, possibly as an acid-base catalyst, thereby extracting a proton from the uronic acid and eliminating the glycosidic bond (Greiling et al., 1975). In addition, histidines participate in a large number of enzymatic reactions by a variety of mechanisms which include general acid/base catalysis, electrophilic catalysis and in the binding of substrate via hydrogen bonding and/or electrostatic interactions (Munier et al., 1992). Histidine can exert its catalytic function by either interacting directly with the substrate or being an intermediate in a two- or three- member proton shuttle system. Such a charge relay catalytic system was frequently seen in reactions catalyzed by hydrolases and transferases (Carter et al., 1988; Nickbarg et al., 1988; Warshel et al., 1989; Dipersio et al., 1991; Miran et al., 1991; Munier et al., 1992).

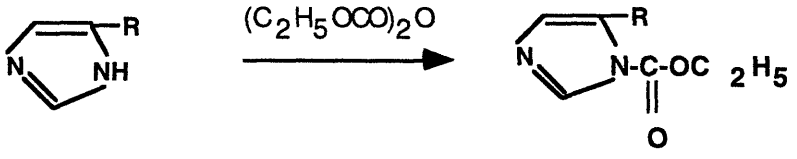
The role of histidine residues in heparinase I activity was investigated. This chapter describes the results obtained from a systematic study which included a combination of chemical modification and site-directed mutagenesis experiments. The key results along with the rationale for conducting the experiments are described in section 4.2 and section 4.3 presents a detailed discussion of the results.

4.2 Results

Chemical Modification with DEPC—Rationale

Diethylpyrocarbonate (DEPC) reacts with histidyl residues in model systems and in proteins according to **equation 4.1** to yield an N-carbethoxy-histidyl derivative (Miles, 1977). The chemical formula of DEPC is $(C_2H_5OCO)_2O$. This reaction can be easily monitored spectrophotometrically by the increase in absorbance, which has a

maximum between 230 and 250 nm. The number of modified residues can be calculated from a molar absorption difference for N-carbethoxyhistidine at 240 nm : ε = 3200 M^{-1}. This reaction is reversible since hydroxylamine can remove the carbethoxy group and the enzyme is reactivated. Use of excess reagent must be avoided since this could result in an incorrect estimation of modified residues.



(4.1)

DEPC reacts with other residues as well (see below) but has been shown to be specific for histidyl residues under certain reaction conditions. In addition, reversibility with hydroxylamine is an indication of histidyl or tyrosyl residue modification and not lysine or sulfhydryl residues. By adjusting the reaction conditions the specificity towards histidine residues can be increased.

4.2.1 Inactivation of heparinase I with DEPC

The effect of DEPC on heparinase I activity was studied as described in Materials and Methods (Chapter 3). In the presence of 0.3 mM DEPC about 95 % of heparinase I was inactivated in less than 3 min and as low as 10 µM DEPC resulted in a 85 % loss of heparinase I activity in 15 min. The inactivation rate was concentration dependent: at 1 mM and 0.05 mM DEPC, inactivation was complete within 1 min and 10 min, respectively. The control sample, which contained 2 % v/v ethanol showed no change in heparinase I activity. Plots of the log % residual activity versus time of incubation at different DEPC concentrations were linear indicating that the inactivation followed pseudo-first order kinetics **(Figure 4.1)**. The pseudo-first-order inactivation rate constants k, were determined from the activity after a determined time, A_t, the initial activity A_i and **equation 4.2**:

$$A_t = A_i \exp(-kt) \qquad (4.2)$$

A plot of the observed pseudo-first-order rate constants as a function of DEPC concentration (0.01-0.1 mM) yielded a straight line passing through the origin

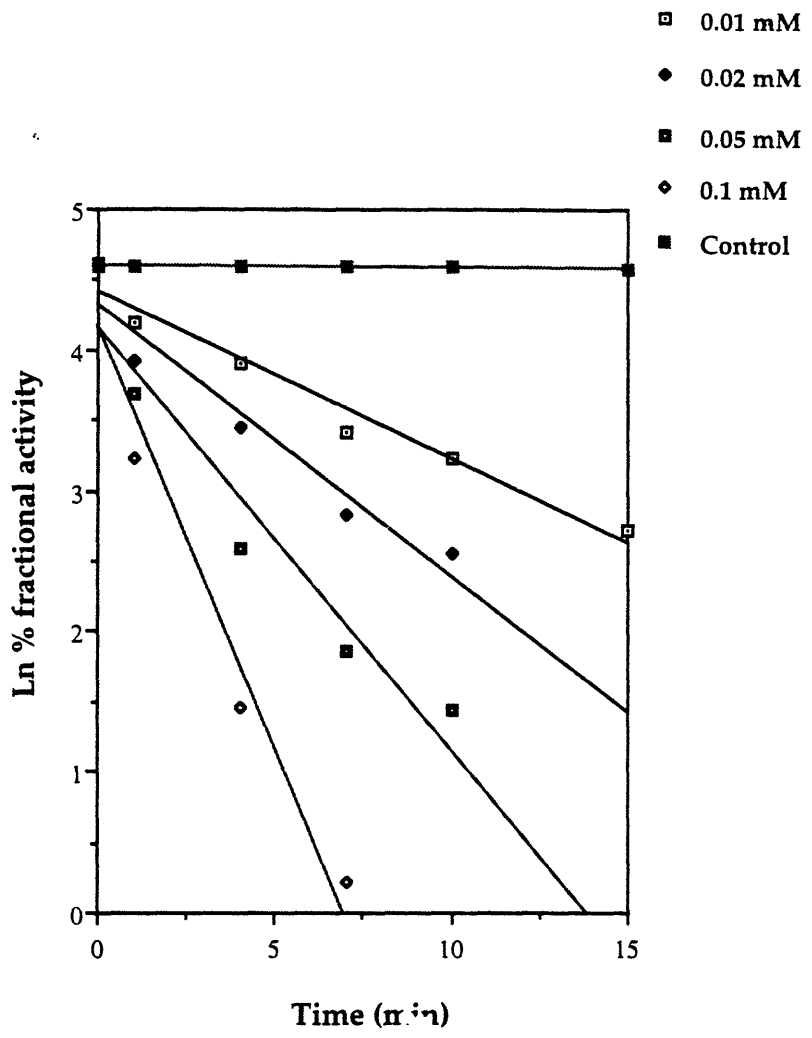


Figure 4.1: *DEPC inactivation of heparinase I.* Heparinase I was incubated with 0 mM, 0.01 mM, 0.02 mM, 0.05 mM and 0.1 mM DEPC at pH 7.0 and at room temperature. The fractional activity was expressed as a percentage of the initial activity and the natural logarithm of the fractional % activity was plotted as a function of time. The values are averages of triplicate determinations. The data were fitted to equation 1; the rate constants were calculated to be k= 0.119 min^{-1} (0.01mM DEPC); k = 0.193 min^{-1} (0.02 mM DEPC); k= 0.3 min^{-1} (0.05mM DEPC) and k = 0.6 min^{-1} (0.1 mM DEPC).

(Figure 4.2), from which a second order rate constant of 5.6 min^{-1} mM^{-1} was obtained. The determination of a second order inactivation rate constant indicated that no reversible complex between the inactivating agent (DEPC) and the enzyme occurred (Battaglia et al., 1994). A plot of log k versus log DEPC concentration **(Figure 4.3)** yields a slope of 0.76 (~1), indicating that the rate of inactivation is first order with respect to DEPC concentration.

DEPC inactivated the enzyme effectively at sufficiently low concentrations, suggesting that active site residues were being modified and it remained to be seen which residues were modified. Reversibility of inactivation using hydroxylamine was next investigated.

4.2.2 Reversibility of inactivation using hydroxylamine

As pointed out earlier, hydroxylamine removes the carbethoxy groups from modified histidyl and tyrosyl residues (Miles, 1977) and hence, experiments were performed to determine if the loss in heparinase I activity due to modification by DEPC, could be recovered by treating the enzyme with hydroxylamine. When heparinase I was incubated with 7.5 μM DEPC, the enzyme was inactivated to approximately 74 % of its original activity in 1 min. The modified enzyme was then incubated with 300 mM hydroxylamine at 4° C. As shown in **Table 4.1**, the enzyme activity was restored to approximately 95 % of its initial activity in about 4 hr.

Table 4.1 Hydroxylamine reversibility of DEPC based inactivation of heparinase I.

Time (min)	1	45	105	165	225
% Initial activity	74	79	88	91	95

Heparinase I inactivation was accomplished with 7.5 μM DEPC in MOPS at room temperature for 1 min. The modified enzyme was then immediately incubated with 300 mM hydroxylamine at 4° C for about 4 hr. Activity measurements were taken every hr.

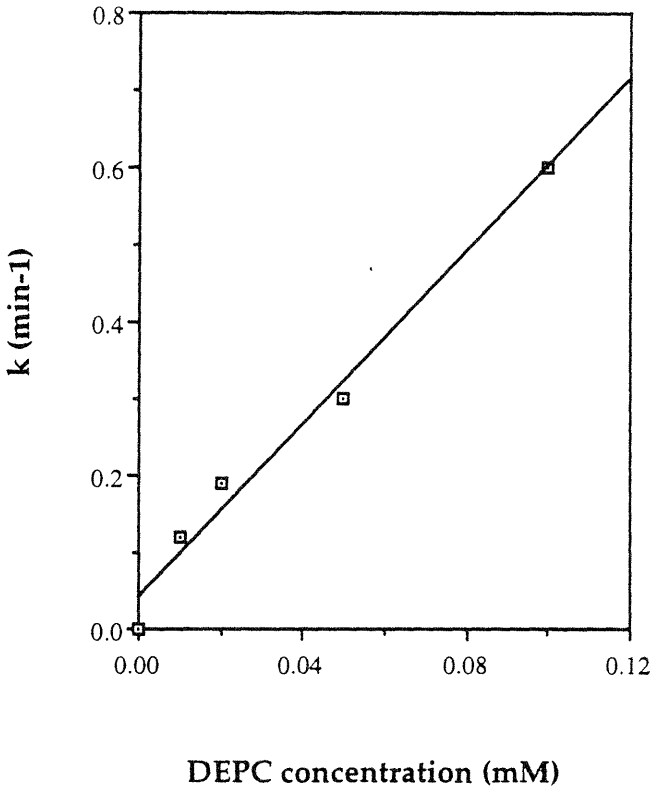


Figure 4.2: *Effect of DEPC concentration on the pseudo-first-order rate constants of inactivation.* The pseudo-first order rate constants are plotted as a function of the DEPC concentration. A second order rate constant of 5.6 min^{-1} mM^{-1} was calculated from the plot.

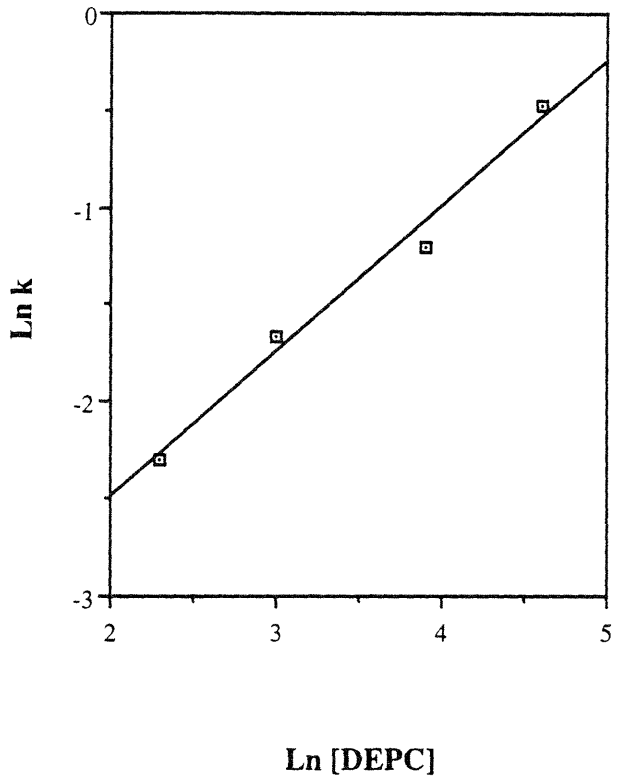


Figure 4.3: *Effect of DEPC concentration on the pseudo-first-order rate constants of inactivation.* The logarithm of pseudo-first order rate constants are plotted as a function of the logarithm of DEPC concentration. A slope of about 0.8 was obtained from the plot indicating the rate of inactivation is first order with respect to DEPC concentration.

4.2.3 *Specificity of inactivation by DEPC*

DEPC is known to react with the side chains of several amino acid residues viz. histidine, lysine, cysteine and tyrosine. By appropriately choosing experimental conditions, the specificity of the reagent towards a residue can be increased. For example, at pH 6.0, DEPC reacts with histidyl residues in various proteins with a high degree of specificity (Fu and Robyt, 1988). Further, lysine residues can be modified by DEPC at pH more than 7.0.

When heparinase I was incubated with 0.025 mM DEPC at pH 6.0, more than 80 % of initial enzymatic activity was lost within 15 min, suggesting specificity towards histidine residues. Further, restoration of heparinase I activity by hydroxylamine ruled out the possibility of lysine and cysteine modification and hence pointed to histidine or tyrosine modification. However, when heparinase I was incubated with DEPC, no decrease in the absorbance at 278 nm was observed, thus precluding the possibility of tyrosine modification (Fu and Robyt, 1988). The above results taken together and the relatively high specificity of the DEPC reaction at pH 6.0 for imidazole groups, strongly suggested that inactivation of heparinase I by DEPC was due to modification of the imidazole ring of histidine residues. Hence it remained to be seen if this inactivation was due to modification of histidine residues in the active site or substrate binding site.

4.2.4 *Substrate protection of DEPC inactivation*

One could argue that if a histidine residue plays a role in enzymatic activity (such as in substrate binding or catalysis), the rate of DEPC inactivation should decrease in the presence of substrate. To investigate the ability of heparin (substrate) to protect heparinase I inactivation by DEPC, the enzyme was pre-incubated with heparin or heparin fragments (see discussion) for about 45 min. As seen in **Figure 4.4**, in the presence of 0.5 mg of heparin/ml (~ 5 K_m) and a DEPC concentration of 0.05 mM, the rate of inactivation significantly decreased. Rate constants were determined by assuming pseudo-first-order kinetics and fitting data to **equation 4.2**. The rate constant in the absence of heparin was k = 0.43 min^{-1} and in the presence of 0.5 mg of heparin/ml the rate constant was reduced by an order of magnitude to k = 0.029 min^{-1}. Since the heparin concentration in the assay medium was much larger than

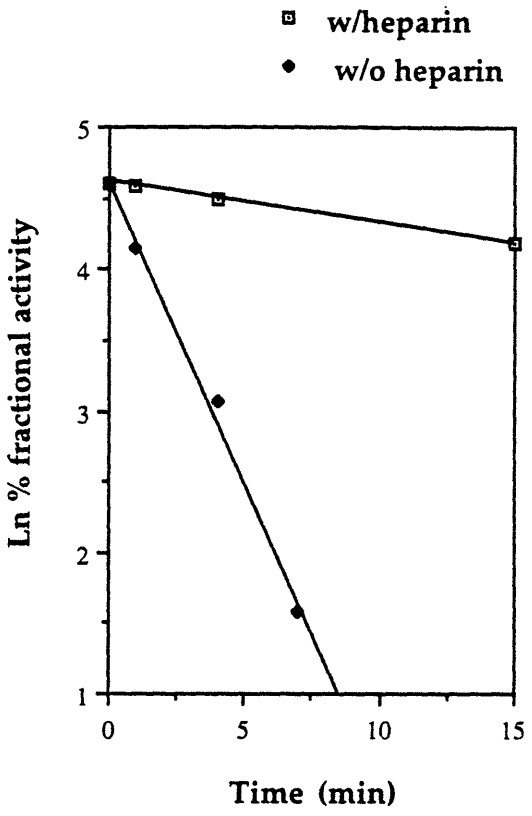


Figure 4.4: *Substrate protection of DEPC inactivation of heparinase I* Heparinase I was incubated with 2 mg/ml heparin (final concentration 0.5 mg/ml) for a period of 45 min prior to the addition of 0.05 mM DEPC. A control experiment with no heparin added was also performed. Aliquots of the enzyme sample were withdrawn every 3 minutes, for a total time of about 15 mins, for enzyme activity determination. The fractional activity was expressed as a percentage of the initial activity and the natural logarithm of the fractional % activity was plotted as a function of time (the values are averages of triplicate determinations). The pseudo first order rate constant was calculated, by fitting the data to equation 1, to be $k = 0.43\ min^{-1}$ in the absence of heparin and in the presence of 0.5 mg of heparin/ml the rate constant was reduced by an order of magnitude to $k = 0.029\ min^{-1}$.

the K_m of 0.1 mg/ml (Yang et al., 1985) any additional heparin introduced did not alter the kinetics.

When the experiment was done by incubating heparinase I in the presence of heparin fragments (primarily disaccharides and tetrasaccharides which are heparinase I derived heparin cleavage products) the inactivation rate did not change significantly **(Figure 4.5)** ($k=0.017\ min^{-1}$) (see discussion). In addition, a smaller incubation time of heparinase I with heparin (10 min) was sufficient to see significant protection.

The data obtained so far from the chemical modification experiments strongly suggested the involvement of histidine residues in either substrate binding or catalysis. The number of histidine residues modified remained to be determined and hence experiments were conducted to quantify the number of modified histidine residues.

4.2.5 Quantification of histidine residues

To determine the number of modified histidine residues, the inactivation of heparinase I by DEPC was plotted against the number of modified histidine residues (calculated from the absorbance increase at 240 nm using an extinction coefficient of 3200 $M^{-1}cm^{-1}$). The plot was non-linear indicating that not all the histidine residues are modified at the same rate, since the modification is dependent on the accessibility of the residues (Fu and Robyt, 1988). Extrapolation of the first linear phase of the plot (where the enzyme is about 40 % inactivated) to zero activity indicated that between one and two histidine residues were modified per mole of heparinase I. However, this does not determine the number of histidine residues essential for enzyme activity but rather only suggests that these residues are modified at a faster rate. To further corroborate the number of histidine residue(s) modified, radiolabeling of heparinase I with [^{14}C]DEPC was performed.

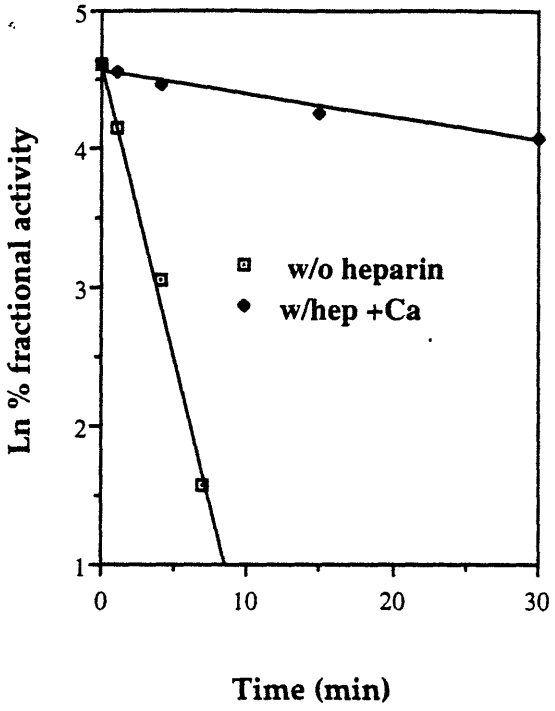


Figure 4.5: *Substrate protection of DEPC inactivation of heparinase I* Heparinase I was incubated with 2 mg/ml heparin (final concentration 0.5 mg/ml) for a period of 45 min (in the presence of 5 mM calcium) prior to the addition of 0.05 mM DEPC. A control experiment with no heparin added was also performed. Aliquots of the enzyme sample were withdrawn every 3 minutes, for a total time of about 15 mins, for enzyme activity determination. The fractional activity was expressed as a percentage of the initial activity and the natural logarithm of the fractional % activity was plotted as a function of time (the values are averages of triplicate determinations). The pseudo first order rate constant was calculated, by fitting the data to equation 1, to be $k = 0.43\ min^{-1}$ in the absence of heparin and in the presence of heparin fragments the rate constant was reduced by an order of magnitude to $k = 0.017\ min^{-1}$.

4.2.6 *Radio-labeling with [^{14}C]DEPC and tryptic mapping of the [^{14}C]histidines of heparinase I*

To investigate the relationship between the extent of DEPC modification and heparinase I inactivation, the amount of isotope incorporation when heparinase I was treated with [^{14}C]DEPC was measured. Heparinase I which had been treated with 0.15 mM [^{14}C]DEPC incorporated ~ 0.33 ± 0.05 x 10^3 cpm [^{14}C]DEPC/µg of heparinase I or 1.3 ± 0.05 x 10^4 cpm [^{14}C]DEPC/nmole of heparinase I. The average [^{14}C] incorporation from four different experiments was about 1.4 moles modified residues/mole of heparinase I (based on the specific activity of 2.6 mCi/mmol). Thus the results consistently pointed towards one or possibly two histidine residue(s) being modified for every mole of non-denatured heparinase I in its native conformation. This is in good agreement with the stoichiometry calculated from the absorbance changes.

Heparinase I modified by non-radiolabeled or [^{14}C]DEPC was digested with trypsin and the tryptic peptides were separated using RPHPLC. Several peptide maps were generated from both the modification experiments and these maps were both identical, as well as, reproducible **(Figure 4.6 a)**. When the tryptic digests of the modified enzyme were compared to a control (unmodified) heparinase I tryptic digest, the maps were almost identical, with minor alterations in a few defined regions **(Figure 4.6 b)**. However, the tryptic mapping experiments, for both non-radiolabeled as well as [^{14}C]DEPC modified heparinase I, under the experimental conditions tested, were inconclusive in mapping the histidine residue(s) essential for enzymatic activity (see discussion). Hence in order to identify the histidine residue(s) critical for enzyme activity, site-directed mutagenesis was performed. Wild-type heparinase I has 4 histidine residues (Sasisekharan et al., 1993) and each of these were individually modified to alanines. As described for cysteine mutagenesis in chapter 2, alanine was chosen for its neutral charge and smaller side group.

4.2.7 *Site-directed mutagenesis of heparinase I*

Four mutant recombinant heparinases were constructed (H129A, H165A, H203A, H339A) and expressed in the BL21(DE3) host. *r*-heparinase I construct devoid of the

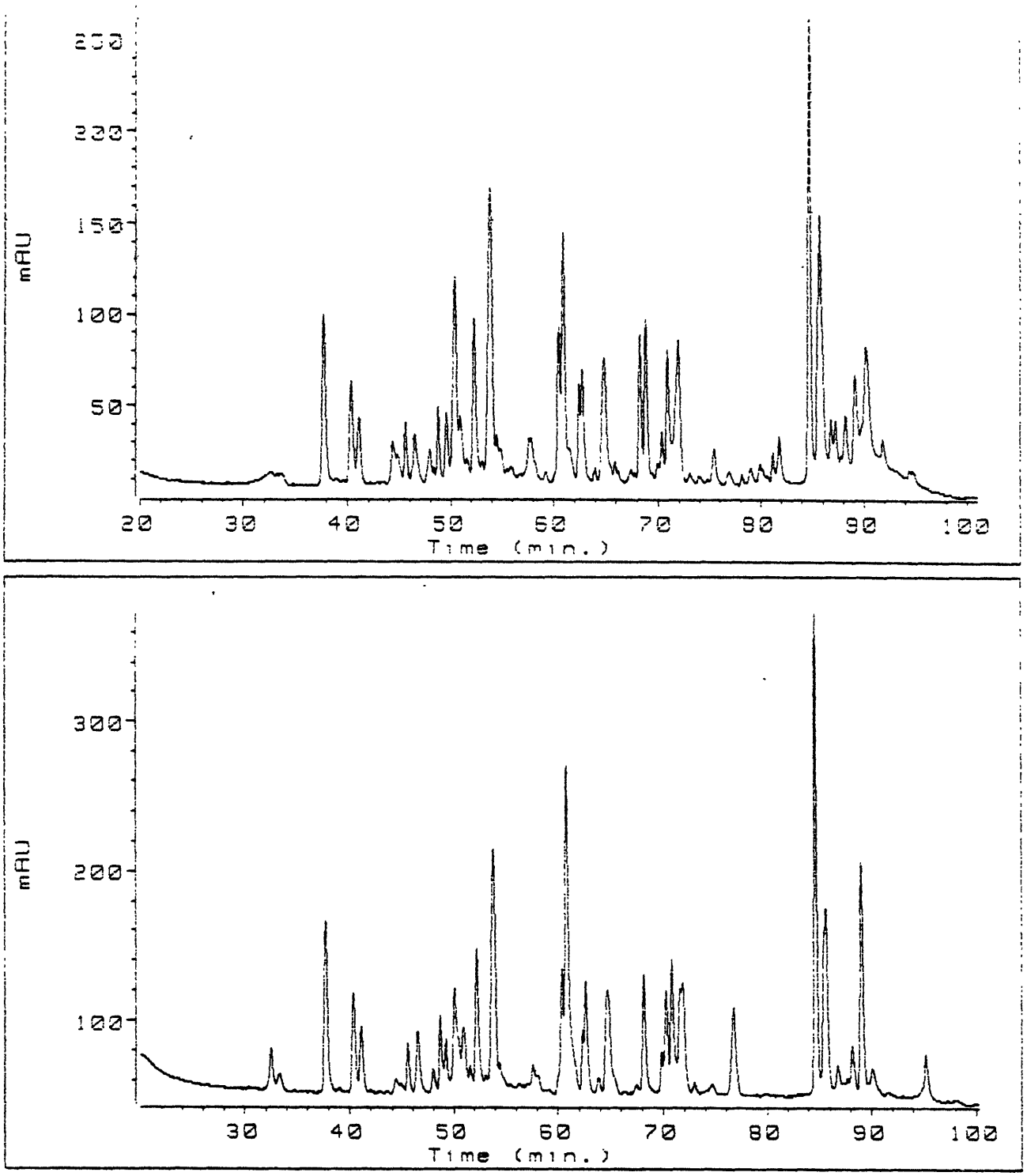


Figure 4.6: a) Tryptic map of heparinase I modified with [^{14}C]DEPC. The enzyme was labeled with [^{14}C]DEPC and digested with trypsin overnight. The tryptic peptides were separated using RPHPLC. b) Tryptic map of control heparinase I.

putative signal sequence (-L *r*-heparinase I) was expressed as a control (Sasisekharan et al., 1993). The level of protein expression for all the recombinant heparinases was identical in the BL21(DE3) host. The purity of *r*-heparinases as determined by SDS-PAGE are estimated to be 80-90% after the first step and greater than 98% with silver stain after the second step **(Figure 4.7)**.

The mutant enzymes were used in online activity assays to determine the kinetic parameters as well as overnight heparin degradation assays to determine the product profile and compare it with wild type heparinase I. It was found that mutants H129A, H165A, and H339A were expressed in *E. coli* with essentially identical activity as the control -L *r*-heparinase I **(Figure 4.8)**. However, the H203A mutation interestingly resulted in an enzyme with no detectable activity, suggesting a role for this residue in heparinase I activity or substrate binding **(Figure 4.9)**. **Table 4.2** lists the kinetic parameters for the recombinant and mutant heparinases.

Table 4.2 Kinetic constants of *r*-heparinase I and the mutant *r*-heparinase I

Enzyme	$k_{cat} sec^{-1}$	K_m μM
wild type *r*-heparinase I	92	10.2
H129A *r*-heparinase I	89	9.1
H165A *r*-heparinase I	88	9.2
H203A *r*-heparinase I	ND[a]	ND
H203D *r*-heparinase I	3.5	5.7
H339A *r*-heparinase I	91	9.1

ND[a]= not determined since the enzyme was inactive.

The kinetic constants reported in the present work for the wild type *r*-heparinase I are consistent with previous reports for the native *F.heparinum* heparinase I (Lohse & Linhardt, 1992).

It is interesting to note that histidine-203 is part of the primary heparin binding site of heparinase I (chapter 2). In order to differentiate between a catalytic and a substrate binding role for histidine-203 (see discussion), an additional H203D mutant was expressed and purified. As shown in **Figure 4.10**, this enzyme was active but with a significant (30 fold) reduction in k_{cat} **(Table 4.2)**. This suggested a

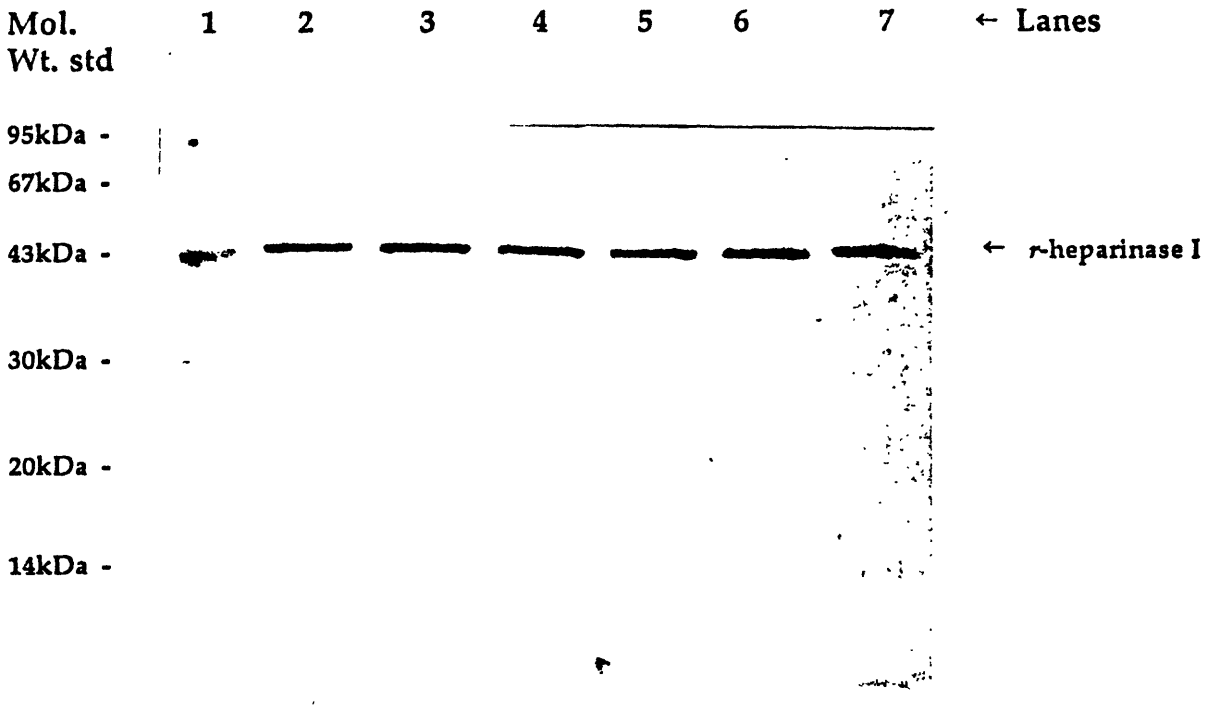


Figure 4.7: *SDS-PAGE of r-heparinase I and mutant r-heparinase I:* A 12% SDS-PAGE of heparinase I. Lanes are described from left to right. **Lane 1:** Molecular weight standards. **Lane 2:** Ni column purified H129A mutant *r*-heparinase I. **Lane 3:** Ni column purified H165A mutant *r*-heparinase I. **Lane 4:** Ni column purified H203A mutant *r*-heparinase I. **Lane 5:** Ni column purified H203D mutant *r*-heparinase I. **Lane 6:** Ni column purified H339A mutant *r*-heparinase I. **Lane 7:** Ni column purified *r*-heparinase I.

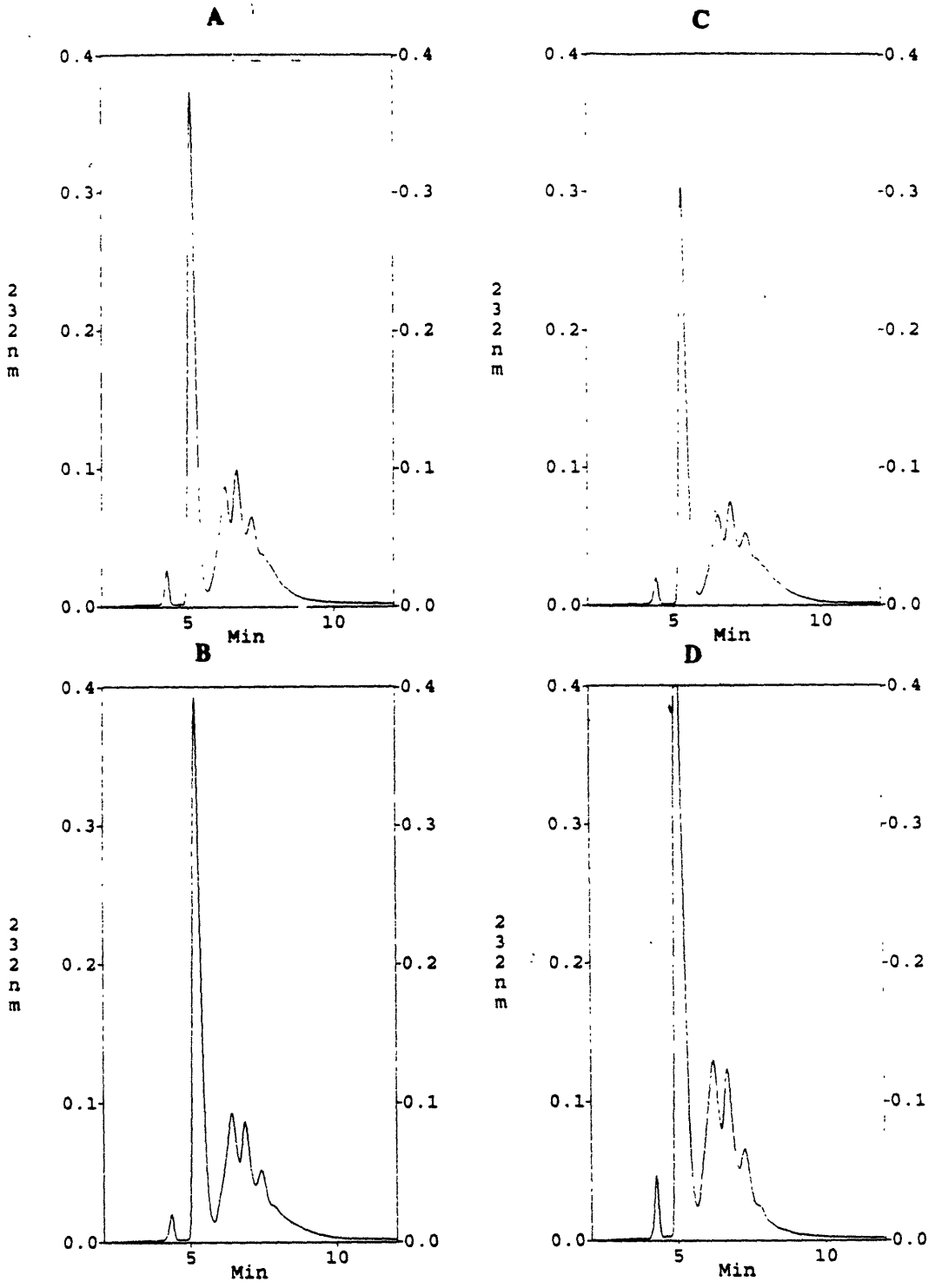


Figure 4.8: *Effect of histidine mutations on activity of heparinase I:* **A** shows the product profile of heparin degradation by *r*-heparinase I; **B** shows the product profile of heparin degradation by the mutant H129A; **C** shows the product profile of heparin degradation by the mutant H165A; **D** shows the product profile of heparin degradation by the mutant H339A.

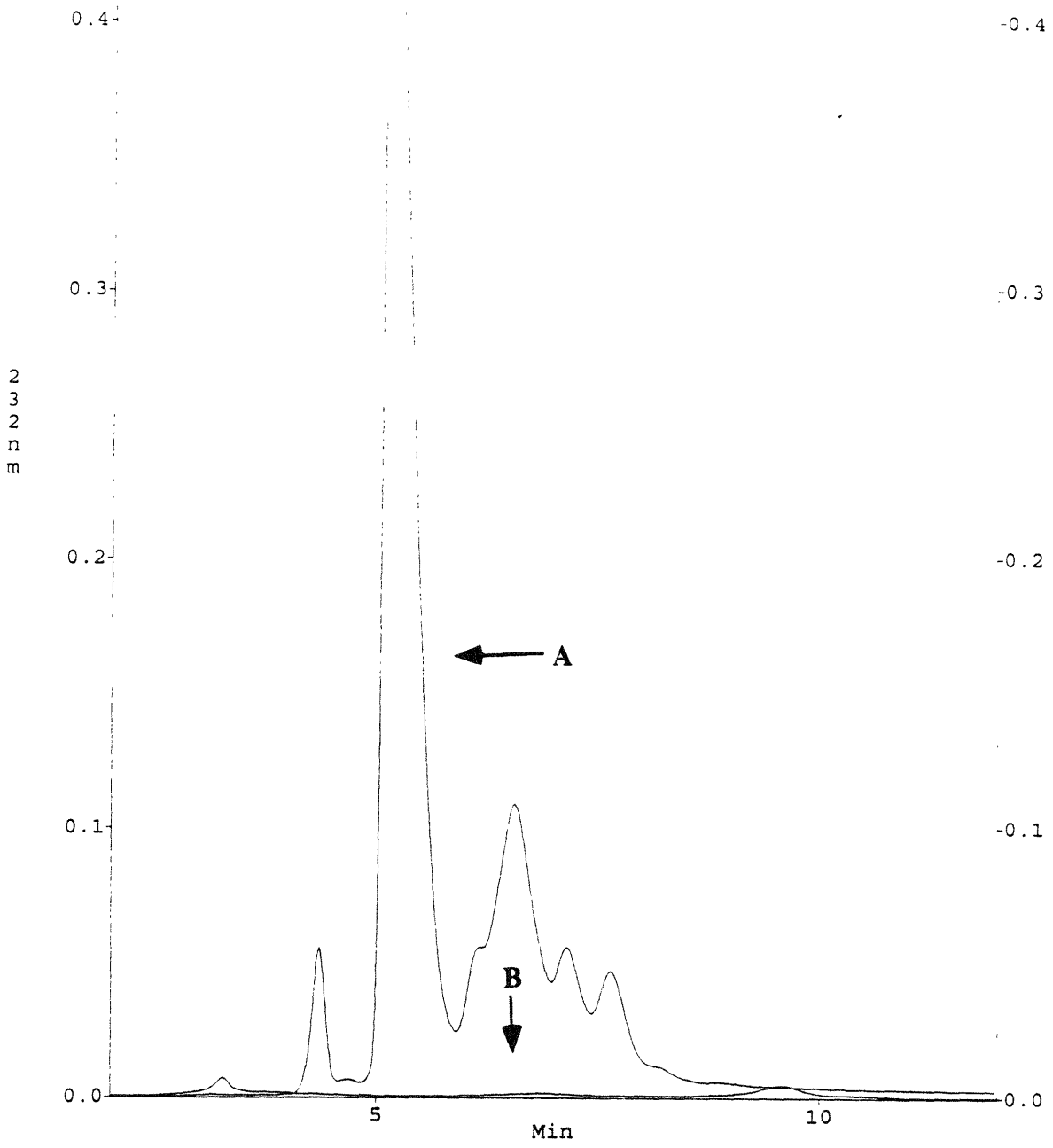


Figure 4.9: *Effect of histidine mutations on activity of heparinase I* **A** shows the product profile of heparin degradation by *r*-heparinase I; **B** shows the product profile of heparin degradation by the mutant H203A.

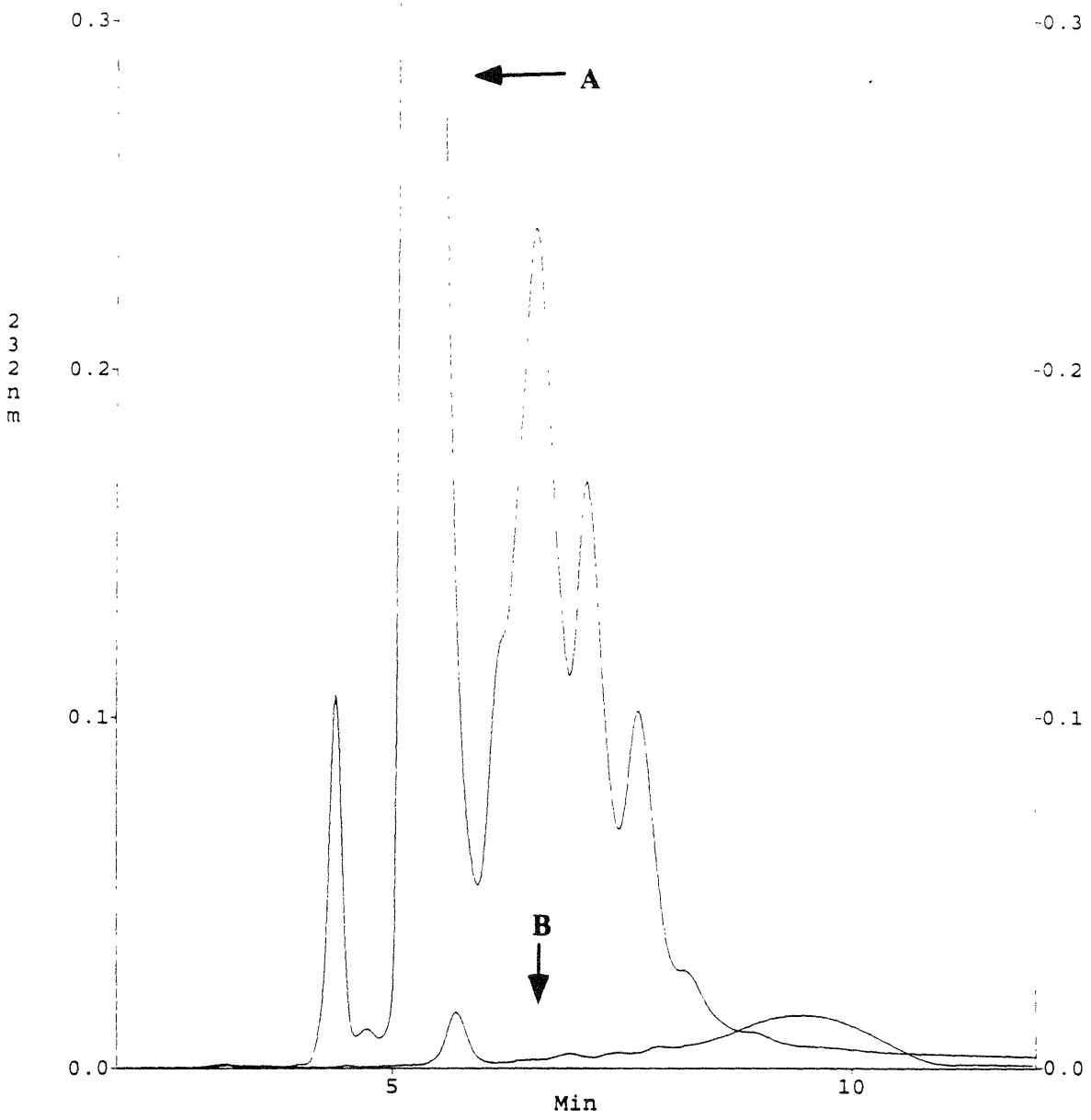


Figure 4.10: *Effect of histidine mutations on activity of heparinase I* **A** shows the product profile of heparin degradation by *r*-heparinase I; **B** shows the product profile of heparin degradation by the mutant H203D.

nucleophilic role for histidine-203, as opposed to a role in heparin binding. The above results from chemical modification studies and site-directed mutagenesis experiments taken together strongly suggest the involvement of histidine-203 in heparinase I activity.

4.3 Discussion

4.3.1 Characterization of functional role of histidine in heparinase I

Heparinase I modification by DEPC at pH 7.0 inactivated the enzyme. The rate of inactivation followed a pseudo-first order mechanism, which suggests the presence of a highly reactive residue critical for enzymatic activity. The value of the second order rate constant for the inactivation of heparinase I by DEPC (5.6 mM^{-1} min^{-1}) was high when compared with that measured for other enzymes under similar conditions (Abdulwajid et al., 1986; Pelton et al., 1992; Battaglia et al., 1994).

DEPC can react with the side chains of several amino acid residues such as cysteines, lysines, histidines and tyrosines (Miles, 1977). The importance of lysines in heparinase I activity was demonstrated (Leckband and Langer, 1991) and chapter 2 described a cysteine residue involved in the catalytic site. When heparinase I was partially inactivated by DEPC, its activity could be restored by addition of 300 mM hydroxylamine. This ruled out the involvement of lysines and cysteines and suggested the possible modification of histidine or tyrosine residues (Miles, 1977). In addition, lysine residues can only be modified by DEPC at pH values higher than 7.0 (Fu and Robyt, 1988). When heparinase I was incubated with 25 µM DEPC at pH 6.0, more than 80 % of its initial activity was lost within 15 min, ruling out the involvement of lysines. Furthermore, the absence of decrease in the 278 nm absorption spectra when heparinase I was incubated with DEPC, indicated that tyrosines were not modified, hence pointing to the specificity of the reaction towards histidine residues.

To quantify the number of modified histidine residues, inactivation of heparinase I by DEPC was correlated with the moles of histidines modified (calculated from absorbance changes at 240 nm and from [^{14}C]DEPC incorporation) per mole of enzyme. The data from both experiments pointed towards one or two residues

being modified per mole of enzyme. DEPC has been used over a wide range of concentrations, from 0.01 to 40 mM (the maximum solubility in water) depending on the reactivity or accessibility of histidyl residues in various proteins (Miles, 1977). It is interesting to note that heparinase I displayed great sensitivity to this reagent, being rapidly and completely inactivated under relatively mild conditions (i.e. as low as 0.01 mM DEPC), suggesting that histidine residues are critical and readily surface accessible.

In addition, the rate of inactivation was significantly (ten fold) reduced in the presence of heparin, suggesting the proximity of the modified residues to the active site or the substrate binding site. Similar heparin protection against inactivation was also observed previously for chemical modification of the surface accessible active site cysteine-135 (Sasisekharan et al., 1995). It is possible that DEPC labeling could sterically hinder heparin access to the active site; another explanation is that heparin bound to heparinase I could alter the charge characteristics and hence lower the labeling kinetics. However, a mechanistic role was established through site-directed mutagenesis experiments (see below). The above results taken together, viz. pseudo-first order inactivation by DEPC; the reversal of DEPC inactivation by hydroxylamine; and the protection against DEPC inactivation by heparin, strongly suggested a role for histidine residues in catalysis.

It is of interest to note that of the four histidines in heparinase I, histidine-129 is in close proximity to the active site cysteine-135 in the primary sequence of heparinase I, and histidine-203 is contained in the primary heparin binding site (which was proposed to be in proximity to the active site cysteine-135 in the folded, tertiary structure) (chapter 2) (see below). As stated above, both non-radiolabeled as well as [^{14}C]DEPC labeling experiments pointed to one or two residues being surface accessible and hence this raised the possibility of either or both of the above mentioned histidine residues being surface accessible.

However, the tryptic mapping experiments with non-radiolabeled DEPC and [^{14}C]DEPC modified enzyme showed that the reversed-phase elution of tryptic peptides td9 and td50 (Sasisekharan et al., 1993), containing histidine-129 and histidine-339 respectively, were unaltered if not identical to that of the control tryptic map, suggesting that these residues were unlikely to be modified. The effect

of chemical modification on the migration of tryptic peptides containing histidine-203 and histidine-165 were inconclusive. It is possible that the very small and hydrophilic tryptic peptide containing histidine-203 might have eluted in the void volume (chapter 2). The [^{14}C]DEPC concentration used in the peptide mapping studies was identical to the non-radiolabeled DEPC (used for both kinetic and mapping studies), in order to minimize non-specific modifications. This resulted in a lower specific activity of [^{14}C]DEPC, thus enormously decreasing the sensitivity of mapping tryptic peptides. Hence, we sought to map the critical histidine residue through systematic site-directed mutagenesis experiments as described below.

4.3.2 *Site-directed mutagenesis of histidines of heparinase I*

In order to determine which histidine residue is involved in enzyme activity and to ascertain a mechanistic role, site-directed mutagenesis was performed. The H129A, H165A and H339A mutations did not affect enzyme activity and the kinetic parameters **(Table 4.2)** suggesting that these residues are not essential for enzyme activity.

The most significant effect on enzyme activity was observed when histidine-203 was replaced with an alanine. The H203A mutation completely inactivated heparinase I, indicating the importance of this residue in enzyme activity. As pointed out earlier, histidine-203 belongs to the heparin binding consensus sequence of the heparin binding site in heparinase I. To differentiate between a nucleophilic role and a role in heparin binding, this residue was replaced with an aspartic acid, which can act as a nucleophile but is not known to participate in heparin binding due to its negative charge. This resulted in an enzyme with residual activity (30 fold reduction in k_{cat}), suggesting a nucleophilic role for this residue.

When heparin protection of DEPC inactivation of heparinase I was conducted in the presence of heparin or heparin fragments such as di- and tetrasaccharides, the inactivation rates were not significantly different. The K_m of heparinase I binding to heparin (0.3 μM) differs significantly from K_m of heparinase I binding to individual tetrasaccharides (15 μM-80 μM) (Rice and Linhardt, 1989). Hence if histidine-203 were to play a role in heparin binding, one might expect the inactivation kinetics in the presence of heparin to be different from that in the presence of primarily di- and

tetrasaccharides. The observation of more or less identical inactivation rates is consistent with the hypothesis that histidine-203 plays a role in catalysis.

4.3.3 Conclusion—the catalytic domain of heparinase I

As pointed out earlier, a surface accessible cysteine (cysteine -135), contained in a basic environment, was essential for heparinase I activity and it was hypothesized that a heparin binding domain provided the basic environment to active site cysteine-135 (chapter 2). The primary heparin binding site of heparinase I (which contains two basic clusters) was identified and mapped. Preliminary results showed that sulfhydryl selective labeling of cysteine-135 in heparinase I protected the lysines of the heparin binding sequence from proteolytic cleavage, corroborating the hypothesis of the close proximity of heparin binding site to the active site cysteine-135 (chapter 2).

However, the most interesting observation from the work presented in this chapter, in conjunction with the above mentioned parallel studies (chapter 2) is the finding that not only is histidine-203 contained in the heparin binding site, but also it is critical for catalysis. The determination of a catalytic role for histidine-203 provides the most compelling evidence for the heparin binding domain being in close proximity to the active site cysteine-135. It is proposed that histidine-203 and cysteine-135 are adjacent to the scissile HI linkage during catalysis and form the active site or the catalytic domain of heparinase I.

Cysteine reactivity has been shown to be enhanced by the proximity of basic amino acids (Rabin and Watts, 1960; Torchinsky, 1981). In the case of heparinase I, it was suggested that a positively charged heparin binding site would tend to keep the thiol group of the active site cysteine-135 negatively charged (by lowering its pKa) so that it can act as a base for proton abstraction (chapter 2). Furthermore, histidine residues in particular have been thought to be involved in promoting the reactivity of cysteine residues (Amuro et al., 1985; Miran et al., 1991) and subsequently act as an acid/base catalyst for proton abstraction and donation. In cysteine proteases such as papain, the active site contains a imidazolium-thiolate ion pair, where the histidine activates the cysteine residue for catalysis (Lewis et al., 1981).

In conclusion, employing chemical modification and site-directed mutagenesis approaches, the importance of histidine-203 in the enzyme's catalytic action was demonstrated. Histidine-203 is a part of the heparin binding site and is critical for catalysis in heparinase I. This study has led to the identification of the second critical residue of the catalytic domain of heparinase I. It remains to be seen if histidine-203 acts as a primary nucleophile (i. e. as a base) by abstracting a proton from the substrate or activates the cysteine-135 residue for catalysis and hence acts as part of a charge relay system, as seen in several other enzymes.

This chapter investigates the interplay between binding and catalysis through extensive site-directed mutagenesis studies. The role of positive charge in the heparin binding domain and residues from the calcium coordinating motif was investigated by systematic mutagenesis. Results show that lysines-198 and 199 from one of the consensus sequences are critical for heparinase I activity. Glycine-213 (a critical conserved residue of the consensus calcium coordination motif) and glutamate-207 are important in calcium coordination. Further, lysine-132, close to active site cysteine-135 in the primary sequence, is also essential for enzymatic activity. Thus a three-dimensional domain of basic residues around the catalytic site is involved in both, heparin binding through charge complimentarity, as well as providing the basic environment for cysteine-135. A model of the substrate binding cleft in heparinase I is proposed. The substrate binding pocket contains the catalytic domain comprising cysteine-135 and histidine-203. Finally, model predictions are tested through the effect of steric mutations on enzyme activity.

Chapter 5

The Heparin Binding Domain of Heparinase I

5.1 Motivation

The previous chapter described the identification of a second critical residue in heparinase I activity that led to the definition of a catalytic domain in the enzyme comprising cysteine-135 and histidine-203. Histidine-203 is part of the primary heparin binding site in heparinase I and hence this established the proximity of the heparin binding site to the active site cysteine-135. Histidine-203 was proposed to play a role in the catalytic mechanism as opposed to heparin binding since a H203A mutation inactivated the enzyme while a H203D mutation retained residual activity.

It was pointed out in chapter 2 that several proteins interact with heparin to mediate a variety of biological processes. However, the molecular basis of these interactions remains poorly characterized. This lack of knowledge results from the structural diversity of the proteins binding to heparin, the extensive heterogeneity in the chemical composition of these complex macromolecules, and the multiplicity of the biological processes which involve such interactions (Kjellèn and Lindahl, 1991; Jackson et al., 1991; Hata et al., 1993).

The binding between polyanionic heparin (containing negatively charged sulfate and carboxylate groups) and proteins primarily involves electrostatic interactions which are disrupted at high salt concentrations (Hata et al., 1993). Such binding interactions would be mediated by positively charged residues in proteins. Indeed, Cardin and Weintraub (1989) reported that several heparin-binding proteins, contain consensus sequences such as XBBBXXBX or XBBXBX (B = basic residues; X = hydrophobic or other residues) which are involved in heparin binding. This has been confirmed through site-directed mutagenesis and binding studies with synthetic or isolated peptides from several of these proteins (Smith and Knauer, 1987; Baird et al., 1988; Bober Barkalow and Schawarzbauer, 1991; Bae et al., 1994; Thompson et al., 1994).

Structural requirements for heparin have been characterized, with the isolation of specific oligosaccharide sequences that confer binding specificity, in the case of AT III, bFGF, LPL and others (Lindahl et al., 1984; Maimone and Tollefsen, 1990;

Maccarana et al., 1993; Parthasarathy et al., 1994). These studies have shown specific heparin binding, dependent on a particular heparin structure with sulfate and carboxylate groups at defined positions. The binding specificity of heparin binding proteins is dependent on certain sequences of the polysaccharide as well as characteristic amino acid sequences of the protein.

As discussed earlier, heparinases I, II and III recognize certain sequences of sulfation and uronic acid epimerization in heparin like molecules with a high degree of specificity (Desai et al., 1993). The substrate specificity of heparinases has been well characterized (Rice and Linhardt, 1989; Linhardt et al., 1990). However, there is limited information on the molecular features of the enzymes that confer this specificity (Linhardt et al., 1986; Lohse and Linhardt, 1992; Sasisekharan et al., 1995). Since heparin-protein interactions are governed by molecular recognition involving specific sequences in both heparin as well as the proteins, it is possible that the substrate specificities of heparinases may be governed by the mechanism of substrate binding.

The identification of a heparin binding site in heparinase I was described in chapter 2. This 18 amino acid region (spanning amino acid residues 196-213 of the heparinase I primary sequence) contains two putative heparin binding consensus sequences (Cardin and Weintraub, 1989) spanning residues 197-204 (segment I) and residues 207-212 (segment II) **(Figure 5.1)**. Furthermore, an additional cluster of basic residues (spanning residues 332-339), that conforms to the consensus sequence, is present in heparinase I. Interestingly, in addition to the above consensus sequences, the tryptic peptide containing cysteine-135 also bound heparin, leading to the hypothesis of a three-dimensional heparin binding domain rather than a linear heparin binding site (chapter 2). Finally, calcium was shown to be required for catalysis and the heparin binding site contains a putative calcium binding loop (residues 207-213) of the 'EF-hand' structural domain (Kretsinger, 1980).

It was proposed that active site cysteine-135 is present in a positively charged environment (Sasisekharan et al., 1995). Since histidine-203 being a part of the catalytic domain indicates the proximity of the heparin binding site to active site cysteine-135, it is possible that the basic clusters in the heparin binding site could provide for the positive charge around cysteine-135. The role of the positive charge

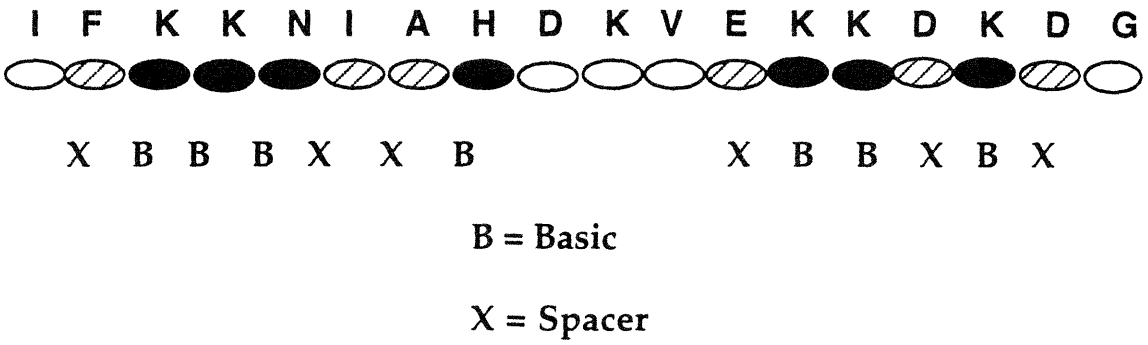


Figure 5.1: The primary heparin binding site of heparinase I showing the two Cardin and Weintraub heparin binding consensus sequences and a calcium coordination motif.

is possibly to provide the necessary charge complimentarity for specific heparin binding and also to provide the basic environment to bias the active site reactivity. Hence, the role of positive charge in heparin binding and enzyme activity was addressed through systematic site-directed mutagenesis studies. The heparin binding domain of heparinase I was dissected in order to investigate the role of the above mentioned regions in heparinase I activity and substrate specificity, and in binding to calcium. Section 5.2 describes the rationale for the site-directed mutagenesis experiments and summarizes the results from these experiments. The results are discussed in detail in section 5.3.

5.2 Results

5.2.1 Heparin Binding Domain of heparinase I—strategy for site-directed mutagenesis studies

The heparin binding site of heparinase I contains two putative heparin binding consensus sequences XBBBXBX and XBBXBX, where B represents a basic residue and X represents a hydropathic residue. Segment I span residues 197-204 (FKKNIAHD) and segment II spans residues 207-212 (EKKDKD). Segment I does not strictly adhere to the consensus sequence since asparagine-200 is not a basic residue. In addition, another region of heparinase I, spanning residues 331-337 (KNKKPQK), contains a basic cluster satisfying the consensus sequence.

In order to map the residues critical for heparin binding, the overall positive charge in the heparin binding sequences was altered, which would presumably alter heparin binding since heparin is negatively charged **(Figure 5.2)**. Hence, positively charged residues from both consensus sequences of the heparin binding site were individually altered to alanines (neutral) and aspartic acids (negative charge). This reduces the overall positive charge and hence reduces heparin binding affinity for the mutant enzymes. The lysines of both domains were altered to arginines (positive charge) as a control and to examine the effect of a bulkier positively charged side group.

In addition to altering the positive charges in the consensus sequences, other lysine residues in the heparin binding site in proximity to the consensus sequences were modified to alanines to probe the functionality of these residues within the heparin

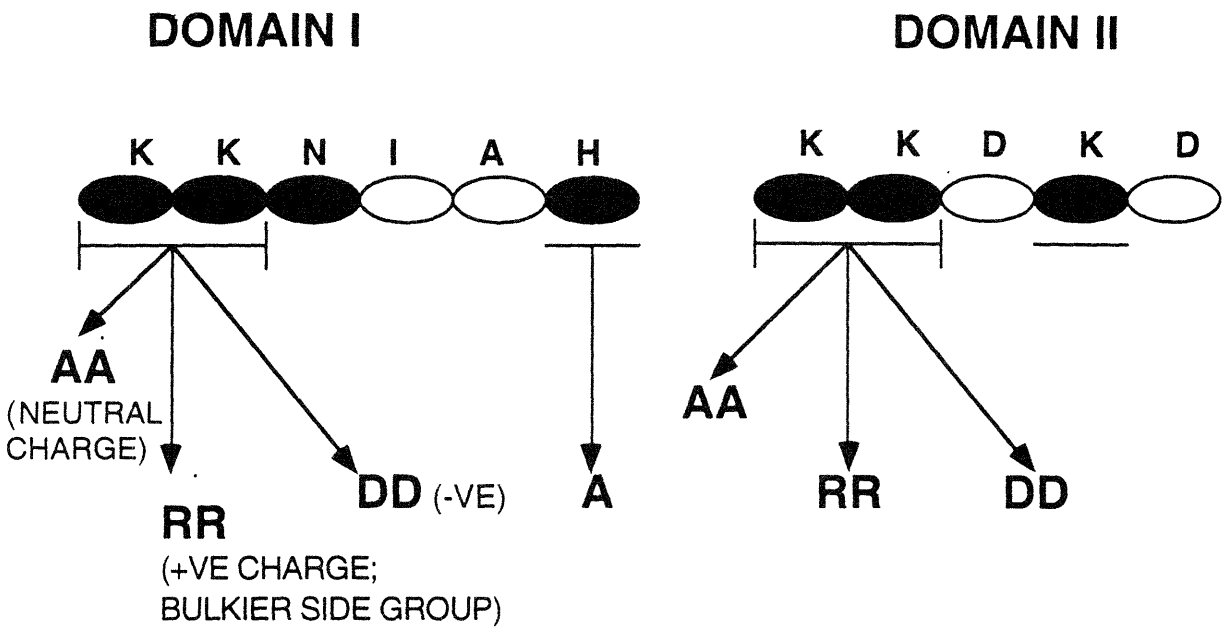


Figure 5.2: Site-directed mutagenesis of the consensus segments of heparin binding site of heparinase I—rationale.

binding site. Further, since tryptic peptide 4 (which contains the active site cysteine-135) was shown to bind heparin (chapter 2), the role of positively charged residues lysine-132 and arginine-141 contained in that peptide, was examined by alanine mutagenesis. Finally, lysines 333 and 334 from the other consensus sequence were altered to alanines.

All the mutant recombinant heparinases were constructed in pET expression system (Novagen, WI) and expressed in the BL21(DE3) host. *r*-heparinase I construct devoid of the putative signal sequence (-L *r*-heparinase I) was expressed as a control (Sasisekharan et al., 1993). The level of protein expression for all the recombinant heparinases was identical in the BL21(DE3) host. The purity of *r*-heparinases as determined by SDS-PAGE is estimated to be 80-90% after the first step and greater than 98% with silver stain after the second step **(Figure 5.3)**. Kinetic parameters, k_{cat} and K_m, were determined for all the mutants as described in chapter 3. The product profiles from overnight heparin degradation reactions were characterized using anion-exchange HPLC (POROS) to resolve the oligosaccharide products.

5.2.2 *Segment I of the heparin binding site*

When lysine-198 and lysine-199 of consensus segment I of the heparin binding site were altered to alanines (K198AK199A) the enzyme activity for the double mutant was drastically affected, with a 15-20 fold reduction in k_{cat}. **Table 5.1** lists the kinetic parameters obtained for the enzymes with mutations in consensus segment I. A similar effect on heparinase I activity was observed for the K198DK199D double mutant.

Table 5.1 Kinetic parameters of segment I mutations

Mutation	k_{cat} (sec^{-1})	K_m (μM)
-L *r*-heparinase I	92	10.2
K198AK199A	4.7	5.4
K198DK199D	4.9	4.2
K198RK199R	90	8.4
N200A	91	8.9
N200K	ND[a]	ND

a = not determined due to very little activity

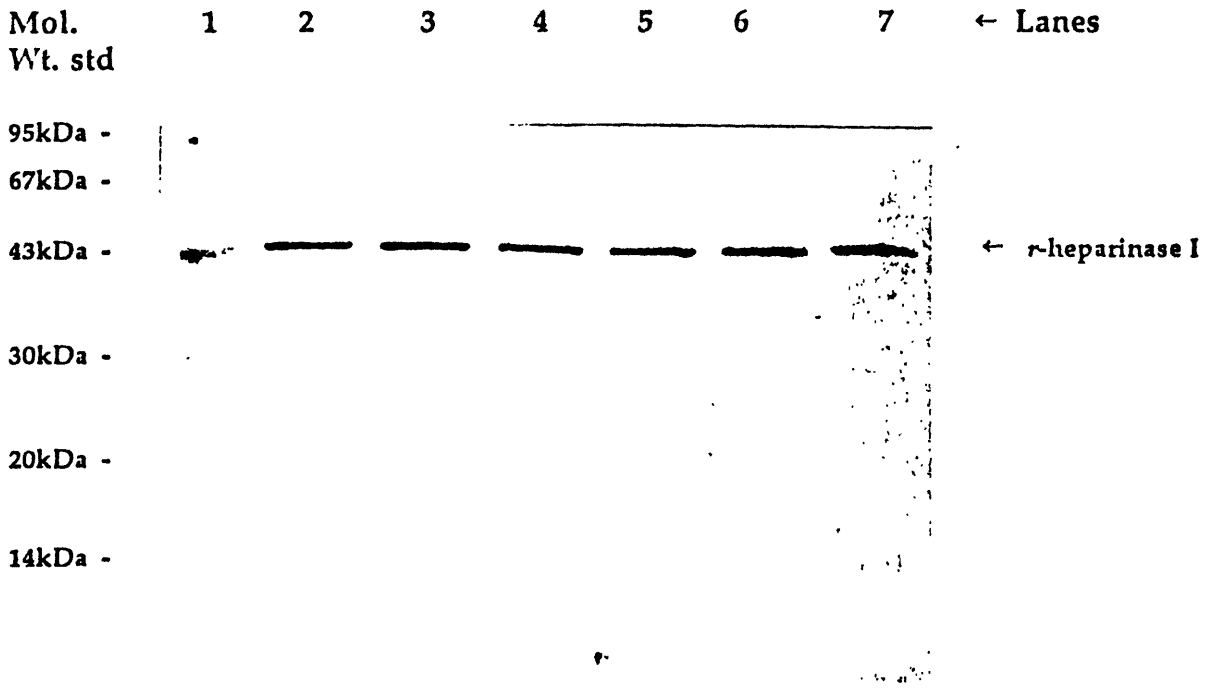


Figure 5.3: *SDS-PAGE of r-heparinase I and segment I mutant r-heparinase I:* A 12% SDS-PAGE of heparinase I. Lanes are described from left to right. **Lane 1:** Molecular weight standards. **Lane 2:** Ni column purified K198AK199A mutant *r*-heparinase I. **Lane 3:** Ni column purified K198DK199D mutant *r*-heparinase I. **Lane 4:** Ni column purified K198RK199R mutant *r*-heparinase I. **Lane 5:** Ni column purified N200A mutant *r*-heparinase I. **Lane 6:** Ni column purified N200K mutant *r*-heparinase I. **Lane 7:** Ni column purified *r*-heparinase I.

When the heparin degradation reactions were allowed to go to completion and the products were resolved using anion-exchange HPLC, the product profiles for both mutant enzymes seemed slightly altered when compared to the control -L *r*-heparinase I (wild type recombinant heparinase I) product profile **(Figure 5.4 a, b, c)** (see discussion). Interestingly, when the lysines-198 and 199 were modified to arginines, the enzyme activity remained unaltered as seen by the k_{cat} value **(Table 5.1)** and the product profile **(Figure 5.5 a, b)**.

As described in chapter 4, when histidine-203 from segment I was altered to an alanine the enzyme was rendered inactive. However, a H203D mutation resulted in residual enzyme activity with 30-40 fold reduction in kcat, suggesting a nucleophilic role in the catalytic mechanism for histidine-203. Since the K198AK199A double mutant retained activity, it is unlikely that these residues are directly involved in catalysis since alanine is a non-functional residue. Further the fact that the K198RK199R mutation resulted in the enzyme activity being unchanged suggested that the positive charge of these residues are important for heparinase I activity.

The above results taken together indicated that lysine-198 and lysine-199 from the Cardin and Weintraub consensus segment I are most likely involved in substrate binding, and the histidine-203 from this segment plays a role in catalysis. This strongly suggested that lysines-198 and 199, histidine-203 and cysteine-135 are in close proximity in the folded, three-dimensional structure of heparinase I.

The N200A mutation resulted in the enzyme activity to be unchanged **(Figure 5.6 a; Table 5.1)** suggesting that this residue had no functional value. Since the positive charge of lysines-198 and 199 were important for heparinase I activity, the effect of increasing the overall charge in the segment was investigated by creating a N200K mutation (which strictly adheres to the Cardin and Weintraub consensus sequence). Interestingly, the enzyme activity was drastically diminished **(Figure 5.6 b)**. This strongly suggested that introducing a bulky group such as lysine could possibly hinder heparin binding in the substrate binding cleft of heparinase I. N200A is an isosteric mutation and hence the results seem to indicate that steric considerations are perhaps more significant at this position rather than functional and charge considerations (see discussion).

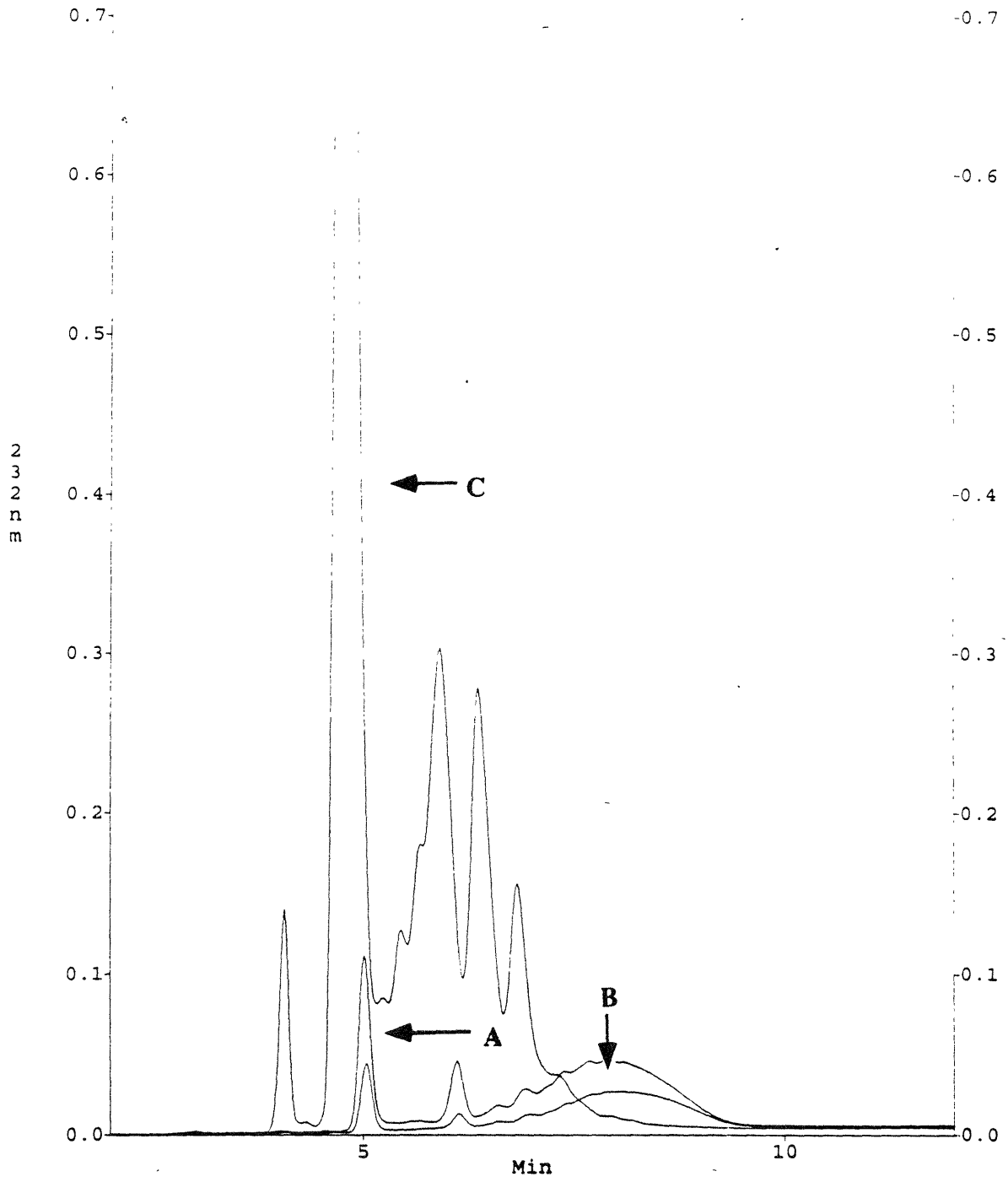


Figure 5.4: *Effect of segment I mutations on activity of heparinase I*
A shows the product profile of heparin degradation by the mutant K198AK199A ; **B** shows the product profile of heparin degradation by the mutant K198DK199D; **C** shows the product profile of heparin degradation by *r*-heparinase I

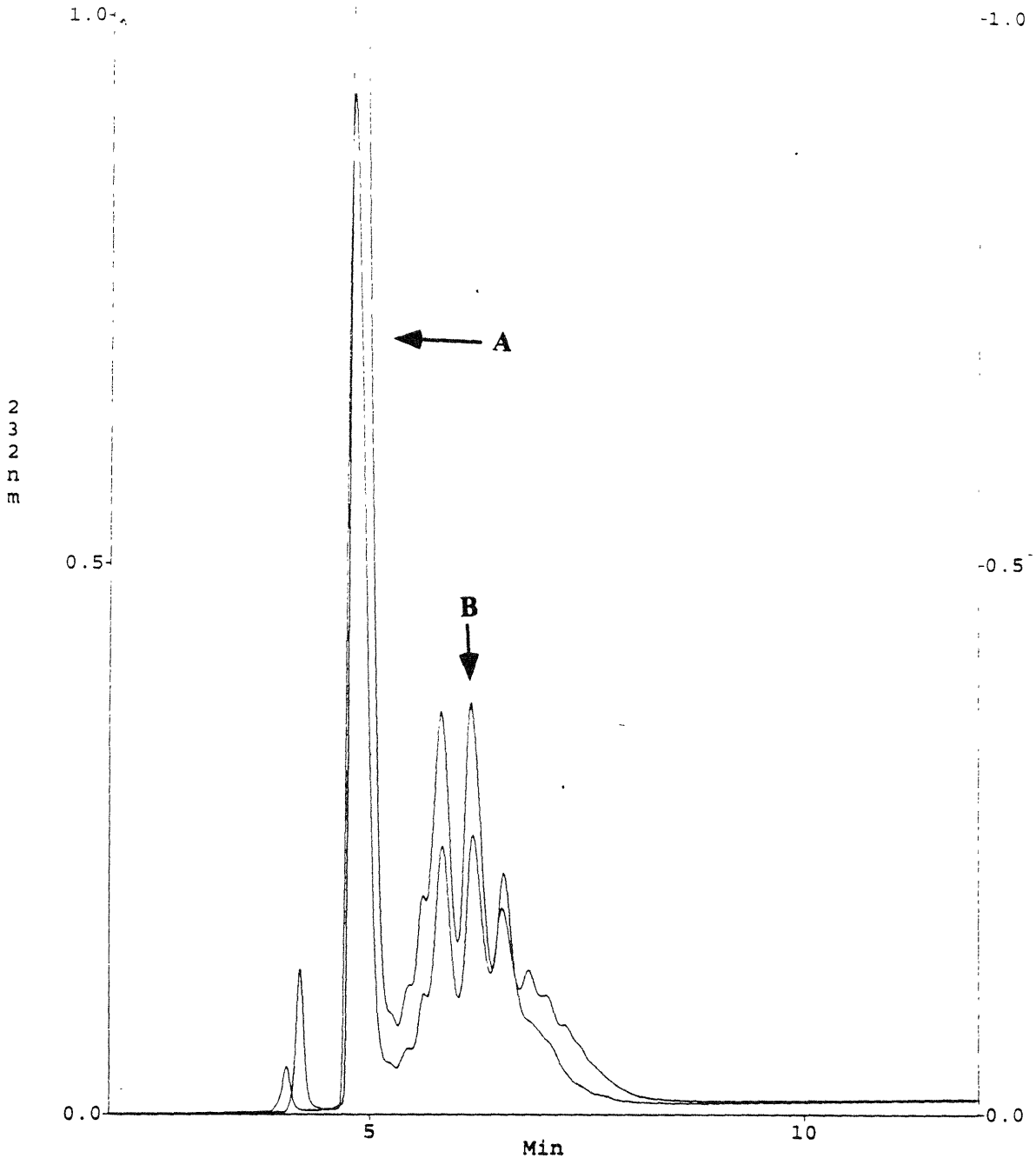


Figure 5.5: *Effect of segment I mutations on activity of heparinase I* **A** shows the product profile of heparin degradation by the mutant K198RK199R ; **B** shows the product profile of heparin degradation by *r*-heparinase I.

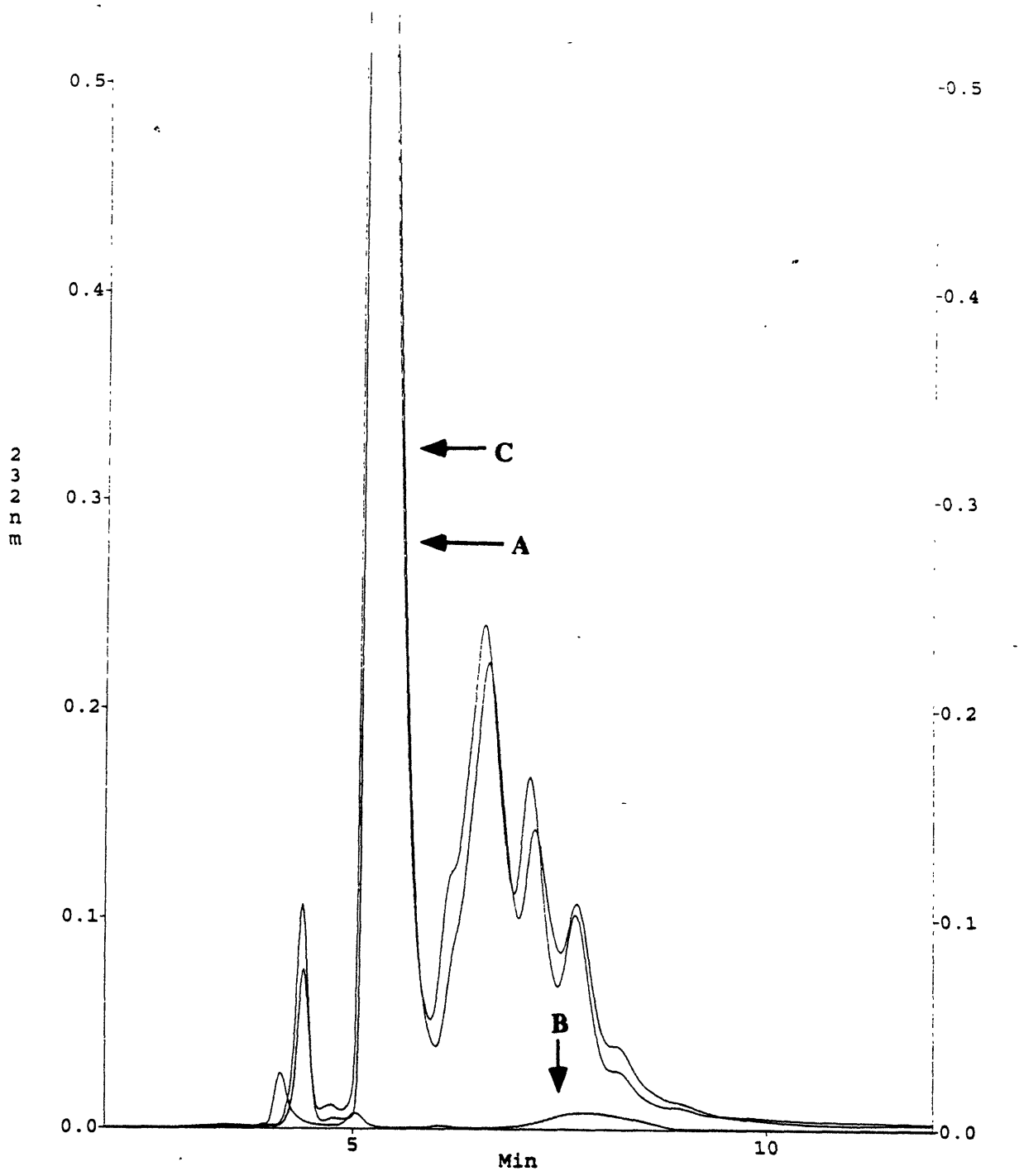


Figure 5.6: *Effect of segment I mutations on activity of heparinase I* **A** shows the product profile of heparin degradation by the mutant N200A ; **B** shows the product profile of heparin degradation by the mutant N200K; **C** shows the product profile of heparin degradation by *r*-heparinase I.

5.2.3 *Segment II of the heparin binding site*

Similar studies as described above were conducted on the other consensus segment in the heparin binding site (KKDKD). Interestingly, when the lysines 208 and 209 were altered to alanines the enzyme activity was only marginally affected as seen in **Table 5.2.** However, the product profiles of the K208AK209A as well as the K208DK209D and K208RK209R double mutants were identical to the wild type enzyme **(Table 5.2; Figure 5.7 a, b, c, d)**. Further, the K211A mutation also did not affect heparinase I activity and product profile **(Figure 5.8)**. Thus, positive charge in segment II of the heparin binding site does not seem to play a role in heparinase I activity. However, this region forms a part of the putative calcium binding loop (see below) and it is possible that it does not participate in heparin binding.

Table 5.2 Kinetic parameters of segment II mutations

Mutation	k_{cat} (sec^{-1})	K_m (μM)
-L *r*-heparinase I	92	10.2
K208AK209A	62	8.6
K208DK208D	90	9.2
K208RK209R	94	9.1
K211A	89	9.8

5.2.4 *Other lysine mutations*

Two other mutations were made to alter the positive charge in proximity to the consensus sequences of the heparin binding site. These two mutations, K205A and K214A, did not alter the activity and the product profile of heparinase I **(Table 5.3; Figure 5.9 b, c)**. Further, the K333AK334A double mutation resulted in unchanged heparinase I activity and product profile, suggesting that this basic cluster is not critical for substrate binding and heparinase I activity **(Table 5.3; Figure 5.9 a)**.

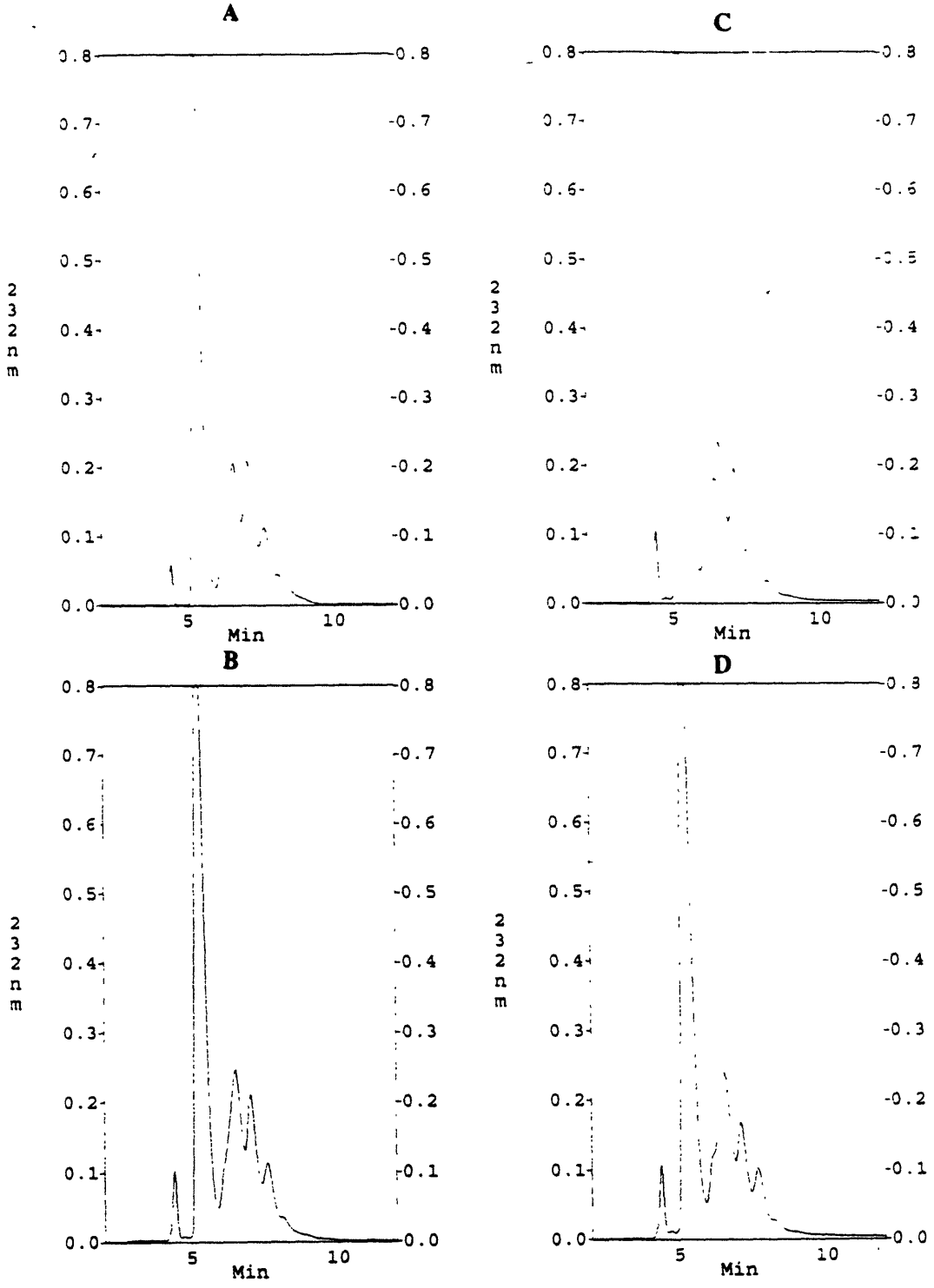


Figure 5.7: *Effect of segment II mutations on activity of heparinase I* **A** shows the product profile of heparin degradation by the mutant K208AK209A; **B** shows the product profile of heparin degradation by the mutant K208DK209D; **C** shows the product profile of heparin degradation by the mutant K208RK209R; **D** shows the product profile of heparin degradation by *r*-heparinase I.

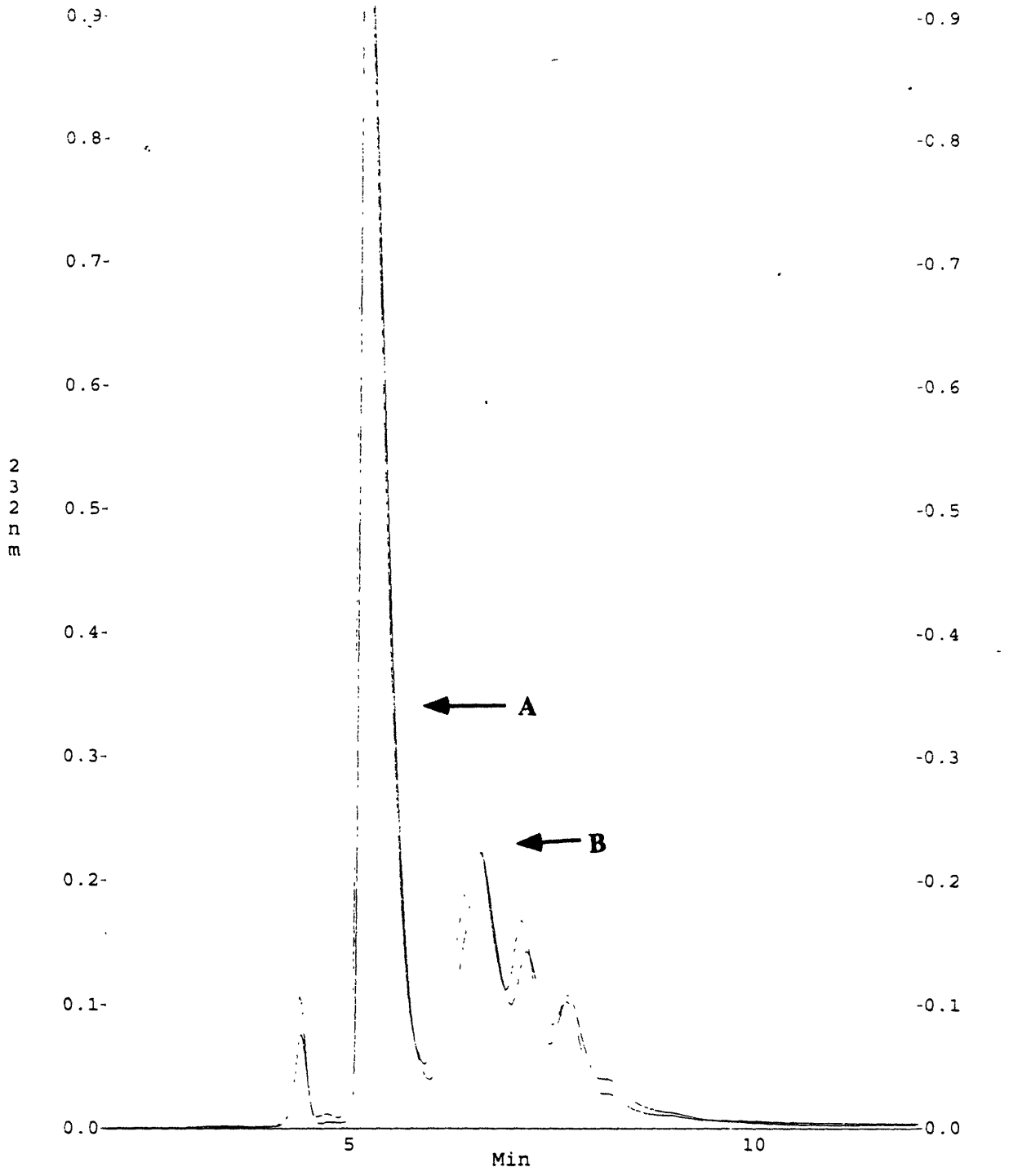


Figure 5.8: *Effect of segment II mutations on activity of heparinase I* **A** shows the product profile of heparin degradation by the mutant K211A; **B** shows the product profile of heparin degradation by *r*-heparinase I.

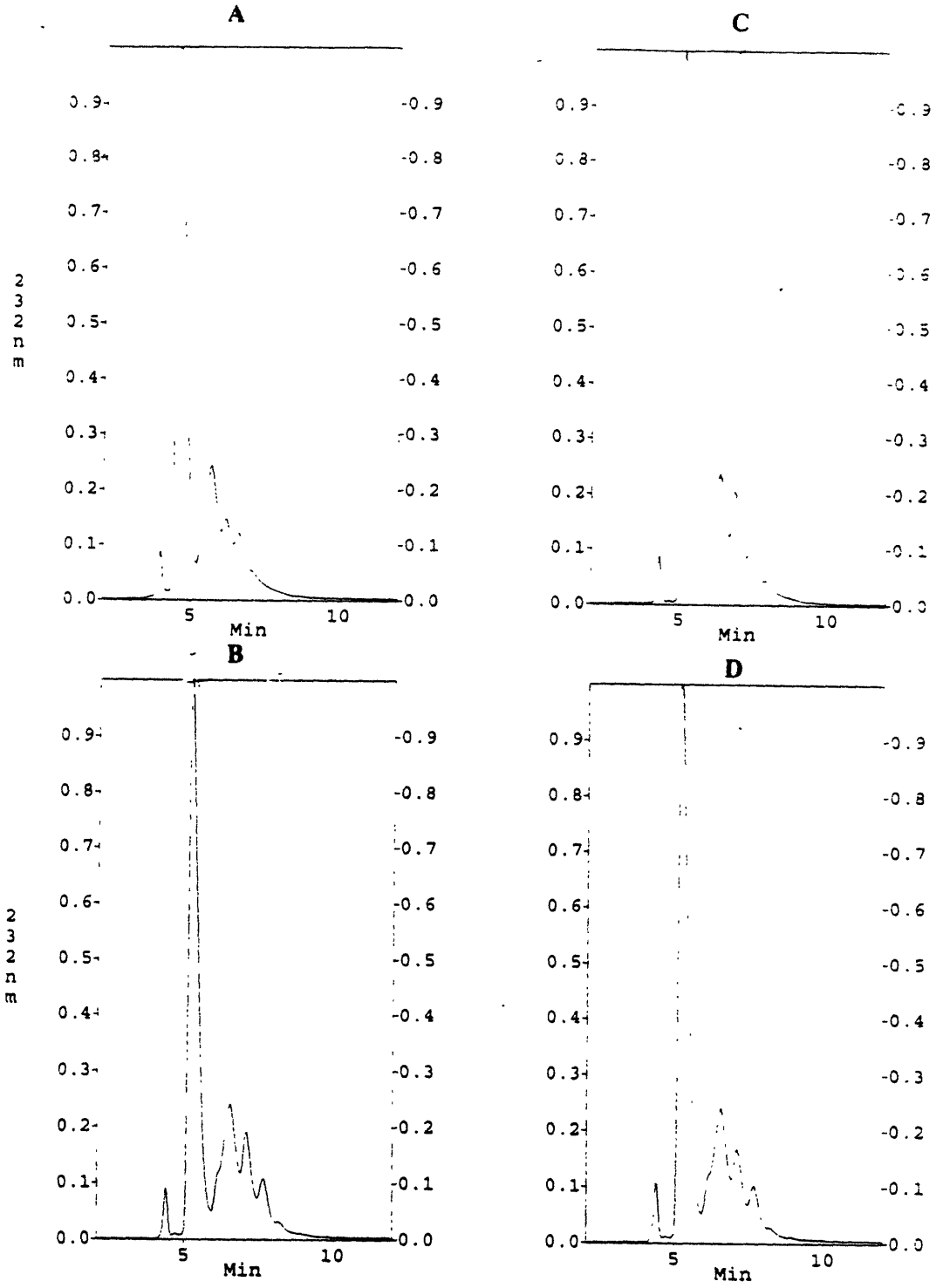


Figure 5.9: *Effect of other lysine mutations on activity of heparinase I* **A** shows the product profile of heparin degradation by the mutant K333AK334A ; **B** shows the product profile of heparin degradation by the mutant K214A; **C** shows the product profile of heparin degradation by the mutant K205A; **D** shows the product profile of heparin degradation by *r*-heparinase I.

Table 5.3 Kinetic parameters of other positive charge mutations

Mutation	k_{cat} (sec^{-1})	K_m (μM)
-L *r*-heparinase I	92	10.2
K214A	90	9.2
K333AK334A	88	9.4
K132A	11	8.1
R141A	48	7.8
K205Y	ND[a]	ND
K205A	82	8.7

a = not determined due to very little activity

5.2.5 *Tryptic peptide 4 (Td 4) mutations*

As pointed out in chapter 2, it was shown through tryptic digests and heparin competition experiments (Venkataraman, 1992), that tryptic peptide td4, containing cysteine-135 binds to heparin. Hence the role of two positively charged residues lysine-132 and arginine-141 in heparinase I activity was investigated. Interestingly the K132A and the R141A mutations did affect heparinase I activity, with a 10 fold reduction in k_{cat} in the case of K132A mutation **(Table 5.3)**, and produced diminished products **(Figure 5.10 a)**. Although the effect of these mutations is not as drastic as the K198K199 double mutations, the results suggest that positive charge of these residues is also important for heparinase I activity or heparin binding. The site-directed mutagenesis data are consistent with the previous observation of td 4 binding to heparin (Venkataraman, 1992) and consistent with the concept of a three-dimensional heparin binding domain rather than a linear basic cluster of residues (see discussion).

5.2.6 *The calcium binding site of heparinase I*

The heparin binding site of heparinase I contains a calcium binding consensus sequence (residues 207-213) of the 'EF-hand' structural domain. Four out of the required five calcium co-ordinating residues and the glycine were conserved in this site **(Figure 5.11)**. To examine the role of these residues, they were each individually

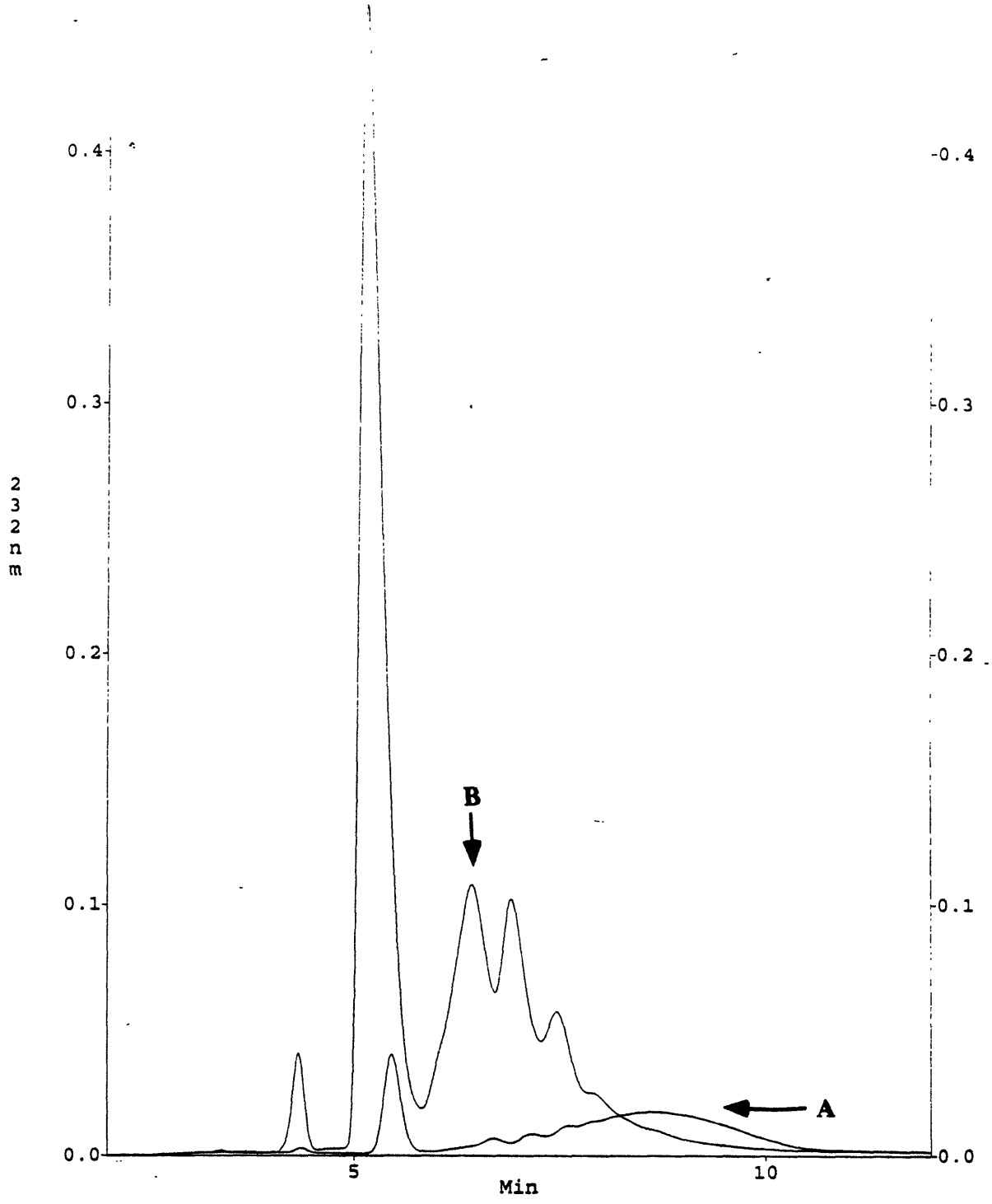
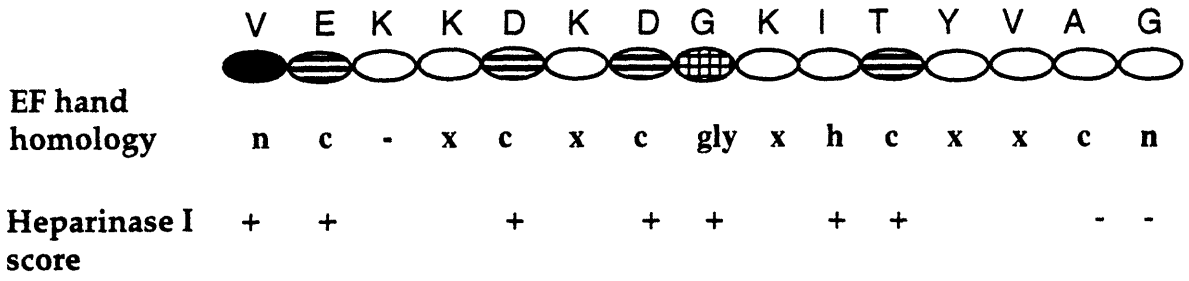


Figure 5.10: *Effect of Td4 mutations on activity of heparinase I*
A shows the product profile of heparin degradation by the mutant K132A ; **B** shows the product profile of heparin degradation by *r*-heparinase I.

Calcium Coordinating Motif of Heparinase I



= n = non-polar amino acid (I, L, V, M, F, Y, or W)

= c = calcium co-ordinating amino acids (D, E, Q, N, S, or T)

= x = any amino acid

= gly = conserved glycine residue

= h = hydrophobic residue (I, L, or V)

Figure 5.11: The calcium coordination consensus motif of heparinase I. In additon to the side-chain oxygen containing residues, the glycine and the isoleucine are also conserved.

altered to alanines and the effect on heparinase I activity was studied. The D210A and the D212A mutations did not seem to affect the enzyme activity **(Figure 5.12 a, b)** (see discussion). However, the G213A mutation resulted in a 5 fold reduction in k_{cat} **(Table 5.4)** and diminished products as shown in **Figure 5.13 a**. E207A also resulted in lower heparinase I activity with a 3 fold reduction in k_{cat}; however the product profile was not significantly altered **(Figure 5.12 c)**.

Table 5.4 Kinetic parameters of the calcium binding motif mutations

Mutation	k_{cat} (sec^{-1})	K_m (μM)
-L *r*-heparinase I	92	10.2
D210A	88	9.2
D212A	90	9.2
E207A	32	9.4
G213A	20	9.8

Thus the data suggest that the conserved glycine-213 plays an important role in calcium coordination through its conformational flexibility (see discussion). Hence the consensus motif could possibly play a role in calcium binding. Interestingly, the COOH terminus of heparinase I (residues 372-383) was also homologous to the calcium co-ordinating loop of the 'EF-hand'. Hence a G378A mutation was made but this resulted in unchanged heparinase I activity.

5.2.7 The substrate binding pocket of heparinase I—A model

The results from the above site-directed mutagenesis experiments give a clear picture of the catalytic domain with reference to the substrate binding pocket of heparinase I. This model is depicted in **Figure 5.14**. The primary heparin binding site is in proximity to the catalytic domain which in turn is contained in the heparin binding domain—thus together forming the substrate binding cleft of heparinase I which possibly contains calcium embedded in the cleft, bound to the calcium binding site (see discussion, section 5.3).

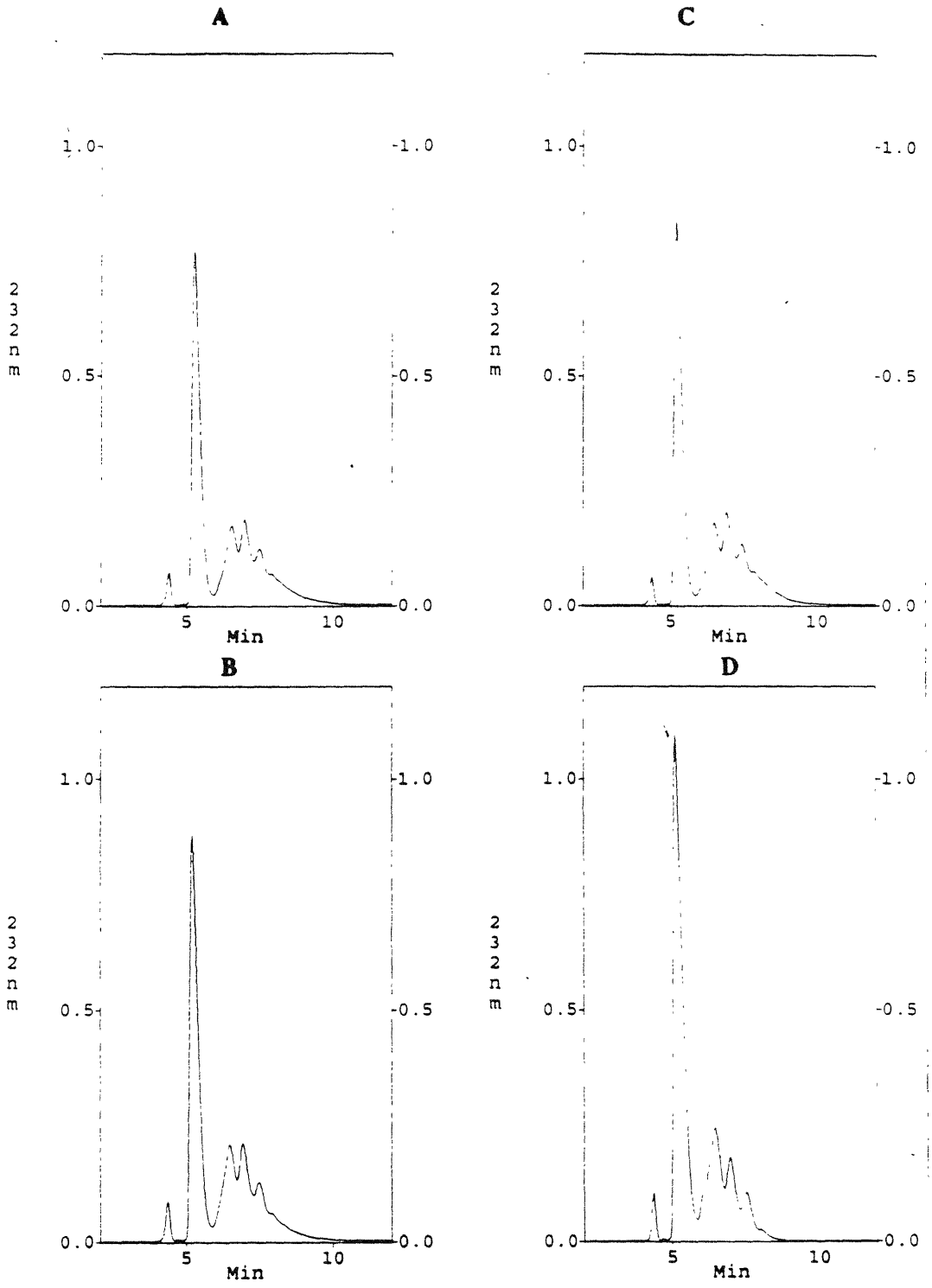


Figure 5.12: *Effect of calcium coordinating motif mutations on activity of heparinase I*

A shows the product profile of heparin degradation by the mutant D210A ; **B** shows the product profile of heparin degradation by the mutant D212A; **C** shows the product profile of heparin degradation by the mutant E207A; **D** shows the product profile of heparin degradation by *r*-heparinase I.

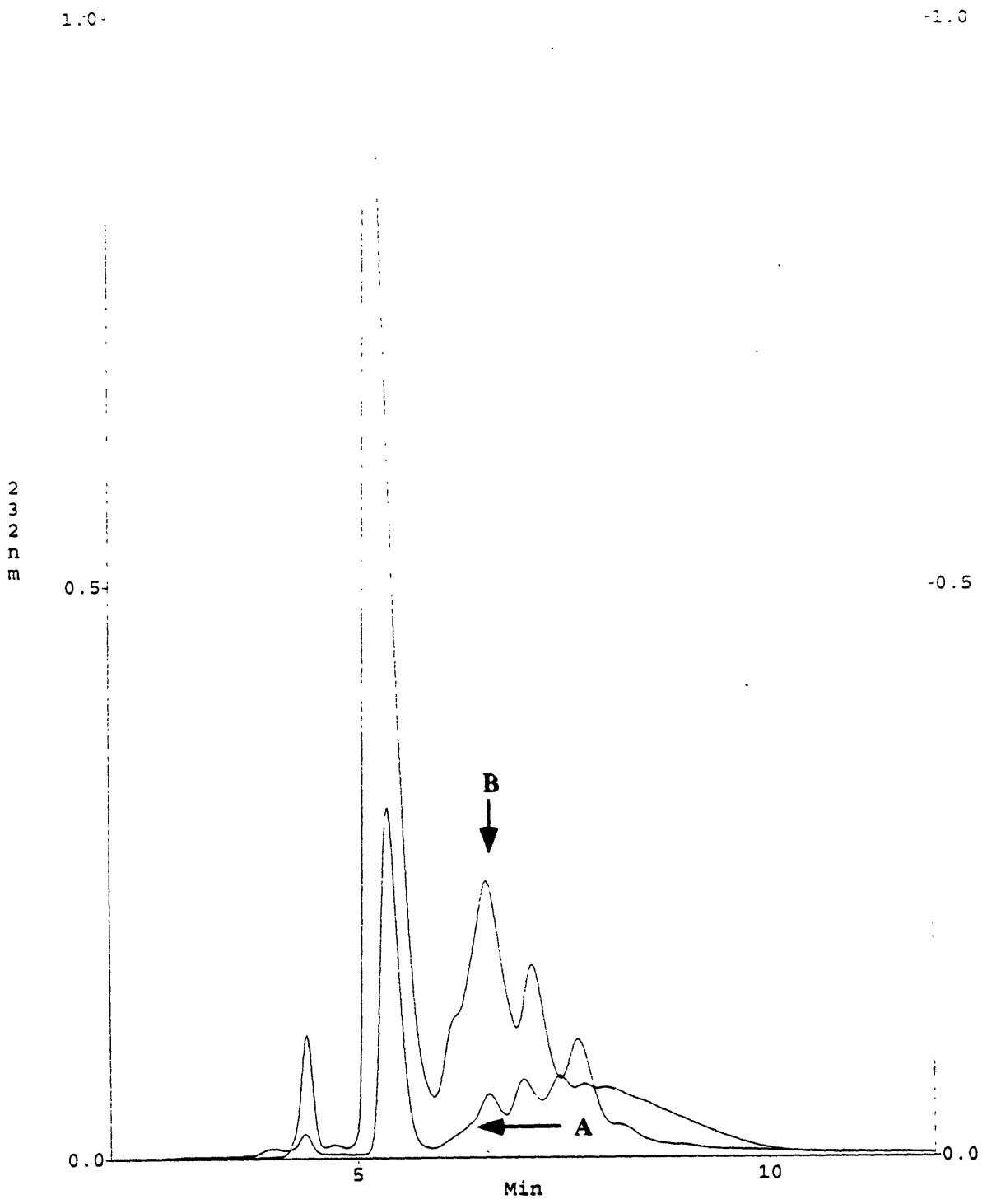


Figure 5.13: *Effect of calcium coordinating motif mutations on activity of heparinase I*
A shows the product profile of heparin degradation by the mutant G213A; **B** shows the product profile of heparin degradation by *r*-heparinase I.

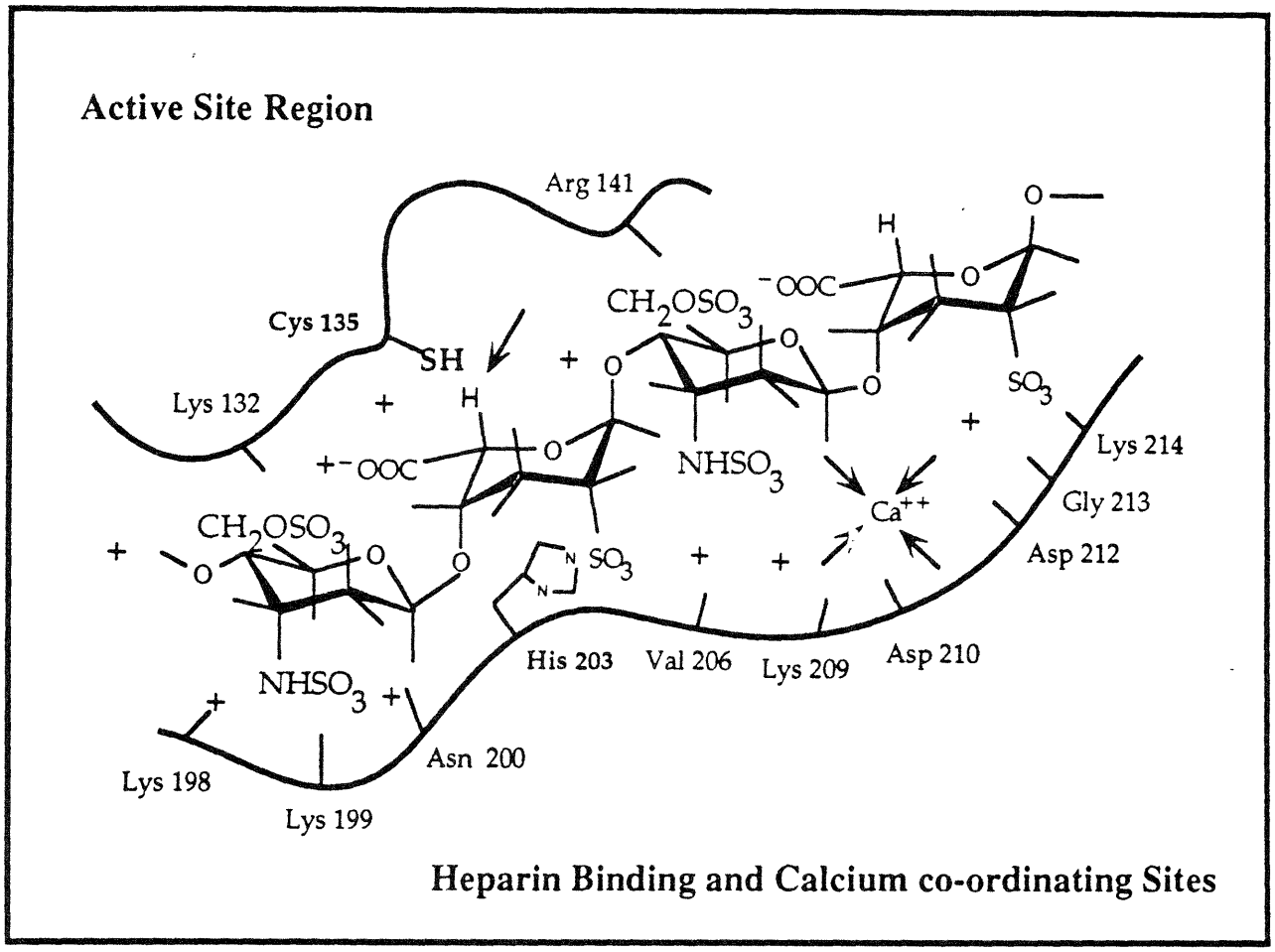


Figure 5.14: *Schematic representation of the substrate binding pocket of heparinase I.* The figure shows a heparin tetrasaccharide (about 0.165 nm in length, Nieduszynski, 1985) bound to the 18 amino acid heparin binding site. It is interesting to note that several secondary structure prediction algorithms suggests that the 18 amino acids heparin binding site contained the main α-helix forming potential in heparinase I, which is otherwise dominated by β-sheet forming potential (Sasisekharan, 1991; Yoder, et al., 1993a). It is hypothesized that the core fold of heparinase I is dominated by β-sheets (as seen in pectate lyase; Yoder, et al., 1993a) and that the region around 196-221 forms a large loop region with perhaps a partial α-helical secondary structure. Cysteine-135 and histidine-203 are in close proximity to each other and to the hydrogen (of the iduronate) marked by an arrow. Flanking these residues are the heparin binding consensus sequence 1 (lys 198, lys 199...), the calcium co-ordinating site (val 206 - gly 213) and K132, R141, which form the heparin binding domain of heparinase I.

5.2.8 *Model prediction: K205Y—a steric mutation*

So far the focus was on probing the functional roles of individual residues in heparin and calcium binding, to isolate critical functional residues. This systematic analysis resulted in developing the above model shown in **Figure 5.14**. Finally, the model of the substrate binding cleft was tested by introducing a steric mutation. The N200K mutation described above proved to be sterically unfavorable but this was part of the heparin binding consensus sequence. Since the K205A mutation did not affect the enzyme activity, this residue was not functionally important. Hence, this was chosen to introduce a bulky group, such as tyrosine. This would either cause steric hindrance to heparin access or interfere with the functioning of catalytically critical histidine-203 as per the model. Indeed, the K205Y mutation drastically affected heparinase I activity with barely noticeable activity **(Figure 5.15)**, consistent with the model of the substrate binding cleft.

5.3 Discussion

5.3.1 *Segment I mutations*

Basic residues in three heparin binding consensus sequences and in other specific regions of heparinase I were altered in a systematic fashion using a site-directed mutagenesis approach. Results presented here strongly suggest that lysine-198 and lysine-199 from the consensus segment I play a significant role in heparin binding and as a result in modulating the enzyme activity. K198AK199A and K198DK199D mutations resulted in substantially reduced enzyme activity and the product profile seemed to be altered when compared to the wild type enzyme. A significant portion of higher order fragments seemed to be present. It will be interesting to see if the fragments eluting between 7 and 9 minutes **(Figure 5.4 a, b)** represent unique oligosaccharide sequences not produced thus far. However, it is possible that the mutations result in a slower enzyme and hence the mixture could represent larger, undigested fragments of heparin.

These enzymes will be useful tools in understanding the kinetics of the cleaving reaction since the product profiles of slower enzymes will provide information about products that appear later in the reaction. For example, when comparing the

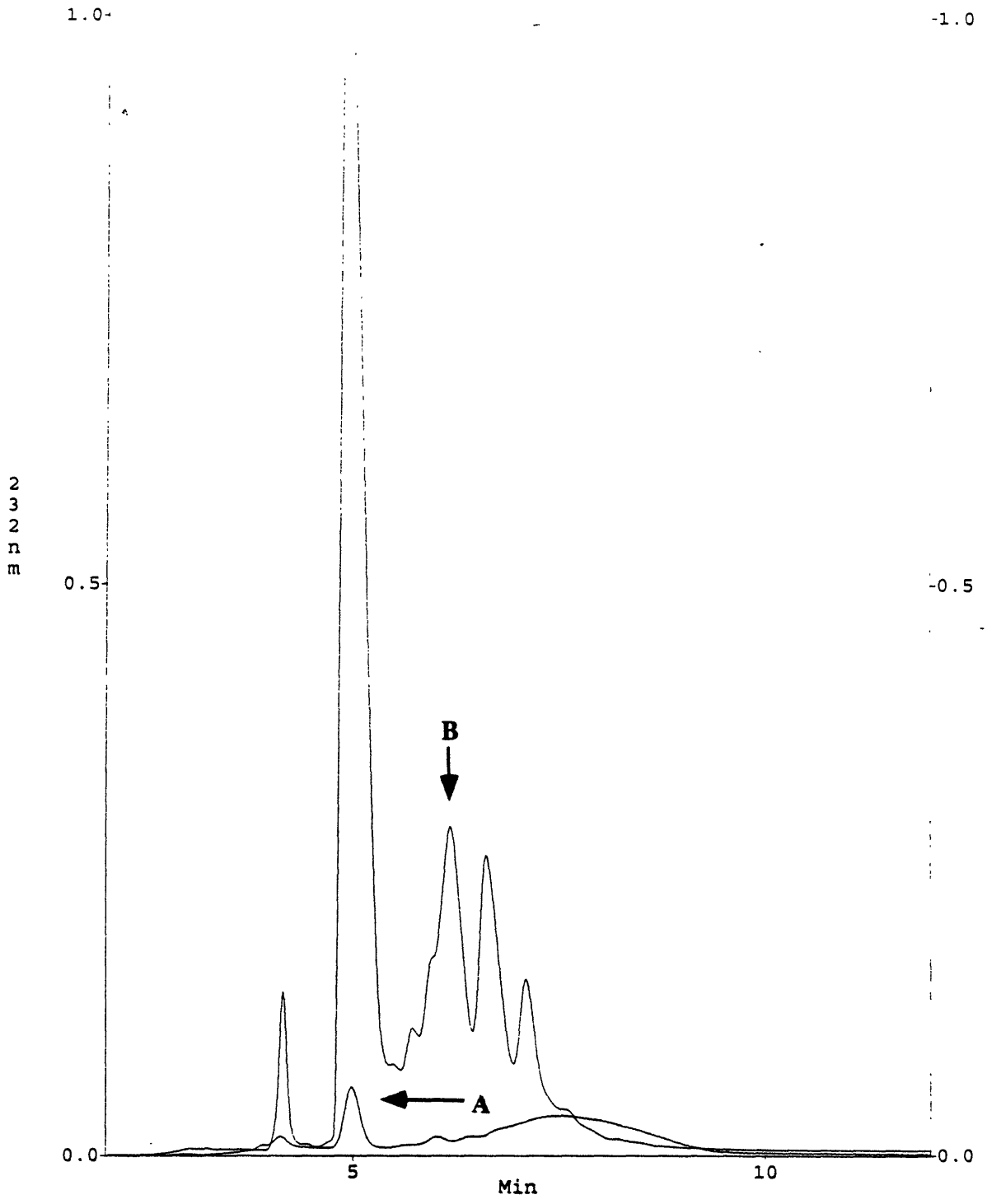


Figure 5.15: *Model Prediction: Effect of a steric mutation on activity of heparinase I*
A shows the product profile of heparin degradation by the mutant K205Y; **B** shows the product profile of heparin degradation by *r*-heparinase I.

product profiles of the wild type enzyme and K198AK199A, it appears that the major disaccharide (tri-sulfated disaccharide) is one of the earlier products and the disulfated disaccharide appears later in the reaction since it is barely present in the product profile of the mutant enzyme.

More importantly, maintaining the positive charge by altering the lysines to arginines did not alter enzyme activity, indicating that the amines or the positive charge in segment I is important for enzyme activity. Previous immobilization studies with heparinase I using amine chemistry resulted in extensive activity losses, suggesting that amines are crucial for enzyme activity (Comfort et al., 1989). Data presented here are consistent with this observation since it is possible that lysines-198 and 199 could have been modified in the earlier study.

The N200K mutation, while maintaining the consensus sequence, increased the overall positive charge density and hence one would have expected improved substrate binding and catalysis. However, interestingly this mutation resulted in drastically reduced enzyme activity. It is possible that introducing a bulkier residue such as lysine sterically hinders heparin access to the active site, since the N200A mutation, with a smaller residue, did not alter enzyme activity. Another possibility is that the bulkier side group of lysine could either interfere with the functioning of lysine-199 or histidine-203, which as described in chapter 4 is involved in the catalytic mechanism. This suggests that asparagine-200 is sterically, and not functionally, important and the spatial orientation of histidine-203 relative to lysines-198 and -199 seems important.

5.3.2 Segment II and other positive charge mutations

Modifying lysines-208 and -209 of consensus segment II (residues 207-210) to charged residues such as aspartic acids or arginines, did not affect heparinase I activity, suggesting that this sequence does not play a role in heparin binding or enzyme activity. Interestingly this region is part of the consensus calcium binding loop. The K208AK209A mutation seemed to exhibit slightly lower activity (2 fold reduction in k_{cat}. It is possible that a charged residue is favored at this position to stabilize the helix dipole in the calcium binding motif (Strydnadka and James, 1989).

The effect of other basic residues in the primary heparin binding site (lysine-205 and lysine-214) near each of the two consensus sequences was investigated. The K205A and K214A mutations did not affect heparinase I activity. In addition to the two basic clusters in the primary heparin binding site, the heparinase I primary sequence contains another basic cluster close to the C-terminus (residues 332-338). This region also did not seem to be implicated in heparin binding since the K333AK334A double mutant resulted in unchanged enzyme activity.

As described in chapter 2, an interesting observation in the earlier study (Venkataraman, 1992) was that the tryptic peptide containing cysteine-135 bound heparin. Hence the role of the two positively charged residues within this peptide was investigated. The K132A and R141A mutations resulted in reduced heparinase I activity suggesting that these basic residues are perhaps important. This points to the importance of a three-dimensional binding domain of basic residues rather than just a linear sequence of positively charged residues of a consensus sequence.

5.3.3 Calcium coordination motif

The description of a calcium binding domain is based on x-ray crystallographic studies on paralbumin and other proteins (Gross et al., 1993). It consists of a 30 residue helix-loop-helix structure with calcium bound to residues in the loop amino acid sequence. Calcium is ligated within the twelve residue loop of the motif, coordinating to 3-5 side-chain oxygen containing groups at positions 1, 3, 5, 9, and 12 (Marsden et al., 1990). A survey of the literature for calcium binding proteins, to study the frequency with which each particular amino acid residue is found at a specific position, showed that at ligand positions at 1, 3, 5, and 12 and non ligand positions 6 (glycine) and 8 (isoleucine) the residues are most conserved (Marsden et al., 1990). For heparinase I, three ligand positions and the non ligand glycine and isoleucine are conserved **(Figure 5.11)**. Further, loops with two acidic ligand residues are expected to have low affinities for calcium. Glycine is conserved at position 6 for reasons of its conformational flexibility. In calcium bound loops of the crystallographic structures observed, the conformation of the glycine residue [(ϕ, ψ = 90°, 0°) is one that is not allowed for any other residue (Herzberg and James, 1985).

The reduced activity for G213A and E207A mutations from the putative calcium binding site indicated that this loop could be indeed involved in calcium co-ordination. EF-hand loops with less than 3 oxygen-containing residues in the conserved positions have very low affinities for calcium. This is consistent with the observation that the D210A or D212A mutations did not have an effect on activity since the mutant enzymes (incorporating point mutations) had 3 oxygen-containing residues at each time for calcium coordination. This is consistent with the observation that E207A did not have a drastic effect on heparinase I activity. One can speculate that creating a double mutant of E207A and either D210A or D212A would possibly affect calcium binding and hence heparinase I activity, since there would be less than 3 calcium binding residues.

5.3.4 The substrate binding pocket of heparinase I

Chapter 2 described the active site environment around cysteine-135 being positively charged, with the role of the positively charged residues being to activate the thiol group for catalysis by lowering its pKa (Sasisekharan et al., 1995). Subsequently a heparin binding site was identified and it was proposed that this provides the necessary charge complementarity for very specific heparin binding on the one hand, while on the other it serves in the active site environment which plays a key role in biasing the active site reactivity (Sasisekharan et al., 1996). Chapter 4 described the identification of histidine-203 as a critical residue for catalysis. An interesting observation from these studies is that histidine-203 is not only part of the consensus segment I of the heparin binding site but also critical for catalysis and hence cysteine-135 and histidine-203 form the catalytic domain (see chapter 4). Consistent with this hypothesis, data presented in this chapter demonstrated that the lysines of consensus segment I are important for substrate binding, thus indicating that positive charge from this segment possibly assists in heparin binding and the histidine from this segment is adjacent to the scissile linkage for catalysis.

Protein tyrosine phosphatases (PTPases) are important enzymes in the signal transduction regulatory pathway, cleaving highly anionic phosphate backbone structures. The active site environment is very similar to that observed for the heparinase I active site. Similarities include a highly reactive cysteine in the active

site in a positively charged environment; preferential reactivity towards IAA than IAM (Pot et al., 1991); and ionic strength dependence of cysteine inactivation with negatively charged compounds (Zhang and Dixon, 1993). Further, it is interesting to note that for PTPases, the active site cysteine is stabilized by surrounding positively charged residues and as a result causes a reduction in pKa of this residue (Zhang and Dixon, 1993). The data presented in this chapter for heparinase I suggested that lysines-198 and 199 could be involved in heparin binding as well as they could be required for stabilizing the thiol group of cysteine-135. These residues are in close proximity to cysteine-135 with histidine-203 close to the scissile bond for catalysis.

As a general mechanism for lyases, it has been proposed (Gacesa, 1987) that an important step in the catalytic mechanism is the neutralization of the negative charge developing on the carboxylate group of the uronate in heparin. This function is most likely performed by a lysine residue, and not normally by arginines or histidines, by the formation of a salt bridge (Gacesa, 1987). Although it is possible that the lysines-198 and 199 could be involved in this function, the K198RK199R double mutation resulted in unchanged heparinase I activity, suggesting no preference for lysines. In addition, basic residues involved in electrostatic interactions with heparin, would not participate in salt bridges (Hata et al., 1992). Hence the role of the positive charge seems to be not only to provide charge complimentarity for heparin but also to activate cysteine-135 by providing necessary charge stabilization.

It has been observed that in addition to linear heparin binding sequences, protein tertiary and quaternary structures may also control heparin binding (San Antonio et al., 1993). In the case of lactoferrin, positively charged residues from 2 different sequences form a unique structural motif that binds negatively charged glycosaminoglycans (Wu et al., 1995). Consistent with this observation, reduced enzyme activities for the K132A and R141A mutations were observed, although the effect was not as dramatic as seen for the K198AK199A mutation. It is possible that lysine-132 and/or arginine-141 are brought close to the heparin binding sequence in the correctly folded structure of heparinase I, and all these residues together constitute a heparin binding domain in heparinase I. Further, it is possible that these residues contribute to stabilizing the thiol group. It has been observed in the case of the PTPases that the catalytic domain consists of a CXXGXXRS/T signature

motif (Guan and Dixon, 1990). The cysteine is in the active site and the arginine plays a critical role in substrate recognition and transition state stabilization (Zhang et al., 1994). It is intriguing to note that the heparinase I primary sequence has CEQGSSRS, with the cysteine-135 essential for activity. While the R141A mutation did not result in as drastic an effect as was seen with the PTPase (Zhang et al., 1994), a conserved motif, with the positive charge in proximity to cysteines in both enzymes possibly plays similar roles. It should be pointed out that the substrates for both enzymes are highly negatively charged and these enzymes possibly share a similar active site structure or environment.

Another interesting feature of the heparin binding site is the presence of a putative calcium coordinating motif. Calcium is required for the catalytic activity of heparinase I and it is known that calcium binds to heparin (Boyd et al., 1980; Hunter et al., 1988), and causes a significant change in the helical structure of heparin, heparan sulfate, and other glycosaminoglycans (Nieduszynski, 1985). The involvement of glycine-213 and glutamate-207 in enzyme activity suggested that these residues could be involved in coordinating calcium at the active site and the calcium bound substrate is perhaps oriented for catalysis. The crystal structures of three pectate lyases, which are enzymes from the same family of acidic polysaccharide eliminases as heparinases, showed that the substrate binding cleft was embedded with one or more calcium ions (at the calcium binding loop of EF-hands), indicating a direct interaction between calcium and the enzyme (Pickersgill et al., 1994; Yoder et al., 1993; Yoder et al., 1993).

5.4 *Summary*

The 18 amino acid heparin binding site in heparinase I contains two clusters of positively charged residues similar to proposed heparin binding consensus sequences, and a segment homologous to a calcium co-ordinating motif of the 'EF-hand' structural domain. This chapter described the dissection of the heparin binding site to identify specific residues critical for heparin binding and calcium co-ordination through extensive site-directed mutagenesis experiments. More specifically, the role of the positive charge in the heparin binding site in substrate binding and enzyme activity was addressed.

The model of the substrate binding cleft of heparinase I is shown in **Figure 5.14**. It contains a catalytic domain comprising cysteine-135 and histidine-203 involved in the catalytic mechanism of the enzyme. The substrate binding is primarily dictated by lysines-198 and 199, which also form the positively charged environment around cysteine-135 and activate this residue for catalysis. Lysine-132 and arginine-141 could also play a role in either substrate binding or in providing positive charge for stabilizing the cysteine anion. Calcium binding to heparin could possibly introduce a conformational change and this is stabilized by oxygen-containing residues of the calcium binding site in heparinase I.

Finally, introducing a bulky group at positions 200 and 205 resulted in a drastic loss in enzyme activity, consistent with our model of the substrate binding cleft. Thus asparagine-200 and lysine 205 dictate the size of the substrate binding cleft as these positions can not accommodate bulky residues. This study has led to the definition of a heparin binding cleft containing the catalytic domain in heparinase I. The role of the individual catalytic residues in catalysis needed to be investigated and this is described in the following chapter.

Chapter 6

Heparinase I—A Mechanism for Heparin Degradation

The catalytic domain of heparinase I comprising cysteine-135 and histidine-203 had been identified. However, the individual roles of these residues in catalysis remained to be seen. In this chapter, through systematic pH analysis and site-directed mutagenesis studies the ionization states of the catalytic residues of heparinase I (when cleaving heparin) are identified. Results from pH-rate analysis show that cysteine-135 with a pKa of 5.5 is unprotonated and histidine-203 with a pKa of 8.2 is protonated for catalysis. The unusually low pKa for cysteine is attributed to the positive charge around it and this charged environment (basic residues from the heparin binding domain) perhaps stabilizes the anionic form of the thiol group. Site-directed mutagenesis studies on cysteine-135 and histidine-203 with residues with varying nucleophilicities suggest that cysteine-135 plays a nucleophilic role and histidine-203 acts as an acid/base catalyst. A mechanism for heparin degradation that is consistent with data, from both this research as well as previous work, is proposed.

6.1 Motivation

Chapter 2 discussed the hydrolases and the mechanisms of a few enzymes including the polysaccharide degrading hydrolase, lysozyme, extensively characterized by a variety of methods. Very little is known about polysaccharide degrading lyases in general and not much is known on the mechanism of enzyme action. Heparinase I is a lyase that cleaves heparin-like molecules via an elimination reaction. However, detailed mechanistic studies to elucidate structure-function relationships have not been undertaken with eliminases and as a result enzymatic degradation involving elimination mechanisms is not well understood

The previous chapter provided a descriptive model of the substrate binding pocket of heparinase I. A catalytic domain comprising cysteine-135 and histidine-203 is placed in a highly positively charged environment which is provided by lysine residues in heparin binding site and in proximity to cysteine-135. It was proposed that the role of the positive charge was to activate the cysteine-residue for catalysis and also to provide charge complimentarity for binding negatively charged heparin.

Histidine-203 could either play a nucleophilic role or as an acid/base catalyst in a 'proton relay' mechanism as is seen with several enzymes (Carter and Wells, 1988; Warshel et al., 1988; Nickbarg et al., 1988; DiPersio et al., 1991; Miran et al., 1991; Munier et al., 1992). Further, it remained to be seen if there were any other residues critical for catalysis. Hence, to develop a more detailed understanding of the catalytic mechanism employed by heparinase I, the ionization states of catalytically critical residues were investigated using pH effects on kinetics, chemical modification and site-directed mutagenesis approaches. The number of residues necessary for catalysis was established and the role of individual residues was examined. A catalytic mechanism which is consistent with these and previous observations is proposed.

6.2 Results

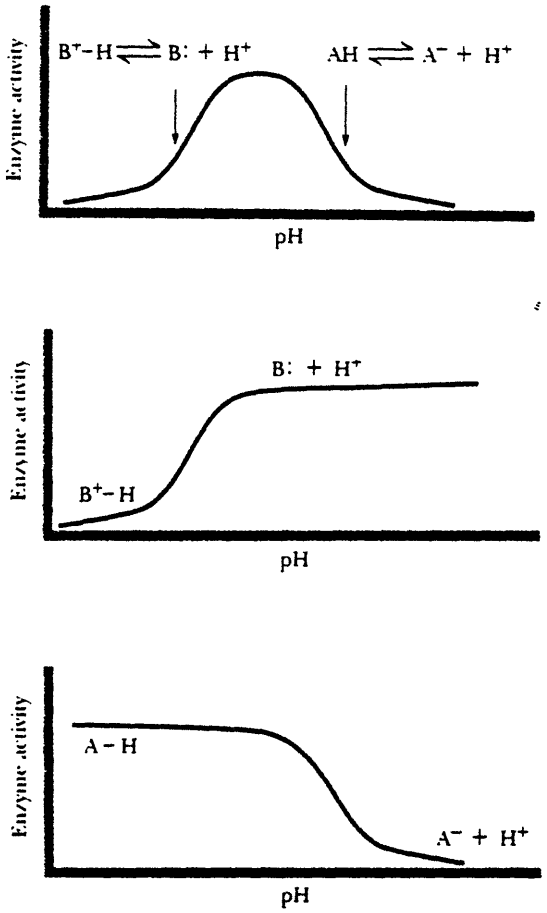
6.2.1 *Identification of critical ionizations for heparinase I using pH profile*

The first question addressed was the number of active site residues involved in the catalytic mechanism through the pH dependence of enzymatic activity. If there are only two active-site residues—one protonated and one ionized, then the pH versus enzyme activity profile is bell-shaped in nature **(Figure 6.1 a)**. If enzyme activity increases with an increase in pH, catalysis may depend on a deprotonated group that may act as a general base, accepting a proton from the substrate or a water molecule (Garrett and Grisham, 1995) **(Figure 6.1 b)**. Protonation of this group at lower pH prevents it from accepting another proton from either the substrate or water, for example. On the other hand, if the enzyme activity dramatically decreases with an increase in pH, the activity may depend on a protonated group that acts as a general acid to donate a proton to the substrate or water molecule **(Figure 6.1 c)**. Hence at high pH the group is ionized and a proton is unavailable for the catalytic reaction. Combining behaviors of figures 6.1b and 6.1c gives rise to figure 6.1a.

In the case of heparinase I, data indicated the involvement of cysteine-135 and histidine-203 in catalysis (chapters 2 and 4). To determine the functional roles of these residues in the catalytic mechanism it was necessary to establish their ionization states in catalysis by determining the pH dependence of enzymatic activity. Hence the kinetic parameter k_{cat} was determined as a function of pH. **Figure 6.2** shows the pH versus rate profiles for the wild type enzyme. The model used for fitting the data (using non-linear least square analysis) was (Denu et al., 1995):

$$k_{cat} = C/(1+H/Ka)(1+Kb/H)$$

C represents the pH independent value of k_{cat} and Ka and Kb are the ionization constants of catalytically critical residues. The experimental values of k_{cat} were in very good agreement with values predicted by the model and a bell shaped curve was obtained. Interestingly in the region between pH 6 and 8, the enzyme activity seemed to be independent of pH. Yang et al. (1985) determined the pH optimum for heparinase I to be 6.5. Lohse (Lohse and Linhardt, 1992) reported a pH optimum of



[Garrett and Grisham, Biochemistry, 1995.]

Figure 6.1: a) pH-activity profile for an enzyme with two residues critical for activity.
b) and c) together give rise to a.

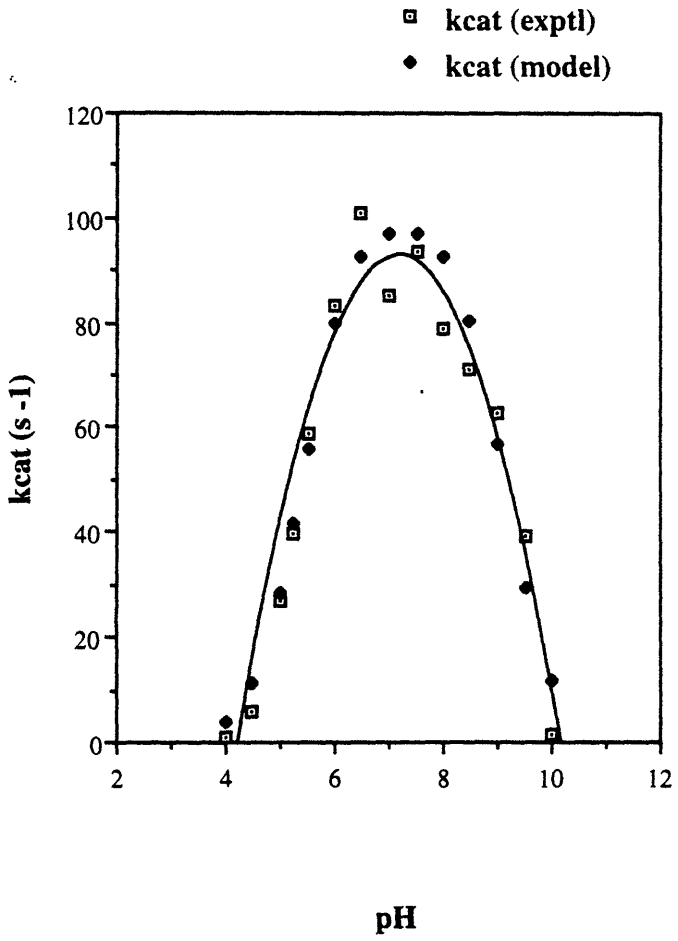


Figure 6.2: *k_{cat} versus pH profile for heparinase I.*
Heparinase I activity was determined (over a range of substrate concentrations to calculate kcat) as a function of pH to give rise to a bell shaped curve. The data were fitted to the model described in the text.

7.15. The difference in the values was attributed to enzyme stability since the assays were performed over different intervals of time (Lohse and Linhardt, 1992). Data from the experiments reported here seem to indicate a range of pH values (6.5-7.5) with optimal enzyme activity.

The pH plot shown in **Figure 6.2** initially rises with a slope of 1, reaches a plateau at around pH 6 and then decreases with a slope of -1, indicating that two groups are involved in catalysis. The values of Ka and Kb obtained from the above model were 5.3 and 8.5. This suggested that one group with an apparent pK value of about 5.3 was unprotonated and a group with an apparent pK value of 8.5 was protonated for catalysis.

The intrinsic pK values of a free cysteine and a free histidine are 8.3 and 6.0 respectively (Fersht, 1985). The pH data were consistent with the hypothesis that cysteine and histidine were involved in the activity of the enzyme. Considering the intrinsic pK values, data seemed to indicate that cysteine was protonated and histidine was unprotonated for catalysis. However, the pK values obtained are apparent values (i. e. not necessarily the intrinsic values found on the free enzyme) since k_{cat} describes the reaction of the enzyme-substrate complex (Denu et al., 1995). To further determine the individual pK's of cysteine-135 and histidine-203 residues in heparinase I, and establish the individual ionization states of these residues inactivation studies with chemical modifying agents were carried out at different pH values.

6.2.2 *DEPC inactivation of heparinase I as a function of pH*

It was described in chapter 4 that DEPC inactivates heparinase I at concentrations as low as 10 μM and that the inactivation was reversible using hydroxylamine, suggesting specificity towards histidine residues (Miles, 1977). Further, the modified residue was mapped to be histidine-203 and confirmed by site-directed mutagenesis.

To determine the pKa of histidine-203 residue, the effect of pH (over the range pH 5.0 to pH 9.0) on heparinase I inactivation with DEPC (at a concentration of 0.025 mM) was studied. As shown in **Figure 6.3**, the pseudo-first order rate constant of inactivation k, was plotted as a function of pH. The rate constant increased with

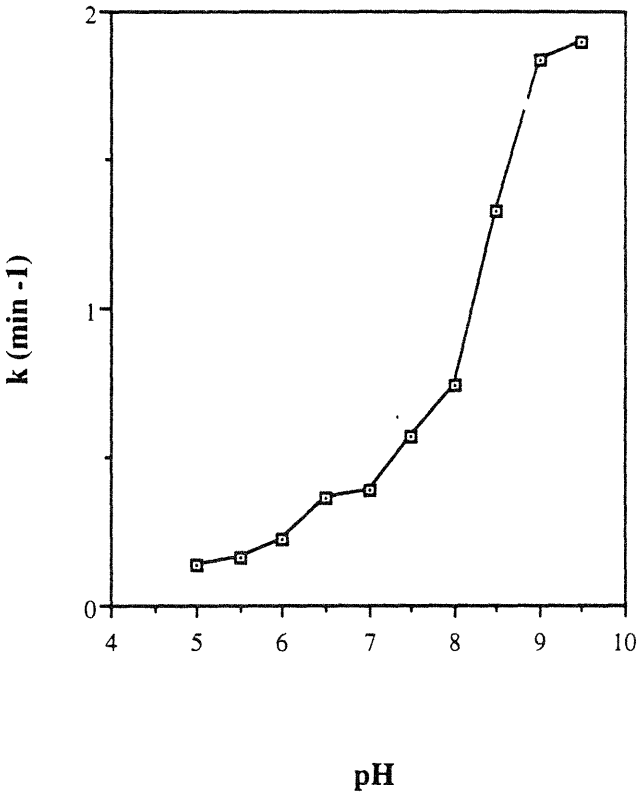


Figure 6.3: *Heparinase I inactivation with DEPC as a function of pH*
Heparinase I was incubated with DEPC at different pH values and the rate of inactivation was determined as a function of pH. The data were fitted to the model described in the text.

increasing pH and the increase was rapid after pH 8.0. The data were fitted to the following model using non-linear least square analysis (Denu et al., 1995):

$$k = C/(1+Kb/H)$$

where C is the pH independent value of k and Kb is the ionization constant of histidine-203. A pKb value of 8.2 was obtained from this equation. This was in good agreement with the value of 8.5 obtained from the pH rate profile for the native enzyme. This indicated that the histidine-203 residue, with a pKa of 8.2, must be protonated for catalysis as opposed to what one would have expected based on intrinsic pK values. Interestingly, this was about 2.2 units higher than its intrinsic value (Fersht, 1985). Hence, the other residue with the pKa of 5.3 observed in the pH rate profile, could be cysteine-135. To determine the pKa of the thiol group, the inactivation of heparinase I with IAA was examined as a function of pH.

6.2.3 IAA inactivation of heparinase I as a function of pH

It was shown previously (Sasisekharan et al., 1996) that cysteine-135 can be selectively carboxymethylated at pH 6.5 by IAA indicating that there were significant amounts of thiolate ions present. The inactivation of heparinase I with IAA was studied over a pH range of 4-7.5. As seen in **Figure 6.4**, the pseudo-first order rate constant of inactivation was pH dependent. The data were fitted to the following model using non-linear least square analysis (Denu et al., 1995):

$$k = C/(1+H/Ka)$$

A pKa value of 5.5 was obtained from the above equation which is unusually low when compared to its intrinsic pKa value of 8.3 (Fersht, 1985). This was in good agreement with the value of 5.3 obtained from the pH-rate profile for heparinase I.

Interestingly, such a behavior i. e. lowered pKa for cysteine has been observed before for the cysteine protease papain (Lewis et al., 1981) and more recently in the case of protein tyrosine phosphatases (Zhang et al., 1993; Denu et al., 1995; see discussion). The above results taken together suggested that cysteine-135 residue with a pKa of 5.5 is unprotonated and histidine-203 residue with a pKa of 8.2 is protonated for

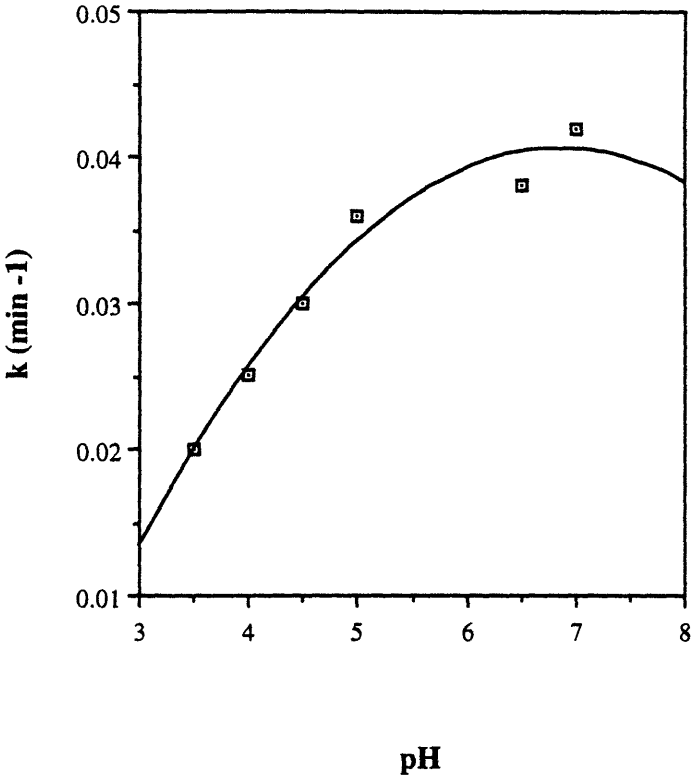


Figure 6.4: *Heparinase I inactivation with IAA as a function of pH*
Heparinase I was incubated with IAA at different pH values and the rate of inactivation was determined as a function of pH. The data were fitted to the model described in the text.

catalysis. As discussed in section 6.1, data pointed towards cysteine-135 being the base, abstracting a proton from the substrate and histidine-203 being the acid catalyst donating a proton to the leaving group. In order to further investigate the ionizations of these residues to gain insight into the protonation states of these residues, they were individually altered to a panel of residues with varying degrees of nucleophilicities i.e. pKas.

6.2.4 Cysteine-135 mutagenesis

Various amino acids were substituted for cysteine-135 with different intrinsic pKa values ranging from 3.5 (aspartic acid) to 9.3 (serine) (Fersht, 1985). Five mutant recombinant heparinases were constructed (C135A, C135D, C135E, C135H, and C135S) and expressed in BL21(DE3) host as described earlier (Sasisekharan et al., 1995). All cultures were grown under similar conditions and yielded similar levels of expression. **Figure 6.5** shows an SDS-PAGE gel of -L *r*-heparinase I and the mutant heparinases. The purity of the proteins was greater than 95% as detected by silver staining. The kinetic parameters were determined and are shown in **Table 6.1**. The k_{cat} values for all the cysteine mutations were reduced about 30-40 fold compared to wild type enzyme, thus confirming the importance of this residue in catalysis.

Table 6.1 Kinetic parameters for cysteine-135 mutations

Enzyme	k_{cat} (sec $^{-1}$)	K_m (μM)	Intrinsic pKa
-L *r*-heparinase I	92	10.2	Cys=8.3
C135A	ND[a]	ND	neutral
C135S	2.0	4.2	Ser=9.2
C135D	3.5	4.4	Asp=3.9
C135E	3.8	4.6	Glu=4.5
C135H	3.0	4.3	His=6.0

a ND= not determined since the enzyme was inactive

The extent of heparin degradation was analyzed by incubating heparin with equal amounts of the different mutant enzymes and the reaction was allowed to go to

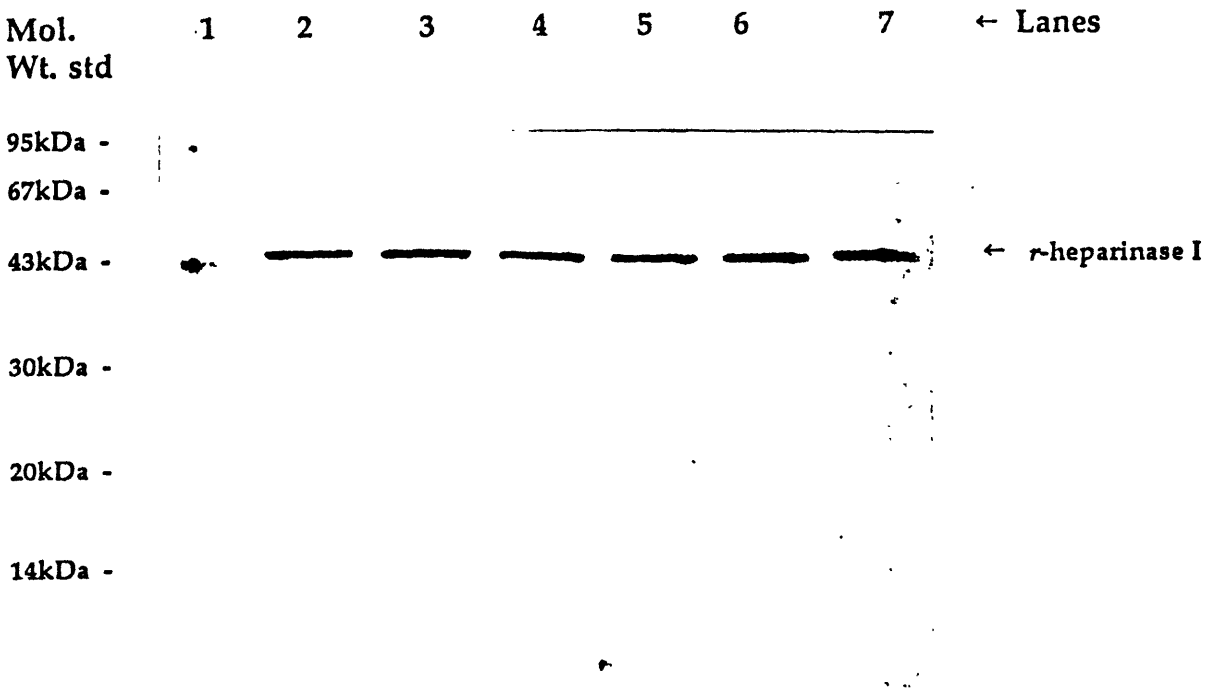


Figure 6.5: *SDS-PAGE of r-heparinase I and mutant r-heparinase I:* A 12% SDS-PAGE of heparinase I. Lanes are described from left to right. **Lane 1:** Molecular weight standards. **Lane 2:** Ni column purified C135A mutant *r*-heparinase I. **Lane 3:** Ni column purified C135D mutant *r*-heparinase I. **Lane 4:** Ni column purified C135S mutant *r*-heparinase I. **Lane 5:** Ni column purified C135E mutant *r*-heparinase I. **Lane 6:** Ni column purified C135H mutant *r*-heparinase I. **Lane 7:** Ni column purified *r*-heparinase I.

completion (~18 hrs) as described earlier (Sasisekharan et al., 1995). The products of degradation were separated on an anion exchange column using perfusion chromatography (POROS) and compared to the wild type enzyme degradation products **(Figure 6.6).** It was interesting to note that although all the mutant enzymes had essentially similar k_{cat} values **(Table 6.1)**, the product profiles after overnight degradation seemed to have a differential effect. The C135H, C135D and C135E mutants had similar effects on the amounts of products; however, the effect of C135S mutant was the most drastic.

Table 6.1 lists the pKas of the residues substituted for cysteine-135 and shows that serine had the highest pKa among the residues substituted. It was intriguing to observe a differential effect on enzyme activity based on the basicity of the residue substituted (see discussion). It is possible that a lower pKa (and hence a higher Ka) would facilitate this residue to function as a nucleophilic base in the degradation reaction. This suggested that the cysteine-135 residue could perhaps function as a general base in the elimination mechanism abstracting the C5 proton from the uronic acid, consistent with the previous inference based on pH data. However, it is also possible that the C135S mutation could perhaps affect the stability of the enzyme to a greater extent.

6.2.5 Histidine-203 mutagenesis

To further investigate the mechanistic role of histidine-203, this residue was altered to nucleophiles such as serine, aspartic acid and cysteine. All mutant enzymes were expressed and purified as described above. **Table 6.2** lists the kinetic parameters observed for the different mutant enzymes. The k_{cat} values were reduced 30-40 fold for all the mutant enzymes indicating the importance of this residue in catalysis.

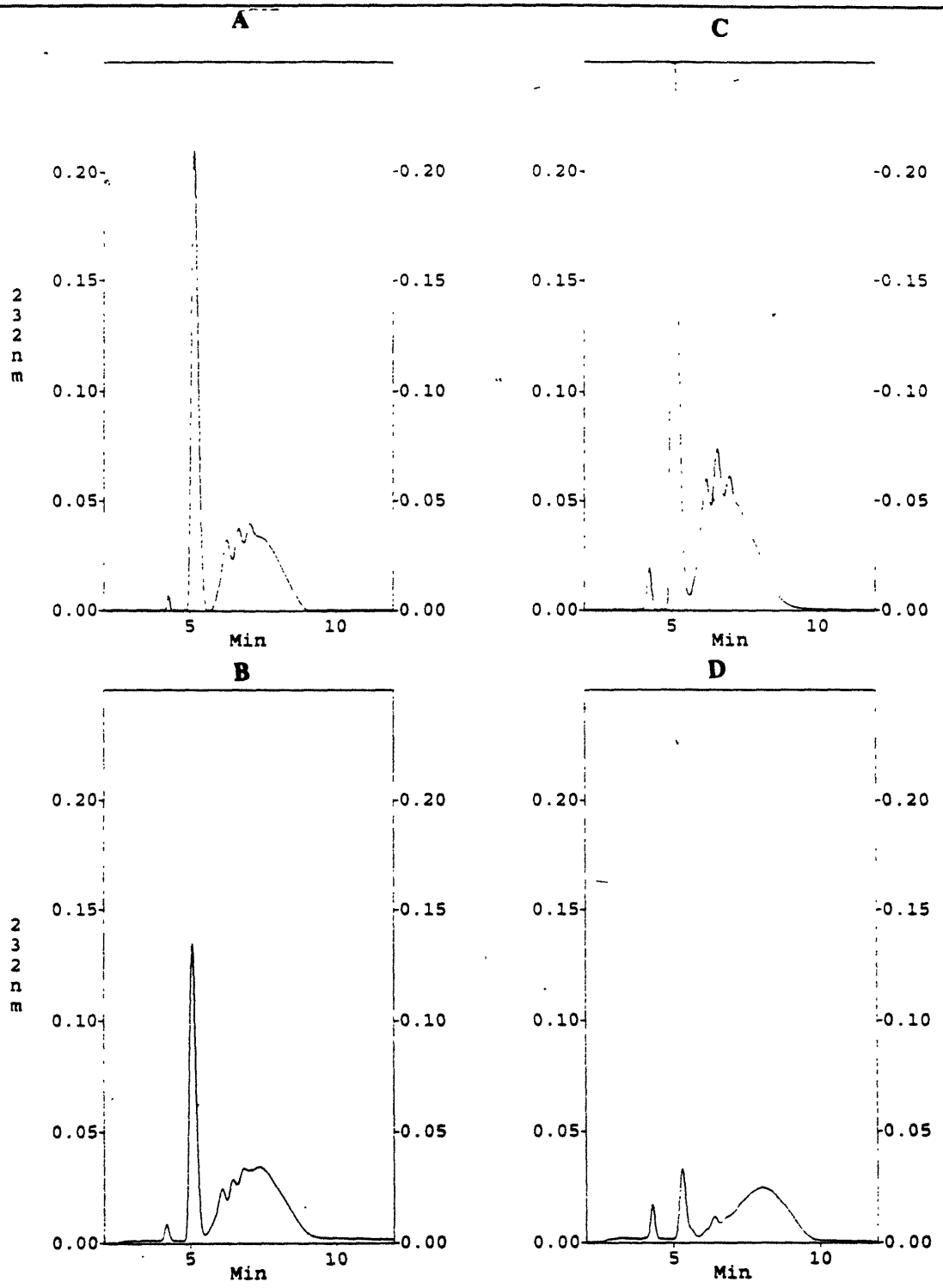


Figure 6.6: *Effect of cysteine-135 nucleophilicity on activity of heparinase I*
A shows the product profile of heparin degradation by the mutant C135D **B** shows the product profile of heparin degradation by the mutant C135E; **C** shows the product profile of heparin degradation by the mutant C135H; **D** shows the product profile of heparin degradation by the mutant C135S.

Table 6.2 Kinetic parameters for histidine-203 mutations

Enzyme	k_{cat} (sec $^{-1}$)	K_m (μM)
-L *r*-heparinase I	92	10.2
H203A	ND[a]	ND
H203D	3.5	5.7
H203S	3.2	6.2
H203C	3.9	5.9

a ND= not determined since the enzyme was inactive

When the reactions were allowed to go to completion and the product profiles were analyzed, the effect of all the mutations on the profiles was similar **(Figure 6.7)**. The imidazole side chain of histidines is versatile in its ability to both, accept and donate protons easily (Fersht, 1985) and it is possible that the side chains of the substituted residues can not replace the imidazole side chain of histidine in shuttling protons. This suggests that histidine-203 may play a role not just as a nucleophile but rather as an acid/base catalyst in a proton-shuttle mechanism, again consistent with the previous observation from pH data.

6.3 Discussion

Enzymes often accelerate the rate of chemical reactions over uncatalyzed reactions by utilizing acid/base catalysis (Fersht, 1985). The existence of acid/base catalysis is commonly established by identification of critical ionizations using pH analysis. k_{cat} versus pH profile for heparinase I reveals the presence of two critical ionizations, with one group (pKa of 5.3) to be unprotonated and another group (pKa of 8.5) to be protonated for catalysis. The data suggested that the ionizations reflected in the pH versus rate profiles could belong to active site residues cysteine-135 and histidine-203 of heparinase I. To determine the pKa of histidine-203, DEPC inactivation of heparinase I was studied as a function of pH and a pKa value of 8.2 was obtained. For cysteine-135, a pKa value of 5.5 was calculated based on IAA inactivation of heparinase I as a function of pH. Both these values were in close agreement to those obtained from the k_{cat} versus pH profile for heparinase I. The above results taken together suggested that cysteine-135 is unprotonated and

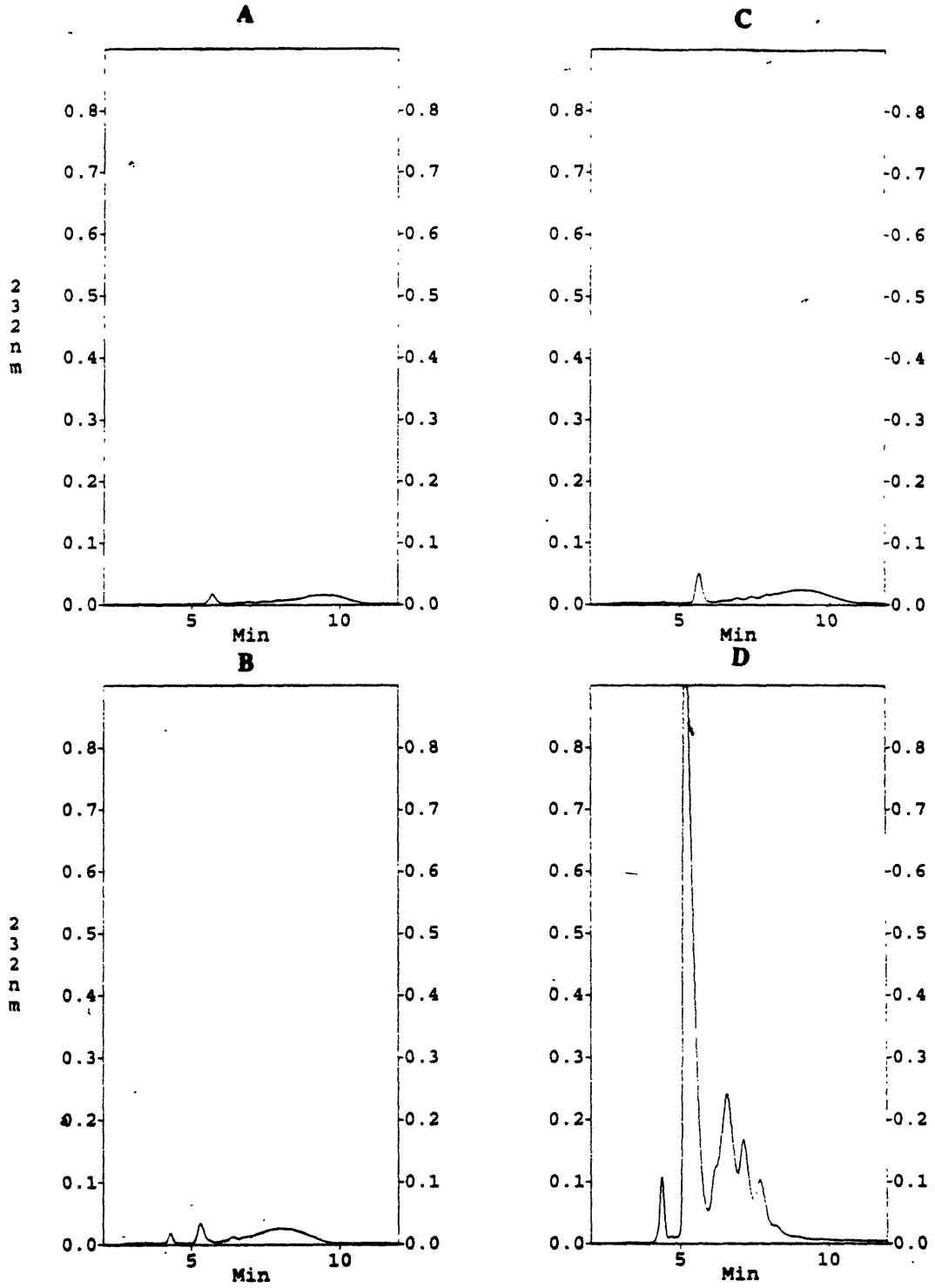


Figure 6.7: *Effect of histidine-203 nucleophilicity on activity of heparinase I*

A shows the product profile of heparin degradation by the mutant H203D; **B** shows the product profile of heparin degradation by the mutant H203C; **C** shows the product profile of heparin degradation by the mutant H203S; **D** shows the product profile of heparin degradation by *r*-heparinase I.

histidine-203 is protonated for catalysis.

In order to ascertain a mechanistic role and to study the nucleophilic potential of each of the two residues, the effect of nucleophilicity (i. e. basicity) on the enzyme reaction was studied by substitution of other nucleophiles. All mutations resulted in a 30-40 fold reduction in k_{cat}. However, when the reaction was allowed to proceed to completion, it was interesting to note that the effect of C135S mutation was far more drastic than that of other cysteine mutations. Although it is possible that different mutations could have different effects on the enzyme stability, all mutations were sterically conservative with C135S being the most conservative mutation (side chain containing OH instead of SH). The above result was consistent with serine (intrinsic pKa of 9.2) (Fersht, 1985) perhaps being least likely to ionize to its basic O^- form at pH 7 and hence least basic among the residues substituted. Thus the role of the cysteine residue seemed to be as a base to abstract a proton. Interestingly all the histidine mutants exhibited similar effects on enzyme activity, suggesting an acid/base catalyst role, role, for histidine-203. The data from all the above experiments taken together corroborated the previous hypothesis that these residues are in close proximity and are together involved in the catalytic mechanism.

6.3.1 The active site environment of heparinase I

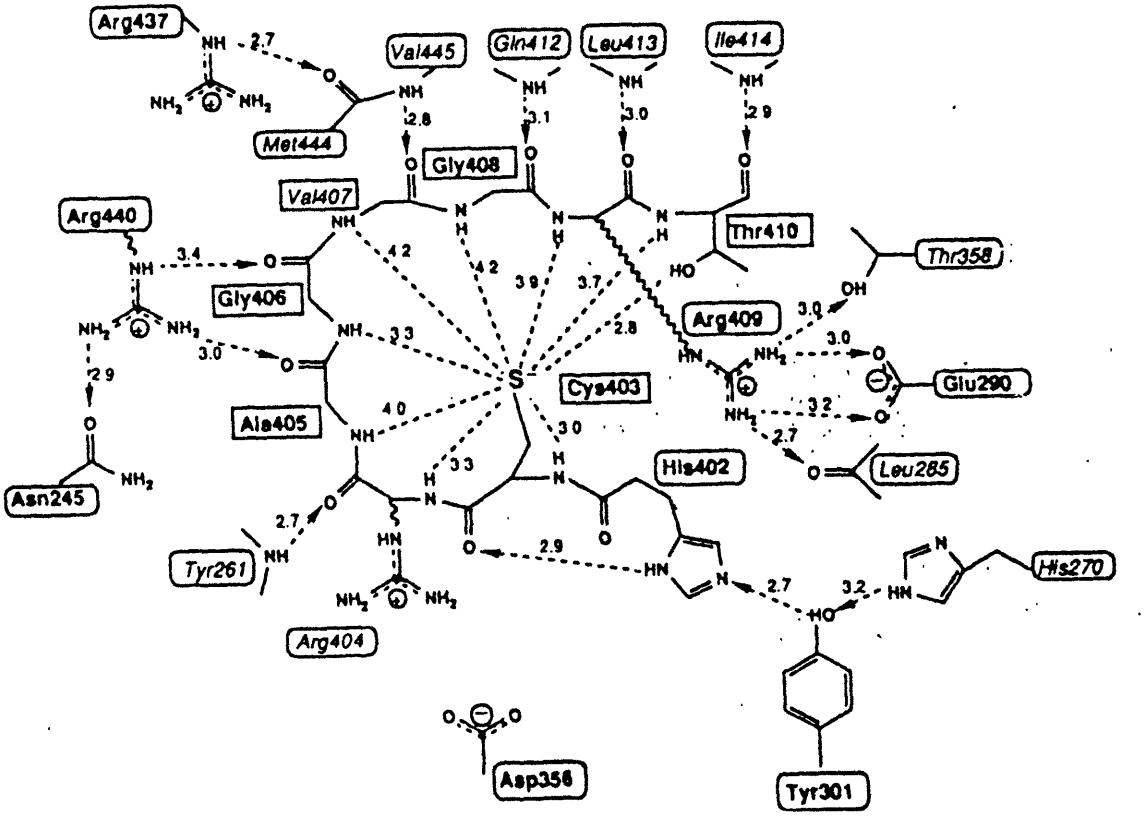
As described in chapter 2, cysteine-135 was preferentially reactive towards negatively charged compounds such as IAA and PCMB. Further, the observation of enhanced reactivity towards IAA compared to IAM and reduced PCMB based inactivation with increasing salt concentration, suggested the influence of a positively charged environment. Further, the previous chapter described that positively charged lysine-198 and lysine-199 of the heparin binding consensus sequence in the primary heparin binding site of heparinase I are important for enzyme activity. It was hence hypothesized that the role of positive charge could be to activate the thiol group by lowering its pKa so that it can act as a base for proton abstraction (Sasisekharan et al., 1995). Indeed, data presented here demonstrated that cysteine-135 had an unusually low pKa of 5.5, which is about 3 units lower than that for a normal cysteine (Fersht, 1985).

It is interesting to note that a similar observation of a highly reactive thiol with a low pKa has recently been reported for dual-specific and tyrosine-specific protein tyrosine phosphatases (PTPases) (Denu et al., 1995). The active site cysteines in the *Yersinia* PTPase and the dual-specific PTPase have been reported to have pKa values of 4.6 and 5.6 respectively (Zhang and Dixon, 1993; Denu et al., 1995). Further, IAA inactivation of PTPase was strongly ionic strength dependent (Zhang and Dixon, 1993) with a decreased inactivation with increasing ionic strength, similar to that observed for heparinase I (chapter 2).

The PTPase active site was thought to contain a highly reactive cysteine in a positively charged environment; this was supported by the preferential reactivity towards negatively charged IAA over its neutral analog, IAM. It is of significance to point out that the PTPases also cleave a highly negatively charged substrate containing a phosphate backbone, very similar to the heparin molecule which is highly sulfated and hence negatively charged. It is possible that a highly reactive thiol group in a positively charged environment is a common feature, with the role of the positively charged environment being not only in substrate binding but also in activating the thiol group for catalysis by lowering its pKa.

A reduction in pKa of this magnitude for a thiol would require stabilization by surrounding residues (Zhang and Dixon, 1993). The crystal structure of the active site of *Yersinia* PTPase indeed revealed that such is the case (Stuckey et al., 1994) **(Figure 6.8)**. The thiol is stabilized as an anion by hydrogen bonding with neighboring positively charged residues. Specifically, a histidine residue, immediately adjacent to the cysteine, plays a role in stabilizing the thiolate anion.

Similar differences in reactivity towards IAA have been reported for papain and other cysteine proteinases (Chaiken et al., 1969; Halasz et al., 1977). It has been proposed that the reactivity of haloacetates is enhanced by the presence of adjacent positively charged residue(s). In the case of papain, the active site cysteine was reported to have an unusually low pKa of about 3.3-4. Interestingly, the thiolate anion of the active site cysteine is stabilized by the formation of a zwitterionic pair with a neighboring imidazolium cation of a histidine (Lewis et al., 1976). The catalytic mechanism assumes an ion-pair model, where the thiolate anionic cysteine attacks the carbonyl carbon of the substrate and protonation of the thiol causes the



[Stuckey et al., Nature, 370, 571, 1994.]

Figure 6.8: The crystal structure of the active site of *Yersinia* PTPase. The thiol is stabilized as an anion by an intricate network of hydrogen bonding with neighboring residues.

pKa of the histidine to drop from 8.5 to 4.3 (Lewis et al., 1981). As a consequence, the acidity of histidine is enhanced and the protonated histidyl residue protonates the leaving group.

6.3.2 Heparinase I—A proposed catalytic mechanism for heparin cleavage

The imidazole side chain in histidine residues has been known to play a role in catalysis as an acid/base catalyst in the active site of several enzymes. For many years, the serine proteases such as chymotrypsin, trypsin and subtilisin have been shown to be members of a super gene family whose activity depends on a charge relay system involving serine, histidine, and aspartic acid (Dipersio et al., 1990). More recently members of the lipase family have been shown to be dependent on a charge relay system involving serine-histidine-aspartic acid catalytic triad for catalytic activity (Winkler et al., 1990; Brady et al., 1990). This has evolved to be a common theme in these hydrolases. Further, histidine has been thought to act as a general base to promote the reactivity of the active site cysteine in various enzymes such as anthranilate synthetase (Amuro et al., 1985) and carbomyl phosphate synthetase (Miran et al., 1991).

While a lot is known on the mechanism of hydrolases in general, there is very little information on the mechanism of polysaccharide degrading lyases. The crystal structure of pectate lyases has been solved (Yoder et al., 1993a; Yoder et al., 1993b) but there is no information on amino acids that are involved in the catalytic mechanism. The observation that pectin and heparin can be degraded by NaOH led to the proposal of a general base catalyzed mechanism for polysaccharide degrading lyases (Gacesa, 1987). Gacesa described a general mechanism for lyases involving a base catalyzed abstraction of the C5 proton and the formation of an unsaturated C4-C5 bond. A prior step to C5 proton abstraction is the neutralization of the negative charge on the carboxylate (C6) of the uronic acid.

While the charge neutralization is thought to be performed by residues such as lysine, a number of amino acids, such as aspartate, glutamate, histidine, lysine or cysteine, could act as a general base for C5 proton abstraction but no direct evidence for the involvement of any of these residues was available. Finally a different amino acid acting as a proton donor protonates the leaving group.

A mechanism for the elimination reaction, involving the degradation of heparin by heparinase I, which is consistent with the results presented in this chapter as well as with other previous data (Chapters 2,4,5), is proposed in **Figure 6.9**. Lysines-198 and 199 from the heparin binding site are in close proximity to the active site and are part of the positively charged environment around cysteine-135. This plays a key role not only in providing the necessary charge complimentarity for heparin binding but also in reducing the pKa of cysteine-135, and thus enhancing its reactivity.

Histidine-203 from the heparin binding site possibly exists as a histidyl cation, as was seen in the case of papain, and acts as a general base to increase the reactivity of cysteine-135 by keeping the thiol group negatively charged **(Step 1)**. The anionic thiol group in turn is stabilized by lysine residues in proximity. The thiolate anion functions as a base to abstract the C5 proton from the iduronate in the heparin molecule **(Step 2)**. This results in a 4-5 unsaturation, and in the cleavage of the C4-O glycosidic bond. The protonation of cysteine could increase the acidity of histidine-203 by reducing its pKa and as a result histidine-203 acts as a general acid catalyst by protonating the leaving hexosamine group **(Step 3)**. Histidine then abstracts the proton from cysteine and returns to the ion-pair mode and the enzyme is thus poised for another round of catalysis **(Step 3) (Figure 6.9)**.

6.4 Significance

A number of different uronic acid-containing polysaccharides undergo post-polymerization modification to produce the final biologically active structure. Enzymes that catalyze C5-epimerization of uronate residues (i. e. epimerases) at the polymer level have been an interesting target for studies, although very little is known about the mechanism of action (Skjak-Braek and Larsen, 1985; Larsen et al., 1986). The post-polymerization modification of polysaccharides has a significant effect on the secondary structure and the biological role of these molecules and hence understanding the mechanism of these epimerases would be very useful.

Polysaccharide lyases are a closely related family of enzymes and understanding the mechanism of these enzymes could provide valuable insight into the mechanism of

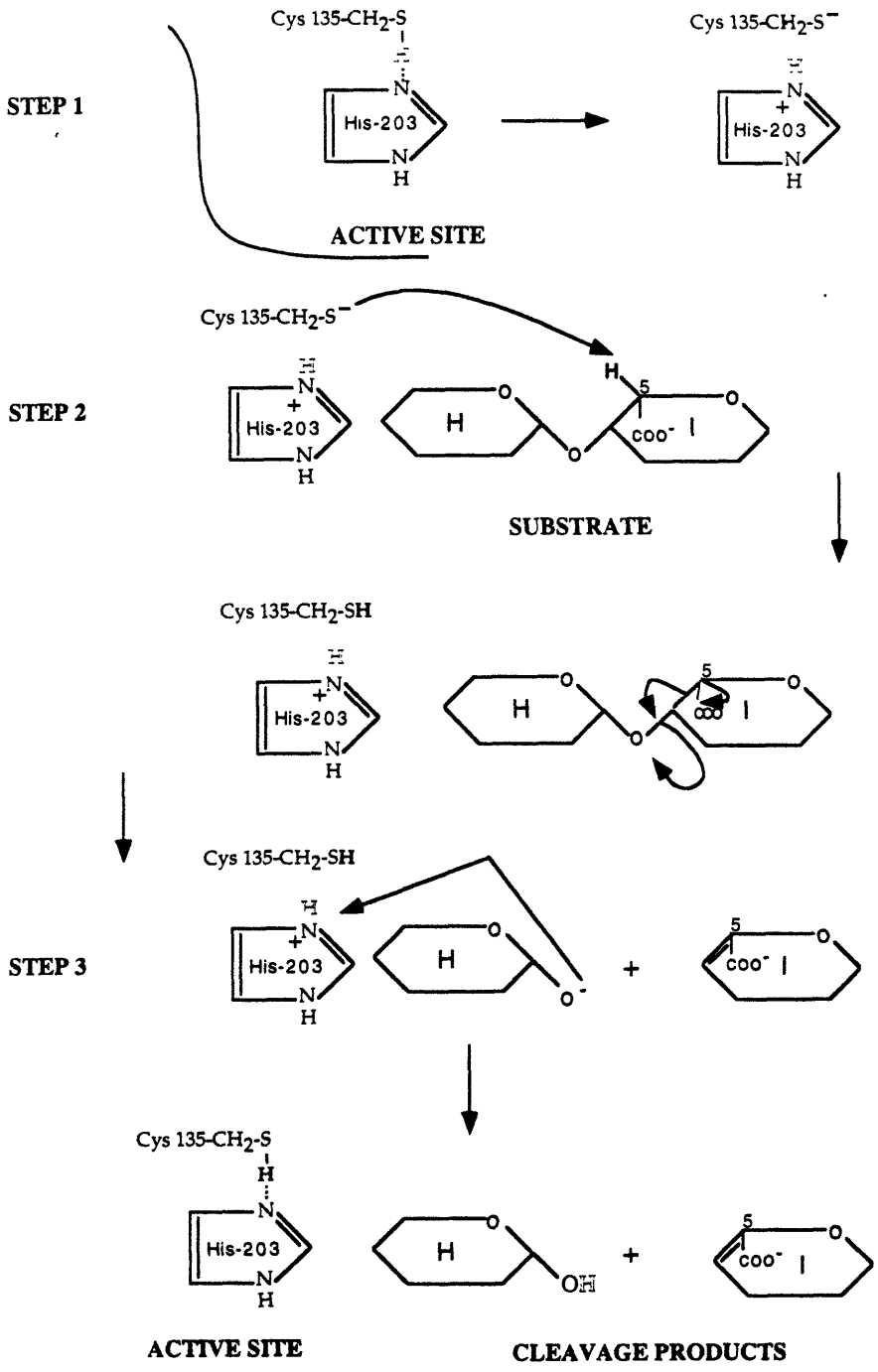


Figure 6.9: *A proposed mechanism for the catalytic cleavage of heparin by heparinase I*
Histidine-203 activates cysteine-135 for catalysis by abstracting a proton. Cysteine-135 acts as a base and abstracts the C5 proton from the uronate resulting in a 4-5 unsaturation. Histidine-203 protonates the leaving hexosamine group and the enzyme is ready for the next round of catalysis.

epimerases. It was proposed that both lyases and epimerases have a common mechanism of action with the difference being in proton donation by enzymes—to the leaving group (lyases) and to carbanion at C5 (epimerases) (Gacesa, 1987). By understanding the mechanism of lyases one could convert lyase action into an epimerase through protein engineering, thus leading to design of novel epimerases and hence novel polysaccharide products. The study presented in this chapter is significant in that it is the first detailed investigation of the catalytic mechanism of a polysaccharide degrading lyase. It will be interesting to see if this mechanism is a general mechanism for all members of the polysaccharide lyase family.

Chapter 7

Cloning & Recombinant Expression of Heparinases II and III in *E.coli*

The use of heparinases has been limited by difficulties in isolation and separation of sufficient quantities of these enzymes produced by *F.heparinum*. Other contaminating polysaccharide degrading and modifying enzymes, such as chondroitinases and sulfatases produced by *F.heparinum*, have complicated the effective use of these enzymes in characterizing heparin-like polysaccharides and in determining their substrate specificity. Hence it was of interest to clone and recombinantly express heparinases II and III. This chapter describes the cloning and recombinant expression in *E.coli* of heparinases II and III from *F. heparinum*. The DNA and protein sequences of both the genes are reported and compared with that of heparinase I for any homologous domains. Based on homology and on knowledge gained from heparinase I in this study, some predictions and speculations on the putative active site and substrate binding site of these enzymes are made.

7.1 Motivation

As described earlier, heparan sulfate proteoglycans (HSPGs) have been shown to play key roles in presentation of growth factors (e.g. fibroblast growth factors, vascular endothelial cell growth factor) to their receptors on various cells including endothelial cells, fibroblasts and smooth muscle cells, thereby influencing a variety of cellular processes including cell adhesion, migration, growth, proliferation, and differentiation (see chapter 2). As the HSPGs themselves play a significant role in modulating a variety of biological functions, enzymes that degrade heparan sulfate are also thought to be important regulatory molecules. Indeed, HSPG degradation, through heparanases, has been observed in inflammation, wound repair, and cancer metastasis (Nakajima et al., 1988).

Involvement of HSPGs in the presentation of growth factors to their receptor and their role in providing structural support to the ECM motivated investigations on the effect of heparinases in physiological processes such as angiogenesis and morphogenesis (Yost, 1992; Itoh and Sokol, 1994). It was shown recently that heparinases I and III are potent inhibitors of neovascularization *in vitro* and *in vivo* (Sasisekharan et al., 1994) while heparinase II has no effect. It was hypothesized that the differential effect could be due to different substrate specificities of these enzymes.

The unique specificity of heparinases has been exploited to map specific linkages in oligosaccharide sequences involved in the binding and modulation of several growth factors and enzymes (Gallagher et al., 1992; Habuchi et al., 1992; Linhardt et al., 1988). Heparinase II displays unusual enzymatic activity as it is capable of cleaving glycosidic linkages containing either a glucuronic or iduronic acid residue i. e. $H_{NY,6X}$-U_{2X}; as a result it cleaves both heparin as well as heparan sulfate. If this activity is associated with a single enzyme, this enzyme would be capable of catalyzing both anti (diaxial) and syn (axial-equatorial) elimination reactions (Linhardt et al., 1986; Linhardt et al., 1990). The only other lyase that displays such catalytic dexterity is chondroitinase ABC from *Bacteriodes.* sp. or *Protease vulgaris* (Linhardt et al., 1986).

Chapter 2 described the properties, purification, substrate specificities and applications of the three heparinases from *F. heparinum*. Our group has focused on the cloning, recombinant expression, and structural and functional characterization of heparinases I, II and III from *F. heparinum* in an effort to understand the catalytic mechanism and substrate specificities of these enzymes (chapters 2, 4, 5, and 6). Recent work in our laboratory also included the development of a scheme for purifying large scale amounts of recombinant heparinase I (Ernst et al., 1996).

The use of heparinases has been limited by difficulties in isolation and separation of sufficient quantities of these enzymes produced by *F.heparinum* (Lohse and Linhardt, 1992). Further, the analysis of substrate specificity of heparin lyases has been difficult due to contrasting observations made by different researchers (Desai et al., 1993). Other contaminating polysaccharide degrading and modifying enzymes, such as chondroitinases and sulfatases produced by *F.heparinum*, have complicated the effective use of these enzymes in characterizing heparin-like polysaccharides and in determining their substrate specificity. Hence it was of interest to clone and recombinantly express heparinases II and III.

This chapter describes the cloning and recombinant expression in *E. coli*, of the heparinase III (section 7.2) and heparinase II (section 7.3) genes from *F. heparinum*. The protein and DNA sequences were compared for homology and the 5' regions were compared to determine conserved promoter regions of *F. heparinum*.

7.2 Heparinase III: Cloning and Recombinant Expression in *E. coli*

7.2.1 Results

7.2.1.1 Heparinase purification and homogeneity

RPHPLC purification of native heparinase III showed the presence of a protein doublet **(Figure 7.1 a)**, similar to what is observed when heparinase I is purified from *F. heparinum* (Sasisekharan et al., 1993). It is possible that the heparinase III doublet peak is due to proteolytic degradation or protein oxidation due to fermentation. It has also been speculated that protein isoforms seen for the heparinases from *F.heparinum* may be due to post-translational modification(s) (see

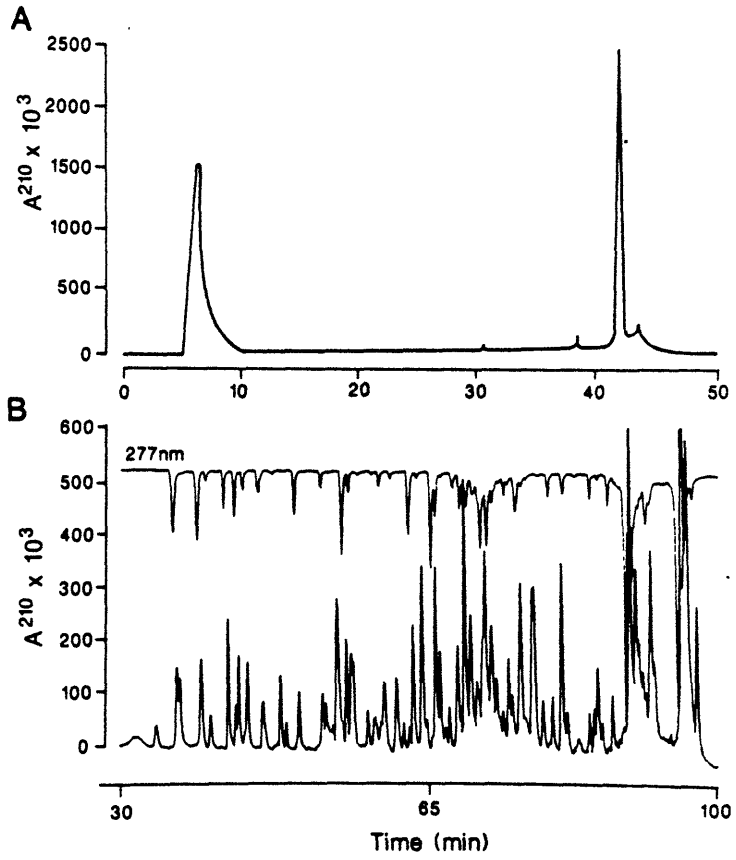


Figure 7.1: *Reverse-phase HPLC profile of purified heparinase III and tryptic digest of heparinase III.*
a) Heparinase III appears as a doublet and the separation of the two isoforms is shown. The major isoform was digested with trypsin.
b) Heparinase III digested with trypsin and resolved on a reverse-phase HPLC column.

discussion and (Zimmermann, 1984)). The doublet was further separated, and the overall tryptic map of the two peaks were virtually identical. The average molecular mass estimated by laser desorption mass spectrometry was 73.539 kDa (Sasisekharan et al., 1994). The N- terminal amino acid is blocked in heparinase III, similar to that observed for heparinase I (Sasisekharan et al., 1993).

7.2.1.2 Cloning strategy, PCR amplification, Southern Blotting

The HPLC-purified heparinase III was reduced, alkylated and digested with trypsin; RPHPLC peptide separation yielded about 64 peaks **(Figure 7.1 b)**. Ten peaks were sequenced and their sequences are shown in **Table 7.1**. **Table 7.2** lists the primer design for the peptide sequences chosen for cloning and PCR amplification.

Table 7.1 Amino acid sequences of tryptic peptides of heparinase III

Peptide	Amino acid sequence
td 46	(K,R) N L F A G V S F P E F K D S P R
td 26	(K,R) P A F V F E K P R
td 17	(K,R) E P D F S N A E K P A D I R
td 25	(K,R) A L L A Y Y R
td 17'	(K,R) A P E I S I R
td 17''	(K,R) L R N W Y R
td 17'''	(K,R) Q G K A P N F L
td 5	(K,R) S G W D K
td 25'	(K,R) S V L F I N K
td 26'	(K,R) I H S T L T L D N Q N M V I T K

This table lists all the heparinase III tryptic peptides sequenced. The sequence begins with (K,R) because trypsin cuts at either lysine or arginine residues.

Table 7.2 Primer design for heparinase III

td peptide 46 A	5'- AAX XTI TTX GCI GGI GTI ZBI TTX CC - 3'
td peptide 46 B	5'- GGY AAI ZBI ACI CCI GCY AAI AYY TT - 3'
td peptide 26 A	5'- CCI GCI TTX GTI TTX GAY AAY CC - 3'
td peptide 26 B	5'- GGX TTX TCY AAI ACY AA- 3'

The table represents the two sets of primers used in the PCR for cloning of the heparinase III gene. I stands for the nucloetide inosine. Z=A/T, B=G/C, X=T/C & Y=A/G

Primers td 46A and td 26B, produced a 1100 bp product. Based on a molecular mass of about 73 kDa, the product represented about 55% of the approximately 2000 bp heparinase III gene. This PCR product was used to probe the heparinase III gene in the *F. heparinum* genomic DNA. Southern blots of *F. heparinum* genomic DNA using the 1100 bp PCR product as the probe, indicated the heparinase III gene was contained in ~ 8 kb *Eco*RI fragment of genomic DNA and in a ~ 3.9 kb *Pst I* fragment.

7.2.1.3 Screening, restriction mapping and sequencing of heparinase III clone

The PCR product obtained above was used to probe a *F.heparinum* λ ZAP II library. Approximately ~1x 10^6 primary clones were screened. Two positive plaques were isolated following tertiary screening and one of the clones was used for *in vivo* excision and sequencing. The plasmids containing the positive clones (pRG5-2 and pRG5-3), identified from the above screening were characterized by restriction mapping. pRG5-3 hep III, contained a 6.1 kb insert, with a 3.9 kb *PstI-PstI* fragment.

The *PstI-PstI* fragment was sequenced beginning with the internal primers td 46A and td 26B. Sequence information thus obtained was used to design new sequencing primers. The sequence **(Figure 7.2)** reveals a single, continuous open reading frame of 1980 base pairs (660 amino acids). The PCR product spans bases 738 to 1824 (from the ATG start site), and corresponds to about 55% of the total gene. In all ten different tryptic peptides mapped onto the ORF constituting about 18% of the total of 660 amino acids.

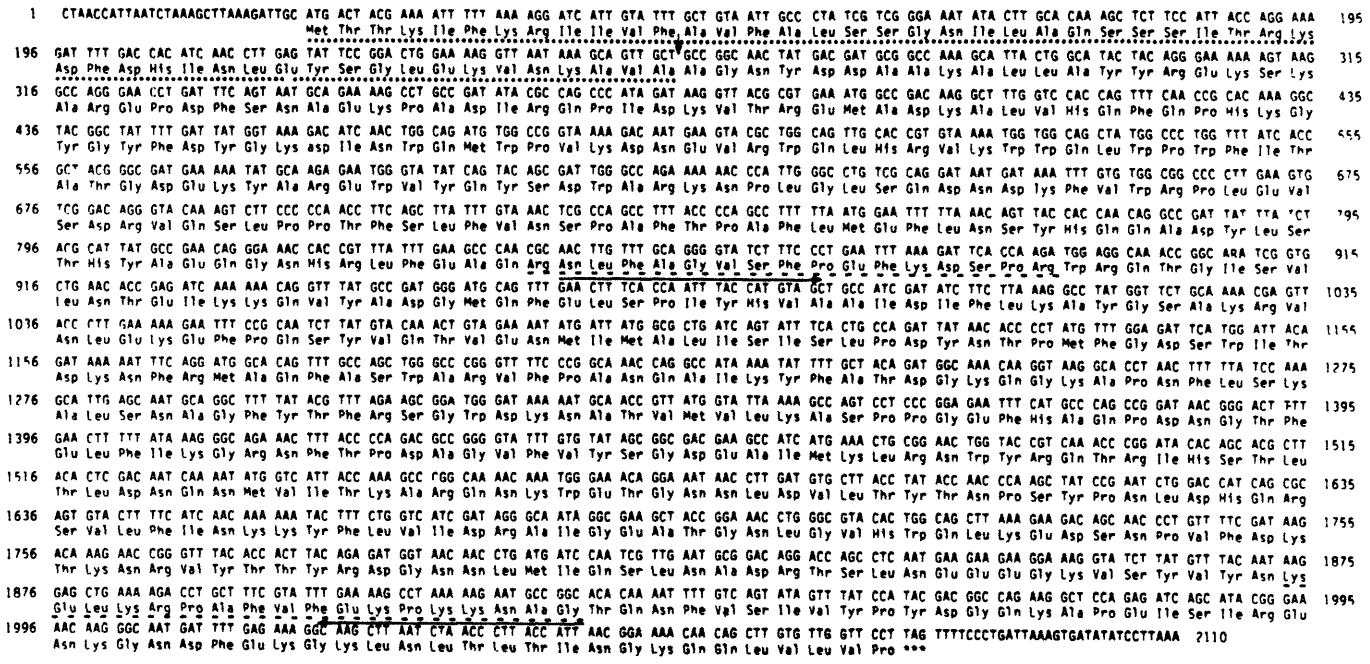


Figure 7.2: *ORF encoding the gene for heparinase III from F. heparinum*

Heparinase III does not have cysteines or any heparin binding consensus sequences in contrast to heparinase I, which has two cysteine residues (one of which is in the active site) and a heparin binding site containing two Cardin-Weintraub heparin binding consensus sequences and one EF-hand calcium coordination motif (Sasisekharan et al., 1996). However, there are two potential EF-hand calcium co-ordinating motifs (spanning residues 390-405 and residues 576-591) in heparinase III. Residues 50-52 represent a consensus Ala xxx Ala site for cleavage, typical of prokaryotic signal sequences (see discussion) (von Heijne, 1988).

7.2.1.4 Recombinant heparinase III

Several approaches to expressing heparinase III from plasmids pET-15b and pET-28a(+) were evaluated. In all cases, heparinase III was expressed as a fusion protein with a His_6-tag to permit affinity purification on a Ni^{2+}-charged resin. The three principal strategies appended the His_6-tag to the N-terminus of heparinase III in plasmids pET-15b or pET-28a(+), or to the C-terminus in plasmid pET-28a(+).

Optimal heparinase III expression was achieved when a His_6-tag was appended to the N-terminus of heparinase III in plasmid pET-15b. Cell lysis yielded a crude extract in which 80 - 85% of r-heparinase III was insoluble; the remaining heparinase III partitioned into the soluble crude extract. **Figure 7.3** shows a Coomasie stained SDS-PAGE gel of recombinant heparinase III. The insoluble fraction was refolded and purified as described earlier for heparinase I (Ernst et al., 1996). However the protein was not active under these conditions. Active heparinase III could only be recovered from the soluble crude extract fraction. The material was incubated with heparan sulfate overnight and the degradation products were separated by anion-exchange chromatography as shown in **Figure 7.4**. Prior to anion exchange separation, the digest was purified using reverse phase HPLC to separate other proteins from the crude extract. The products of the digest did not bind to the column and were collected and subsequently passed through the anion exchange column.

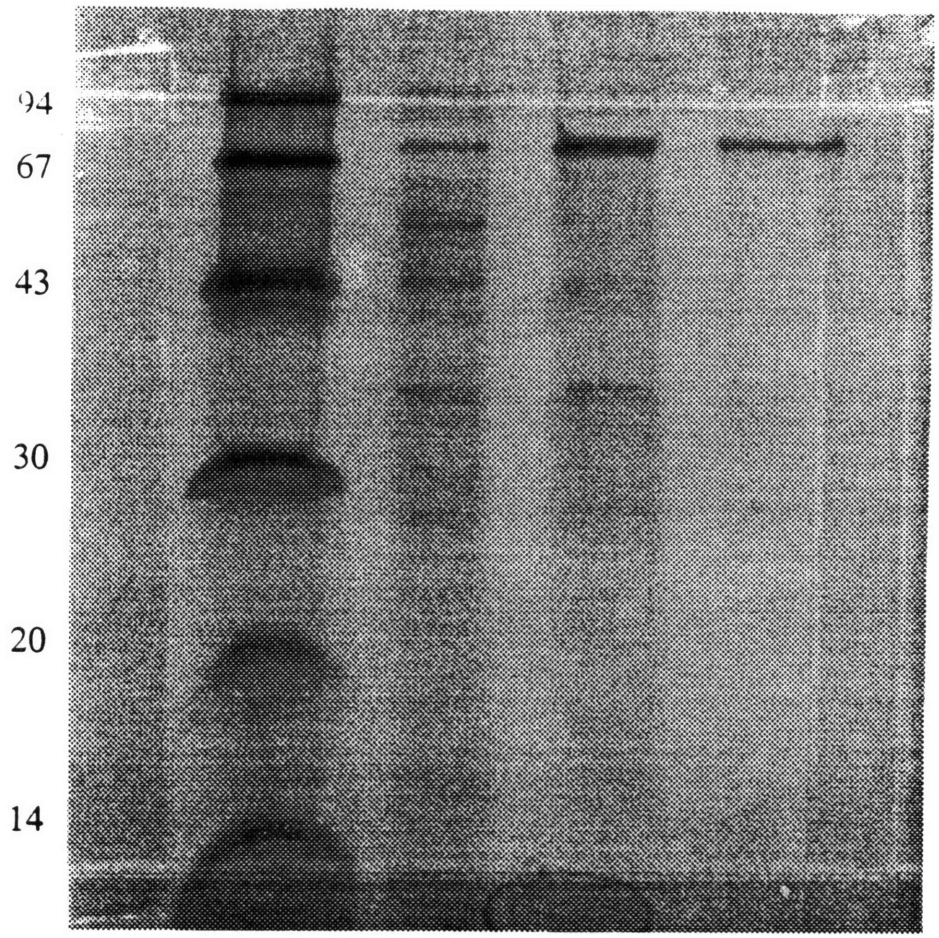


Figure 7.3: *SDS-PAGE gel of recombinant heparinase III.*
A 12% SDS-PAGE of heparinase III. Lanes are described from left to right. **Lane 1:** Molecular weight standards. **Lane 2:** soluble crude extract. **Lane 3:** insoluble crude extract. **Lane 4:** Ni column purified insoluble fraction.

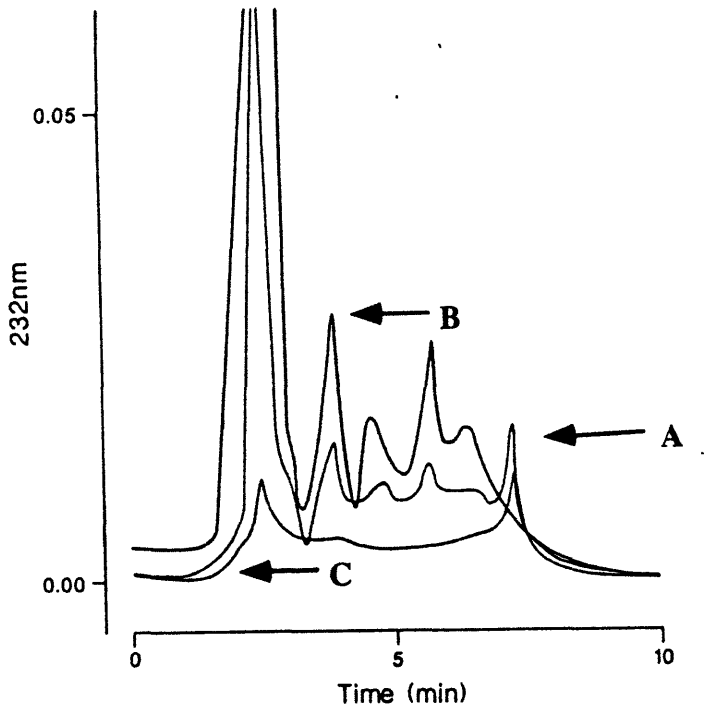


Figure 7.4: *Native and recombinant heparinase III activity*
A shows the product profile of heparan sulfate degradation by *r*-heparinase III; **B** shows the product profile of heparan sulfate degradation by native heparinase III; **C** shows the product profile of heparan sulfate degradation by *E.coli* soluble crude extract.

7.2.2 Discussion

The heparinase III gene was sequenced in both directions to deduce the 1980 bp ORF. It has a putative signal sequence composed of 52 amino acids based on the prokaryotic Ala xxx Ala cleavage site (von Heijne, 1988). Since heparinase III is N-terminally blocked it was not possible to determine the actual start site for the mature protein. Based on the putative signal sequence the deduced molecular weight of mature heparinase III is 70, 300 Da. However, the molecular weight of the wild-type *F.heparinum* heparinase III protein, as determined by mass spectrometry is 73,540 Da. Further, including the leader sequence (ORF encoded protein) does not account for the discrepancy in molecular weight. Thus, the actual start site for heparinase III and the type of N-terminal modification remain unclear at this time. However, the discrepancy in the molecular weights is possibly due to additional protein post-translational modifications (see below).

Recombinant heparinase III was expressed using the pET expression system. A significant fraction of the recombinant protein was insoluble; a smaller fraction was soluble. Crude soluble extract contained heparinase III activity with heparan sulfate. The soluble fraction did not bind to the Ni affinity column, suggesting that when heparinase III is in an active conformation, an N-terminal His_6-tag is not accessible to the affinity resin. Purification and refolding of the insoluble material resulted in very little activity, suggesting improper folding or instability due to absence of glycosylation (see below).

Interestingly, a new type of O-linked glycosylation has been discovered for heparinase I from *F.heparinum* (Huang, 1995). This is consistent with an additional recent finding of the O-linked type glycosylation seen on several proteins expressed by a related *Flavobacterium menigosepticum* (Plummer, 1995). This glycosylation (involving an acidic oliogosaccharide) is novel in that it is unique among both prokaryotes and eukaryotes. In case of heparinase I, it was speculated that glycosylation is essential for enzyme thermal stability but not activity (Ernst et al., 1996). It is of importance to point out that the multiple isoforms of heparinases I, II and III observed during purification (Lohse and Linhardt, 1992; Sasisekharan et al., 1993) of the enzymes to homogeneity, is typically seen of several glycoproteins

(Varki, 1993). It is possible that the poor activity observed for recombinant heparinase III is due to the absence of glycosylation when expressed in *E.coli*.

The previous chapters described the characterization of heparinase I with an active site cysteine residue in a basic environment and a heparin binding site, with two Cardin-Weintraub heparin binding consensus sequences and a EF-hand calcium coordinating motif, in close proximity. Interestingly, in the case of heparinase III, the primary amino acid sequence does not contain any cysteine residues or identifiable heparin binding consensus sequences. Heparinase III is selective in cleaving heparan sulfate-like regions, which have a relatively lower degree of sulfation and hence lower negative charge. As a consequence, the active site environment of heparinase III is likely to be different from that observed for heparinase I.

It is known that calcium catalytically activates heparinases I and III (Lohse and Linhardt, 1992). In the case of heparinase III, the primary amino acid sequence contains two putative calcium co-ordinating motifs (spanning residues 390-405 and residues 576-591) which satisfy the 'EF-hand' structural domain, similar to what is observed for heparinase I (Sasisekharan et al., 1996). However, apart from the presence of calcium coordinating motifs, it is of significance to note that there is little sequence homology (about 15%) at the DNA and the amino acid levels between heparinases I and III. It is pertinent to point out that in the case of pectate lyases, while there is less than 30 % primary sequence homology between the different enzymes, the proteins appear to preserve a unique parallel β-sheet scaffold (Pickersgill et al., 1994; Yoder et al., 1993a; Yoder et al., 1993b). It remains to be seen if the heparinases, like the pectate lyases, share a structurally similar scaffold.

In summary, heparan sulfate degrading enzymes appear to play a key role in regulating several important molecules signaling through the ECM. Recently, a platelet-derived peptide was shown to possess heparin degrading activity (Hoogewerf, 1995). This peptide belongs to the platelet basic protein (PBP) family of chemokines that function in inflammation, wound healing and growth regulation. Further, heparan sulfate degrading enzymes by themselves have significant therapeutic value since heparinase III is a very potent inhibitor of neovascularization *in vivo* and *in vitro* (Sasisekharan et al., 1994). The work

presented here is significant in being the first description of the cloning and recombinant expression in *E. coli* of a heparan sulfate degrading enzyme.

7.3 Heparinase II: Cloning and Recombinant Expression in *E. coli.*

7.3.1 Results

7.3.1.1 Cloning of heparinase II

The strategy adopted for cloning the heparinase II gene was similar to that described for heparinase III. **Figure 7.5 a** shows the RPHPLC purification of native heparinase II and the presence of a protein doublet, similar to heparinase I and III profiles. The average molecular mass estimated by laser desorption mass spectrometry was 85.984 kDa (Sasisekharan et al., 1994). The N- terminal amino acid is blocked in heparinase II, similar to that observed for heparinases I and III (Sasisekharan et al., 1993). Tryptic digest of the purified peak yielded approximately 78 peaks **(Figure 7.5 b).** Eleven peaks were sequenced and **Table 7.3** lists the sequences of the peptides. Four of the eleven peptides were chosen for synthesizing oligonucleotides for PCR amplification; the sequences are shown in **Table 7.4.**

Primers td 51A and td 48B produced a ~ 800 bp product and primers td 27C and 48B yielded a ~ 600 bp product. As a control, the 800 bp product was isolated and used as a template for PCR with primers 27C and 48 B to confirm the appearance of a 200 bp product. Based on a molecular mass of about 85 kDa, the ~ 800 bp product represented about 33 % of the approximately 2400 bp heparinase II gene.

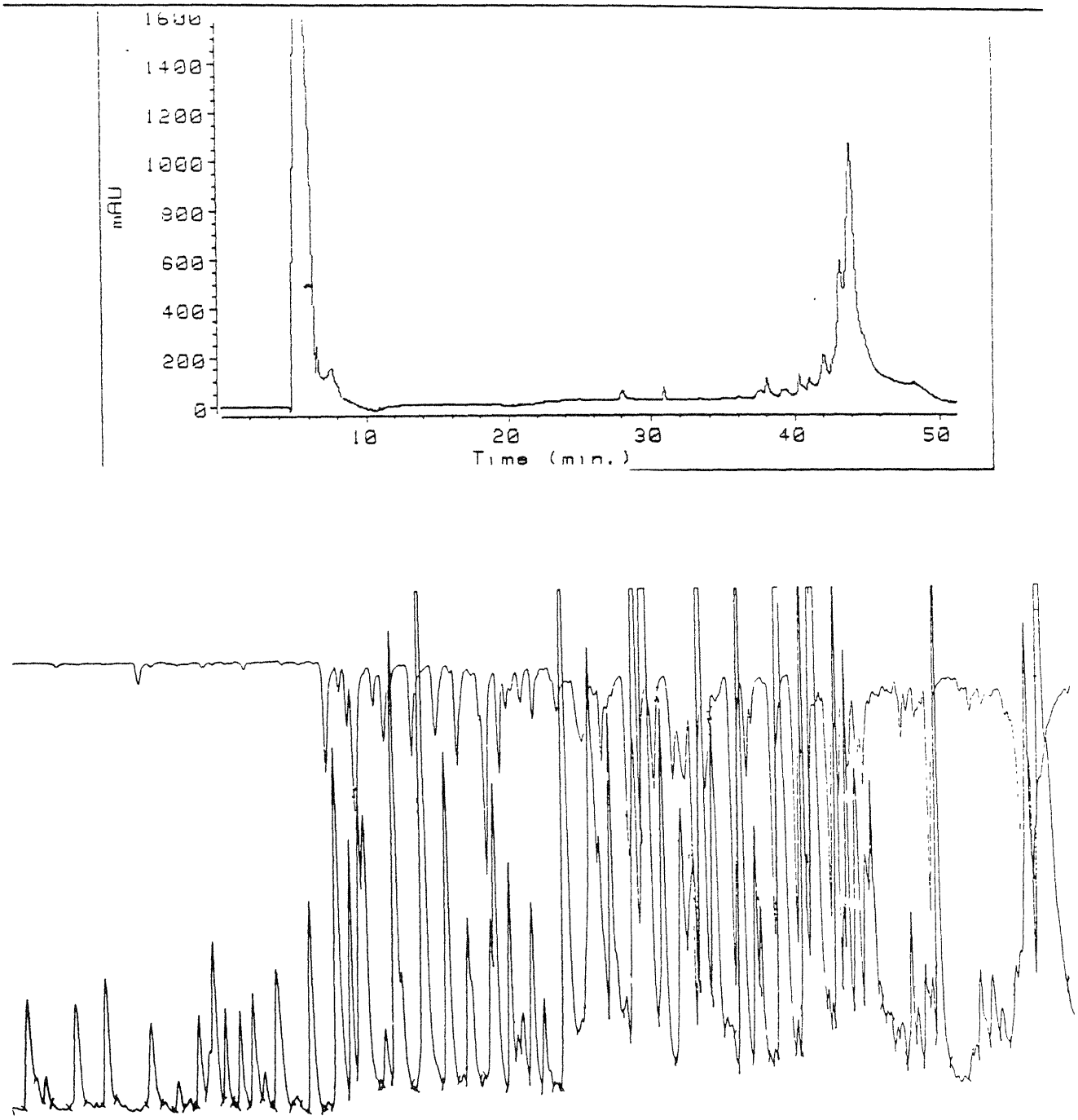


Figure 7.5: *Reverse-phase HPLC profile of purified heparinase II and tryptic digest of heparinase II*
a) Heparinase II appears as a doublet and the separation of the two isoforms is shown. The major isoform was digested with trypsin.
b) Heparinase II digested with trypsin and resolved on a reverse-phase HPLC column.

Table 7.3 Amino acid sequences of tryptic peptides of heparinase II

Peptide	Amino acid sequence
td 21	(K,R) L Y L R
td 21'	(K,R) K L P D D L P L S R
td 27	(K,R) P L A I D A G S Y Y G S S G G Y N S P - N K
td 27'	(K,R) R I D G D
td 27''	(K,R) F F K E H
td 27'''	(K,R) D G - K V V - - Q L A I
td 8	(K,R) D P N V E P H Q K
td 48	(K,R) G K D F - V F G T N Y T N D P K P G
td 48'	(K,R) M V T G A I L Y
td 72	(K,R) N W F Y P S H N Y H Q G M S Y L N V R F T N D L F A L - I L
td 51	(K,R) Y S G S P F G - M I A R - G

This table lists all the heparinase II tryptic peptides sequenced. The sequence begins with (K,R) because trypsin cuts at either lysine or arginine residues.

Table 7.4 Primer design for heparinase II

td peptide 27 C	5'- GGI TAX AAX ZBI CCI III AAX AA - 3'
td peptide 27D	5'- TTY TTI III GGI BZY TTY TA - 3'
td peptide 48 A	5'- GTI TTX GGI ACI AAX TAX ACI AAX GAX CC - 3'
td peptide 48 B	5'- GGY TCY TTI GTY TAY TTI GTI CCY IAC- 3'
td peptide 51 A	5'- GGI ZBI CCI TTX GGI III ATG AT - 3'
td peptide 51 B	5'- ATC ATI III CCY AAI GGI BZI CC- 3'

The table represents the two sets of primers used in the PCR for cloning of the heparinase III gene. I stands for the nucloetide inosine. Z=A/T, B=G/C, X=T/C & Y=A/G

7.3.1.2 *Screening and sequencing of the heparinase II clone*

The *F.heparinum* λ ZAP II library was screened using the 792 bp PCR product obtained above. Approximately 1 X 10^6 primary clones were screened. Two positive clones were isolated following tertiary screening and one of the clones was used for *in vivo* excision and sequencing. The sequence **(Figure 7.6)** revealed a single, continuous open reading frame of 2316 base pairs (772 amino acids). Interestingly a consensus Ala xxx Ala site for cleavage (typical of prokaryotic signal sequences, (On Heijne, 1988)) which was observed for heparinases I and III, was absent in the heparinase II sequence. The 792 bp PCR product spanned bases 1123 to 1915 (from the ATG start site), and corresponds to about 32 % of the total gene. In all eleven different tryptic peptides mapped onto the ORF, constituting about 23 % of the total of 772 amino acids.

The sequence reveals three cysteine residues in contrast to heparinase I, which has two cysteine residues (one of which is in the active site) and heparinase III which does not have cysteines. The protein is highly basic with a pI of about 8.9 with approximately 15 % basic residues in the primary sequence. Interestingly, there are three basic clusters (amino acid residues 312-316; 446-451; and 594-597) that satisfy the heparin binding consensus sequence (Cardin and Weintraub, 1989). Two such consensus sequences were observed in the heparin binding site for heparinase I (Sasisekharan et al., 1996).

7.3.1.2 *Recombinant Heparinase II*

Heparinase II was expressed in *E. coli* using the pet-28a expression system as described above for heparinase III. As observed for heparinase III, 80 - 85% of r-heparinase II was insoluble; the remaining heparinase II partitioned into the soluble crude extract. The soluble crude extract was incubated with heparin overnight and the products of degradation were separated using anion exchange chromatography as shown in **Figure 7.7**. The major products of degradation are two disaccharides with minor amounts of other products. The product profile was more or less similar to that obtained from purified native heparinase II degradation of heparin (Linhardt et al., 1990). The elution times of the two disaccharide peaks were identical to the elution times of the two disaccharide peaks (i. e. di-sulfated and tri-

```
1/1                                   31/11                                 61/21
ATG AAA CGA CAA TTA TAC CTG TAT GTG ATT TTT GTT GTA GTT GAA CTT ATG GTT TTT ACA ACA AAG GGC TAT TCC CAA ACC AAG GCC GAT
Met lys arg gln leu tyr leu tyr val ile phe val val val glu leu met val phe thr thr lys gly tyr ser gln thr lys ala asp
91/31                                 121/41                                151/51
GTG GTT TGG AAA GAC GTG GAT GGC GTA TCT ATG CCC ATA CCC CCT AAG ACC CAC CCG CGT TTG TAT CTA CGT GAG CAG CAA GTT CCT GAC
val val trp lys asp val asp gly val ser met pro ile pro pro lys thr his pro arg leu tyr leu arg glu gln gln val pro asp
181/61                                211/71                                241/81
CTG AAA AAC AGG ATG AAC GAC CCT AAA CTG AAA AAA GTT TGG GCC GAT ATG ATC AAG ATG CAG GAA GAC TGG AAG CCA GCT GAT ATT CCT
leu lys asn arg met asn asp pro lys leu lys lys val trp ala asp met ile lys met gln glu asp trp lys pro ala asp ile pro
271/91                                301/101                               331/111
GAA GTT AAA GAC TTT CGT TTT TAT TTT AAC CAG AAA GGG CTT ACT GTA AGG GTT GAA CTA ATG GCC CTG AAC TAT CTG ATG ACC AAG GAT
glu val lys asp phe arg phe tyr phe asn gln lys gly leu thr val arg val glu leu met ala leu asn tyr leu met thr lys asp
361/121                               391/131                               421/141
CCA AAG GTA GGA CGG GAA GCC ATC ACT TCA ATT ATT GAT ACC CTT GAA ACT GCA ACT TTT AAA CCA GCA GGT GAT ATT TCG AGA GGG ATA
pro lys val gly arg glu ala ile thr ser ile ile asp thr leu glu thr ala thr phe lys pro ala gly asp ile ser arg gly ile
451/151                               481/161                               511/171
GGC CTG TTT ATG GTT ACA GGG GCC ATT GTG TAT GAC TGG TGC TAC GAT CAG CTG AAA CCA GAA GAG AAA ACA CGT TTT GTG AAG GCA TTT
gly leu phe met val thr gly ala ile val tyr asp trp cys tyr asp gln leu lys pro glu glu lys thr arg phe val lys ala phe
541/181                               571/191                               601/201
GTG AGG CTG GCC AAA ATG CTC GAA TGT GGT TAT CCT CCG GTA AAA GAC AAG TCT ATT GTT GGG CAT GCT TCC GAA TGG ATG ATC ATG CGG
val arg leu ala lys met leu glu cys gly tyr pro pro val lys asp lys ser ile val gly his ala ser glu trp met ile met arg
631/211                               661/221                               691/231
GAC CTG CTT TCT GTA GGG ATT GCC ATT TAC GAT GAA TTC CCT GAG ATG TAT AAC CTG GCT GCG GGT CGT TTT TTC AAA GAA CAC CTG GTT
asp leu leu ser val gly ile ala ile tyr asp glu phe pro glu met tyr asn leu ala ala gly arg phe phe lys glu his leu val
721/241                               751/251                               781/261
GCC CGC AAC TGG TTT TAT CCC TCG CAT AAC TAC CAT CAG GGT ATG TCA TAC CTG AAC GTA AGA TTT ACC AAC GAC CTT TTT GCC CTC TGG
ala arg asn trp phe tyr pro ser his asn tyr his gln gly met ser tyr leu asn val arg phe thr asn asp leu phe ala leu trp
811/271                               841/281                               871/291
ATA TTA GAC CGG ATG GGC GCT GGT AAT GTG TTT AAT CCA GGG CAG CAG TTT ATC CTT TAT GAC GCG ATC TAT AAA CGC CGC CCC GAT GGA
ile leu asp arg met gly ala gly asn val phe asn pro gly gln gln phe ile leu tyr asp ala ile tyr lys arg arg pro asp gly
901/301                               931/311                               961/321
CAG ATT TTA GCA GGT GGA GAT GTA GAT TAT TCC AGG AAA AAA CCA AAA TAT TAT ACG ATG CCT GCA TTG CTT GCA GGT AGC TAT TAT AAA
gln ile leu ala gly gly asp val asp tyr ser arg lys lys pro lys tyr tyr thr met pro ala leu leu ala gly ser tyr tyr lys
991/331                               1021/341                              1051/351
GAT GAA TAC CTT AAT TAC GAA TTC CTG AAA GAT CCC AAT GTT GAG CCA CAT TGC AAA TTG TTC GAA TTT TTA TGG CGC GAT AGC CAG TTG
asp glu tyr leu asn tyr glu phe leu lys asp pro asn val glu pro his cys lys leu phe glu phe leu trp arg asp thr gln leu
1081/361                              1111/371                              1141/381
GGA AGT CGT AAG CCT GAT GAT TTG CCA CTT TCC AGG TAC TCA GGA TCG CCT TTT GGA TGG ATG ATT GCC GTA CCG GAT GGG GTC CGG GAA
gly ser arg lys pro asp asp leu pro leu ser arg tyr ser gly ser pro phe gly trp met ile ala val pro asp gly val arg glu
1171/391                              1201/401                              1231/411
GGT GTG ATT GCA GAG ATG AAA GTC AAC GAA TAT TCC TTT CTT AAC CAT CAG CAT CAG GAT GCA GGA GCC TTC CAG ATC TAT TAC AAA GGC
gly val ile ala glu met lys val asn glu tyr ser phe leu asn his gln his gln asp ala gly ala phe gln ile tyr tyr lys gly
1261/421                              1291/431                              1321/441
CCG CTG GCC ATA GAT GCA GGC TCG TAT ACT GGT TCT TCA GGA GGT TAT AAC AGT CCG CAC AAC AAG AAC TTT TTT AAG CGG ACT ATT GCA
pro leu ala ile asp ala gly ser tyr thr gly ser ser gly gly tyr asn ser pro his asn lys asn phe phe lys arg thr ile ala
1351/451                              1381/461                              1411/471
CAC AAT AGC TTG CTG ATT TAC GAT CCT AAA GAA ACT TTC AGT TCG TCG GGA TAT GGT GGA AGT GAC CAT ACC GAT TTT CCT GCC AAC GAT
his asn ser leu leu ile tyr asp pro lys glu thr phe ser ser ser gly tyr gly gly ser asp his thr asp phe pro ala asn asp
1441/481                              1471/491                              1501/501
GGT GGT CAG CGG CTG CCC GGA AAA GGT TGG ATT GCA CCC CGC GAC CTT AAA GAA ATG CTG GCA GGC GAT TTC AGG ACC GGC AAA ATT CTT
gly gly gln arg leu pro gly lys gly trp ile ala pro arg asp leu lys glu met leu ala gly asp phe arg thr gly lys ile leu
1531/511                              1561/521                              1591/531
GCC CAG GGC TTT GGT CCG GAT AAC CAA ACC CCT GAT TAT ACT TAT CTG AAA GGA GAC ATT ACA GCA GCT TAT TCG GCA AAA GTG AAG GAA
ala gln gly phe gly pro asp asn gln thr pro asp tyr thr tyr leu lys gly asp ile thr ala ala tyr ser ala lys val lys glu
1621/541                              1651/551                              1681/561
GTA AAA CGT TCA TTT CTA TTC CTG AAC CTT AAG GAT GCC AAA GTT CCG GCA GCG ATG ATC GTT TTT GAC AAG GTA GTT GCT TCC AAT CCT
val lys arg ser phe leu phe leu asn leu lys asp ala lys val pro ala ala met ile val phe asp lys val val ala ser asn pro
1711/571                              1741/581                              1771/591
GAT TTT AAG AAG TTC TGG TTG TTG CAC AGT ATT GAG CAG CCT GAA ATA AAG GGG AAT CAG ATT ACC ATA AAA CGT ACA AAA AAC GGT GAT
asp phe lys lys phe trp leu leu his ser ile glu gln pro glu ile lys gly asn gln ile thr ile lys arg thr lys asn gly asp
1801/601                              1831/611                              1861/621
AGT GGG ATG TTG GTG AAT ACG GCT TTG CTG CCG GAT GCC CGG AAT TCA AAC ATT ACC TCC ATT GGC GGC AAG GGC AAA GAC TTC TGG GTG
ser gly met leu val asn thr ala leu leu pro asp ala arg asn ser asn ile thr ser ile gly gly lys gly lys asp phe trp val
1891/631                              1921/641                              1951/651
TTT GGT ACC AAT TAT ACC AAT GAT CCT AAA CCG GGC ACG ATG AAG CAT TGG AAC GTG GAG AAT GGC GTG TGG AAA TCA CTC CAA AAA AGG
phe gly thr asn tyr thr asn asp pro lys pro gly thr met lys his trp asn val glu asn gly val trp lys ser leu gln lys arg
1981/661                              2011/671                              2041/681
CAG CAG CAG GAG GAT TAC TAC CTG AAT GTG ATA CAG ATT GCC GAC AAT ACA CAG CAA AAA TTA CAC GAG GTG AAG CGT ATT GAC GGT GAC
gln gln gln glu asp tyr tyr leu asn val ile gln ile ala asp asn thr gln gln lys leu his glu val lys arg ile asp gly asp
2071/691                              2101/701                              2131/711
AAG GTT GTT GGT GTG CAG CTT GCT GAC AGG ATA GTT ACT TTT AGC AAA ACT TCA GAA ACT GTT GAT CGT CCC TTT GGC TTT TCC GTT GTT
lys val val gly val gln leu ala asp arg ile val thr phe ser lys thr ser glu thr val asp arg pro phe gly phe ser val val
2161/721                              2191/731                              2221/741
GGT AAA GGA ACA TTC AAA TTT GTG ATG ACC GAT CTT TTA CCC GGT ACC TGG CAG GTG CTG AAA GAC GGA AAA ATA CTT TAT CCT GCG CTT
gly lys gly thr phe lys phe val met thr asp leu leu pro gly thr trp gln val leu lys asp gly lys ile leu tyr pro ala leu
2251/751                              2281/761                              2311/771
TCT GCC AAA GGT GAT GAT GGA GCC CTT TAT TTT GAA GGA ACT GAA GGA ACC TAC CGT TTT TTG AGA TAA
ser ala lys gly asp asp gly ala leu tyr phe glu gly thr glu gly thr tyr arg phe leu arg OCH
```

Figure 7.6: *ORF encoding the gene for heparinase II from F. heparinum*

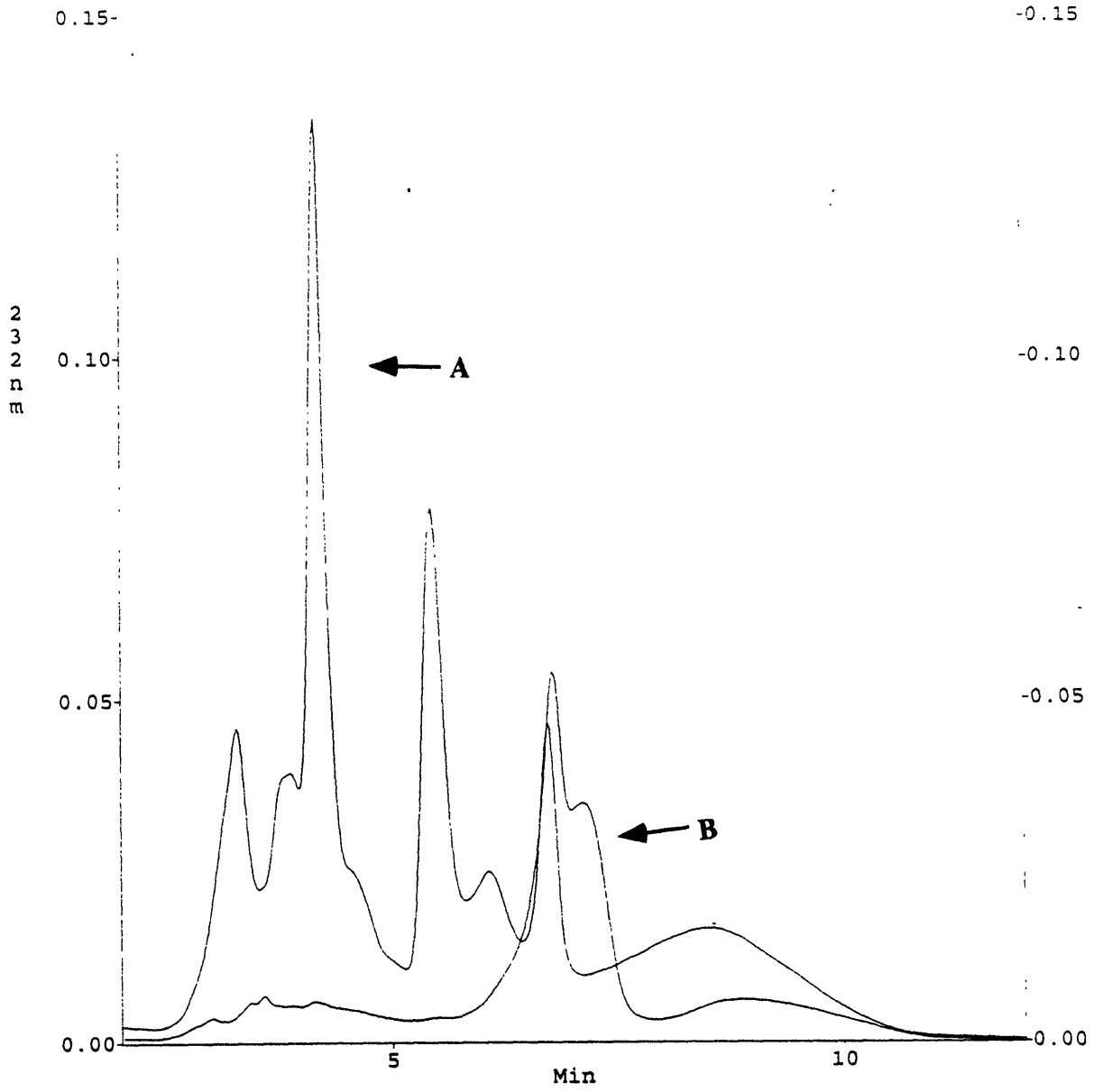


Figure 7.7: *Recombinant heparinase II activity*
A. shows the product profile of overnight degradation of *r*-heparinase II soluble crude extract with heparin; **B** shows the profile of *E. coli* soluble crude extract with heparin.

sulfated disaccharides) produced from heparinase I digestion of heparin and separated under the same gradient.

7.3.2 Discussion

The heparinase II gene contains a 2316 bp ORF and eleven heparinase II tryptic peptides representing 23 % of the protein sequence mapped onto the ORF. Conspicuously absent from the heparinase II amino acid sequence was the typical prokaryotic signal sequence with an Ala xxx Ala cleavage site (On Heijne, 1988). Such a cleavage site was observed in heparinase I and III protein sequences. Heparinase II like the other heparinases is also N-terminally blocked suggesting post-translational modifications. Further, similar to observations made for heparinases I and III, there is a discrepancy in the molecular weight calculated from the sequence and as observed from mass spectrometry. It is possible that this is due to additional protein post-translational modifications (see below). For heparinase II the molecular weight deduced from the amino acid sequence (87.9 kD) is approximately 2 kD higher than that obtained from mass spectrometry, suggesting that the protein gets proteolytically processed and the true start site is not known. There are two potential glycosylation sites based on the consensus DS sequence observed for heparinase I (Huang, 1995) and one such site in heparinase III.

Heparinase II was recombinantly expressed using pET-28a expression system. A major portion of the expressed protein was insoluble and activity could be recovered in the crude soluble fraction. The product profile of heparinase II degrading heparin revealed two disaccharides with elution times identical to the disulfated and trisulfated disaccharides produced by heparinase I acting on heparin.

Heparinase II, contains three cysteine residues and three putative heparin binding consensus sequences. Interestingly one of the consensus sequences shares a high degree of homology to the consensus sequence in heparinase I **(Table 7.5)** which was shown to be important for activity (chapter 5). It will be interesting to see if the histidine or positive charges from this region are important for heparinase II activity.

Table 7.5 Heparin binding consensus sequences of heparinases I and II

Heparinase I consensus sequence	196 **IFKKNIAH** 203
Heparinase II consensus sequence	444 **FFKRTIAH** 451

It has been reported that heparinase II is relatively the most stable among the three heparinases (Lohse and Linhardt, 1992). It is possible that two of the three cysteines are involved in a di-sulfide bond to stabilize the enzyme while a third is involved in catalysis. From the hydrophilicity and surface probability charts for heparinase II it appears that cysteine-348 is in a more hydrophilic region and has a higher probability of being surface accessible and hence could possibly be a candidate for the active site. As pointed out earlier for heparinase I, the surface accessible cysteine-135 was involved in catalysis (Sasisekharan et al., 1995). It will be interesting to identify the catalytic domain of heparinase II to understand the versatility of this enzyme in being able to cleave linkages involving both isoforms of uronic acid.

It is known that calcium catalytically activates heparinases I and III and is not required for heparinase II activity (Lohse and Linhardt, 1992). Consistent with this observation there are calcium coordinating motifs in heparinases I and III which satisfy the 'EF-hand' structural domain and no such sequences in heparinase II.

The promoter region sequences for heparinases I, II and III were compared for homologous domains. Several interesting stretches of homologous sequences were observed. Among them, heparinases I and II had a CAAT box at the -13 region and heparinases I, II and III share a TAACC box at close to the -30 region. Not much is known about the regulatory sequences in *F. heparinum* in general.

Comparison of heparinases I, II and III sequences showed little sequence homology at the DNA and the amino acid levels. It is interesting to note that heparinase II, which is able to cleave both heparin and heparan like sequences, also shares little homology with both enzymes. The sequences of the three genes when compared to the Genbank database, revealed no significant homology with any other protein. It is possible that the heparinases, like the pectate lyases, share a structurally similar three-dimensional scaffold.

Chapter 8

Conclusions & Significance

A framework was provided, to not only address structure-activity relationship issues of heparinase I but also to create a battery of mutant heparinases with altered properties. This chapter lists the conclusions from the work and describes the overall significance of the work and its relevance to biotechnology and biochemical engineering. Some recommendations for future work are provided.

8.1 Conclusions

In conclusion, for this thesis work:

1) A rational experimental approach to rapidly engineer mutations in heparinase I was developed by applying protein engineering tools.

2) A rapid screening assay system was developed to measure, both qualitatively and quantitatively, the effect of these changes on heparinase I properties.

3) The use of the above experimental approach to understand the enzyme's catalytic mechanism and relationship between the surface properties (structure) and function (activity) was demonstrated.

4) A framework was provided, to not only address structure-activity relationship issues of heparinase I, but also an opportunity to create a battery of mutant heparinases with a) altered substrate specificity, b) altered catalytic efficiency, and c) improved thermostablility.

   This would eventually lead towards creating "designer enzymes" for specific applications including thermostable heparinase I or stable heparinase I immobilization for prolonged bioreactor use in applications such as chronic dialysis.

More specifically, a systems approach towards elucidating structure-activity relationships of heparinase I was applied. A combination of chemical modification and site-directed mutagenesis approaches were applied to determine the structure and properties of the active site and substrate binding site of heparinase I. The key conclusions from these experiments are as follows:

- **Active Site**

1) Cysteine-135 and histidine-203 are part of the catalytic domain of heparinase I.

2) Cysteine-135 has an unusually low pKa (~5.5), while histidine-203 has an unusually high pKa (~ 8.5).
3) Cysteine-135 is present in a positively charged environment and is perhaps stabilized by the network of positive charges.
4) Cysteine-135 perhaps plays a nucleophilic role and histidine-203 acts as an acid/base catalyst in the catalytic mechanism.

- **Heparin Binding Domain**

1) Heparin binding site is in proximity to cysteine-135. Basic residues of heparin binding site perhaps provide the positively charged environment.
2) Lysines-198 and -199 from the primary heparin binding site are important for heparinase I activity.
3) Positive charge of lysine-132 is also important for heparinase I activity and is possibly part of the heparin binding domain.

- **Calcium Coordination Loop**

1) Conserved residue glycine-213 from the calcium binding loop in the heparin binding site is important for heparinase I activity, possibly due to its conformational flexibility.
2) Aspartic acids-210, -212 and glutamic acid-207 are possibly involved in calcium coordination.

- **Heparinases II and III**

The genes for heparinases II and III were cloned, sequenced and expressed in *E. coli*. The key conclusions from this work are:

1) Heparinase III gene comprises 1980 bps (660 amino acids) while heparinase II is made up of 2316 bps (772 amino acids).
2) Heparinase III amino acid sequence does not contain any cysteine residues; heparinase II has three cysteine residues.
3) Heparinase III does not possess any heparin binding consensus sequences as was observed with heparinase I. On the other hand, heparinase II possesses

three putative heparin binding consensus sequences, with one of them sharing homology to the heparinase I consensus sequence.

4) Two potential consensus calcium binding loops were identified in the heparinase III sequence while heparinase II did not possess any such sequences.

8.2 Overall Significance

- ***Protein engineering—a multidisciplinary approach to molecular design***

The biotechnology industry uses enzymes in various catalytic reactions and the industrial enzyme market is currently valued at over $ 1 billion. One of the greatest challenges in biochemical engineering and biotechnology has been to be able to engineer enzymes with desired and improved properties such as enhanced activity; altered specificity (selectivity); improved stability etc. A highly multi-disciplinary approach to molecular design, involving biology, chemistry and chemical engineering, is emerging. The ultimate goal is to produce cost-effective and novel applications of modified enzymes in various areas such as food, pharmaceutical, detergents, textiles, pulp and paper, and specialty chemicals.

Protein engineering is an example of a multi-disciplinary approach, where using knowledge of enzyme structure and defining its ligand binding surface would enable one to develop a rational approach towards engineering improved properties such as stability, activity or potency in enzymes, in order to adapt them for clinical and industrial applications through mutagenesis (Smith, 1985; Blow et al., 1986; Blundell, 1994). Examples include stabilization of proteins (Arnold, 1994); altering enzymes to function under different pH conditions to better suit an industrial process; altered enzyme specificities (El Hawrani et al., 1994); and development of stable, calcium insensitive protease in biological detergents (Procter & Gamble, OH; GEN, 1996).

Progress towards this front for heparinase I has been limited by a lack of fundamental understanding of the enzyme's catalytic and kinetic mechanism and relationship between the surface properties (structure) and function (activity). This work represents a significant progress made towards understanding the structure-

activity relationship of heparinase I, with the eventual goal of using this information and protein engineering techniques in creating "designer heparinases I" for a variety of applications.

8.3 Recommendations for Future Work

Heparinases I, II and III

This work has resulted in the elucidation of the active site structure and the substrate binding cleft in heparinase I. The substrate binding cleft (containing the catalytic domain) is positively charged, due to highly anionic substrate heparin. This charged environment, in the enzyme's substrate binding cleft, also influences the active site and, as a result, the catalytic mechanism. Thus, understanding the structure and environment of the substrate binding cleft can result in a better fundamental understanding of the enzyme's catalytic properties and mechanism.

The approach developed in this research can be extended to heparinases II and III to address several interesting questions. As was discussed in chapter 7, the primary sequences of heparinases I, II and III revealed some interesting similarities and differences. For example, heparinases I and II have positively charged heparin binding consensus sequences while heparinase III does not. Heparinases I and II have cysteine residues, (with cysteine-135 in heparinase I in the catalytic domain) while heparinase III does not. On the other hand, heparinases I and III are activated by calcium and possess calcium binding consensus motifs, while such motifs are absent in heparinase II. The charge properties of the substrate could explain these observations. Heparinases I and II cleave heparin (highly negatively charged substrate) and hence possibly have positively charged clusters while heparinase III primarily cleaves heparan sulfate like regions which are not as negatively charged. These sequence comparisons lead to several questions such as:

- What is the basis for the unique substrate specificity of heparinases I, II and III? Is it driven by the charge properties of the substrate?

- Does substrate binding influence specificity or does calcium binding to the substrate facilitate molecular recognition?

- How do the properties of the substrate binding clefts in heparinases II and III compare to that of heparinase I?

- How does the heparinase II active site (catalytic domain) possess the unique versatility to cleave both heparin like and heparan like regions of the substrate?

- It is known that heparinase I while cleaving heparin leaves behind tetra- and hexasaccharides while heparinase II leaves behind primarily disaccharides. Can the catalytic and substrate binding domain structures explain this difference in kinetic mechanisms? Does the size of the substrate binding cleft dictate the cleavage patterns?

- This work described a possible mechanism for heparin degradation by heparinase I—is there a unifying mechanism that could be extended not only to heparinases II and III but to other complex polysaccharide degrading lyases as well?

- Do the three enzymes share a similar three-dimensional scaffold, similar to that seen with the pectate lyases?

To address the above questions, the approaches and experiments described in this work could be applied to heparinases II and III. Using chemical modification and site-directed mutagenesis approaches the active site and the heparin binding domain of heparinases II and III can be isolated. It is possible that a free cysteine is involved as a nucleophile in the heparinase II active site. Isolating the catalytic domain would provide important insight into the catalytic versatility of heparinase II. Swapping the heparin binding domains of heparinases I and II would answer the question of the involvement of heparin binding in modulating substrate specificity.

Heparinase I

Several interesting questions with heparinase I—at a fundamental level and from an applications perspective, can be addressed. This work described a catalytic mechanism for heparin degradation; however, an understanding of the kinetic mechanism is lacking. The mutant enzymes can prove to be valuable tools towards developing this model. The questions that need to be addressed include:

- What is the kinetic mechanism for heparinase I? Is it purely random endolytic, exolytic or a combination of both?

- Using the mutant enzymes, can a kinetic model be developed to explain and predict the product distribution for heparinase I?

- What are the large fragments produced by the enzymes with mutations in the heparin binding site i. e. K198AK199A and K198DK199D? Do they represent unique fragments of heparin (i. e. altered substrate specificity) or are they larger, uncleaved heparin molecules resulting from cleavage by a slower enzyme?

- The role of calcium was investigated in this work and a putative consensus calcium binding domain was identified. However, what is the stoichiometry of calcium binding i. e. how many calcium ions are bound to a molecule of heparinase I? What is the role of the C-terminal calcium binding consensus motif?

These questions can be addressed using the following approaches and experiments:

- Stoichiometry of calcium binding to heparinase I can be determined by fluorescence experiments. The role of the side-chain oxygen containing residues in the C-terminal calcium binding motif in heparinase I needs to investigated through extensive site-directed mutagenesis. Double mutations of such residues in the first consensus motif can be created to impair calcium binding. These experiments will confirm the role of the calcium binding loops in heparinase I activity.

- To develop a kinetic model, heparin degradation with wild-type and mutant heparinases can be studied at various time points by observing the product distribution by anion-exchange HPLC. Using a combination of the different mutant enzymes which cleave heparin at different rates, and by stopping the reaction at various times, the cleavage fragments can be isolated, mapped and characterized. An empirical kinetic model can be developed to predict product distributions by assuming an average chain length for heparin and a cleavage rate from the kcat values of the respective mutant enzymes. This approach would also help in confirming the cleavage pattern and mechanism of heparinase I, i. e. endolytic or exolytic or both.

- The fragments of heparin degradation generated from K198AK199A; K198DK199D and other mutants that cleave heparin differently compared to wild type heparinase I, can be collected, desalted and characterized by mass spectrometry. It is possible that these fragments could have potential therapeutic value as LMW heparins, which in turn could have significant clinical implications as described previously (Linhardt and Loganathan, 1989).

From an applications perspective, it was mentioned in chapter 2 that the two major applications of the heparinases are as inhibitors of angiogenesis and in enzyme immobilization (heparinase I). Some important questions that need to be addressed include:

- Heparinase I and III are potent inhibitors of angiogenesis, while heparinase II has no effect. It is believed that this differential effect is due to the different specificities and fragment sizes left behind on cleaving heparin like molecules. The mutant enzymes have different cleavage patterns and it would be interesting to see if they behave differently in such models.

- Is it possible to engineer heparinase I for creating an enzyme with improved stability; enhanced activity, or altered specificity? Stable enzymes would be useful in immobilization applications involving prolonged use of the enzyme such as acute hemodialysis.

- Is it possible to engineer heparinase I for improved immobilization?

- Since cysteine, histidine, and lysines are involved in heparinase I activity, chemistries involving these residues for immobilization would result in reduced immobilization efficiency. Could other chemistries, for example involving carboxylic acid containing residues such as asp or glu, be applied for heparinase I immobilization?

Approaches to address these questions include:

- The mutant enzymes that cleave heparin leaving larger fragments, can be purified and used for CAM and endothelial cell proliferation assays to study their behavior in such models.

- For heparinase I immobilization, chemistries involving amines, thiol or imidazole results in extensive activity losses. The surface residues of heparinase I have been elucidated in this study and suggests that chemistries involving carboxylic acids can be investigated as an approach towards immobilization. This chemistry is powerful since it results in strong peptide bonds between the enzyme acidic residues and the support, which will significantly reduce enzyme losses through leaching from the support.

- Further site-directed mutagenesis experiments can be targeted at the catalytic and substrate binding domains of heparinase I to improve properties such as stability, activity, and specificity since these domains represent the surface of the enzyme. For example:

1) The surface residues (i. e. residues of the primary heparin binding site or the residues close to cysteine-135) could be altered to modulate activity and pH dependence. The net positive charge on the surface in proximity to histidine-203; cysteine-135 and lysines-198 and -199 can be enhanced by suitable mutations (e.g. Asp to Asn; Glu to Gln) to improve substrate binding to negatively charged heparin and hence catalytic activity.

2) A double mutation containing two adjacent histidine residues within the heparin binding site (but away from histidine-203) can be introduced to facilitate heparinase I immobilization through metal chelation.

3) For improved stability, a random mutagenesis approach, targeting the residues of the heparin binding domain, can be applied. This approach, using either phage display or PCR based random mutations, is very powerful since it would yield a library of mutants with various mutations within the heparin binding domain. The mutants can be screened for thermally stable enzymes and sequenced and can also be used for screening for other properties as required.

Thus, understanding the heparin-heparinase interaction can have several significant implications. It enables one to rationally engineer improved properties such as stability, altered specificities, enhanced catalytic efficiencies etc. In addition, modified heparinases could be a useful tool in elucidating heparin structure information; in heparin sequencing and in generating LMW heparin fragments with potential therapeutic value. One can eventually conceive the possibility of designing smaller enzyme mimics in future, which could utilize the groove formed by the catalytic and substrate binding sites in a compact structure.

References

(1995) Genetic Engineering News. **15:** 29.

(1996). Genetic Engineering News. **16:** 1.

Abdulwajid, A. W., and Wu, F. Y. -H. (1986). *Biochemistry.* **25**: 8167-8172.

Ameer, G. (1996). *Personal Communication*

Amuro, N., Paluh, J. L., and Zalkin, H. (1985). *J. Biol. Chem.* **260**: 14844-14849.

Arnold, F. H. (1993). *FASEB J.* **7**: 744-49.

Atha, D. H., Lormeau, J.-C., Petitou, M., Rosenberg, R. D. and Choay, J. (1985). *Biochemistry.* **24**: 6723-29.

Bae, J., Desai, U., Pervin, A., Caldwell, E., Weiler, J. and Linhardt, R. (1994). *Biochem. J.* **301**: 121-29.

Baird, A., Schubert, D., Ling, N. and Guillemin, R. (1988). *Proc. Nat. Acad. Sci., USA.* **85**: 2324-28.

Battaglia, E., Pritchard, M., Ouzzine, M., Fournel-Gigleux, S., Radominska, A., Siest, G., and Magdalou, J. (1994).. *Arch. Biochem. Biophys.* **309**(2): 266-272.

Barrer, R. M. (1982). Hydrothermal Chemistry of Zeolites. London, Academic Press.

Belkin, S. and. Jannasch., H. W. (1985). *Arch. Microbiol.* **141**: 181-186.

Bellamy, R. W. and Horikoshi, K. (1992). *United States Patent.* **5,145,778**:

Bengtsson, G., Olivecrona, T., Hook, M., Riesenfeld, J., and Lindahl, U. (1980). *Biochem. J.* **189**: 625-633.

Bernstein, H., Yang, V. C., and Langer, R. (1987).*Appl. Biochem. and Biotech..* **16**: 129-43.

Bernstein, H., Yang, V. C., Cooney, C. L., and Langer, R. (1988). *Meth. Enzym.* **137**: 515-29.

Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C., Phillips, D. C., and Sarma, V. R. (1965). *Nature.* **206**: 757.

Blow, D. M. (1976). *Acc Chem Res.* **9**: 145-152.

Blow, D. M., Fersht, A. R., and Winter, G. (1986). Design, Construction and Properties of Novel Protein Molecules.

Blundell, T. L. (1994). *Trends Biotechnol.* **12**: 145-48.

Bober Barkalow, F. J. and Schwarzbauer, J. E. (1991). *J. Biol. Chem.* **266**: 7812-18.

Boyd, J., Williamson, F. W. and Gettins, P. (1980). *J. Mol. Biol.* **137**: 175-90.

Brady, L., Brzozowski, A. M., Derewenda, Z. S., Dodson, E.,Dodson, G., Tolley, S., Turkenburg, J. P., Christiansen, L., Huge-Jensen, B., Norskov, L., Thim, L., and Menge, U. (1990). *Nature*. (343): 767-770.

Breck, D. W. (1974). Zeolite Molecular Sieves. New York, Wiley-Interscience.

Brocklehurst, K., Willenbrock, F., and Salih, E. (1987). Hydrolytic Enzymes. Amsterdam, Elsevier.

Bryan, P. N., Rollence, M. L., Pantaliano, M. W., Wood, J., Finzel, B. C., Gilliland, G. L., Howard, A. J., and Poulas, T. L. (1986). *Proteins: Struct. Func. Gent.* **1**: 326-334.

Cardin, A. D. and Weintraub, H. J. R. (1989). *Arteriosclerosis*. **9**: 21-32.

Carter, P., and Wells, J. A. (1988). *Nature*. **332**: 564-568.

Catellot, J. J. J., Beeler, D. L., Rosenberg, R. D., and Karnovsky, M. J. (1984). *J. Cell. Physiol.* **120**: 315.

Catellot, J. J. J., Choay, J., Lormeau, J. -C., Petitou, M., Sache, E., and Karnovsky, M. J. (1986). *J. Cell Biol.* **102**: 1979.

Celesia, G. G. (1991). *Sem. Thromb. Hem.* **17 (suppl 2)**: 158-60.

Cerbelaud, E. C., Conway, L. J., Galliher, P. M., Langer, R. S. and Cooney, C. L. (1986). *Appl. Env. Micr.* **51**: 640-46.

Chaga, G., Widersten, M., Andersson, L., Porath, J., Danielson, U. H., and Mannervik, B. (1994). *Protein Engineering*. **7**: 1115-1119.

Chaiken, I. M., and Smith, E. L.(1969) *J. Biol. Chem.* **244**, 5095-5099

Chang, C. D., Hellring, S. D., Miale, J. N., Bridger, R. F., and Calvert, R. B. (1985). *J. Chem. Soc. Faraday Trans. I,*. **81**: 2215.

Chen, K., and Arnold, F. H. (1993). *Proc. Natl. Acad. Sci. USA.* :

Cheng, C.-F., Oosta, G. M., Bensadoun, A., and Rosenberg, R. D. (1981). *J. Biol. Chem.* **256**: 12893-12898.

Chipman, D. M., and Sharon, N. (1969). *Science*. **165**: 454.

Cohen, F. E. (1993). *Science*. **260**: 1444-45.

Coltrini, D., Rusnati, M., Zoppetti, G., Oreste, P., Isacchi, A., Caccia, P., Bergonzoni, L. and Presta, M. (1993). *Eur. J. Biochem.* **214**: 51-58.

Comfort, A. R., Albert, E. and Langer, R. (1989). *Biotech. Bioeng.* **34**: 1374-82.

Comper, W. D. (1981). Heparin (and related polysaccharides). New York, Gordon & Breach.

Craik, C. S., Roczniak, S., Largman, C., and Rutter, W. J. (1987). *Science*. **237**: 909-913.

Csicsery, S. M. (1984). *Zeolites.* **4**: 202.

Cunningham, B. C., and Wells, J. A. (1987). *Protein Engineering.* **1**: 319-325.

Dalhquist, F. W., Rand-Meir, T., and Raftery, M. A. (1968). *Proc. Natl. Acad. Sci USA.* **61**: 1194.

Davis, R. C., Wong, H., Nikazy, J., Wang, K., Han, Q., and Schotz, M.C. (1992). *J. Biol. Chem.* **267**: 21499-21504.

Denu, J., Zhou, G., Guo, Y., and Dixon, J. (1995) *Biochemistry.* **34**: 3396-3403.

Desai, U. R., Wang, H. and Linhardt, R. J. (1993a). *Biochemistry.* **32**: 8140-45.

Desai, U. R., Wang, H. and Linhardt, R. J. (1993b). *Arch. Biochem. Biophys.* **306**: 461-68.

Dessau, R. M., and Kerr, G. T. (1984). *Zeolites.* **4**: 315.

Dietrich, C. P., Silva, M. E. and Michelacci, Y. M. (1973). *J. Biol. Chem.* **248**: 6408-15.

Dipersio, L., Fontaine, R. N., and Hui, D. Y. (1991). *J. Biol. Chem.* **266**: 4033-4036.

Duncan, T. M., and Dybowski, C. (1981). *Surf. Sci. Reports.* **1**: 157.

El Hawrani, A. S., Moreton, K. M., Sassions, R. B., Clarke, A. R., and Holbrook, J. J. (1994). *Trends Biotechnol.* **12**: 207-211.

Erdodi, F., Gergely, P., and Bot, G. (1984). *Int. J. Biochem.* **16**: 1391.

Ernst, E., Venkataraman, G., Winkler, S., Godavarti, R., Langer, R., Cooney, C., & Sasisekharan. R. (1996). *Biochem J.* **In Press**:

Ernst, S. E., Langer, R. S., Cooney, C. L., & Sasisekharan, R.,. (1995). *Critical Reviews in Biochemistry and Molecular Biology.* **30**((5)): 387-444.

Fersht, A. (1985). Enzyme Structure and Mechanism. New York, W.H. Freeman and Co.

Folkman, J., Langer, R., Linhardt, R., Haudenschild, C., and Taylor, S. (1983). *Science.* **221**: 719-725.

Folkman, J., and Shing, Y. (1992). *J. Biol. Chem.* **267**: 10931-10934.

Fu, D. and. Robyt, J. F. (1988). *Carbohydr. Res.* **183**: 97-109.

Gacesa, P. (1987). *FEBS lett.* **212**: 199-202.

Gallagher, J. T., Lyon, M. and Steward, W. P. (1986). *Biochem. J.* **236**: 313-25.

Gallagher, J. T., Turnbull, J. E. and Lyon, M. (1992). *Int. J. Biochem.* **24**: 553-60.

Garrett, R., and Grisham, C. (1995). Biochemistry. Florida, Saunders College Publishers.

Getzoff, E. D., Cabelli, D. E., Fisher, C. L., Parge, H. E., Viezzoli, M. S., Banci, L. and Hallewell, R. A. (1992). *Nature*. **358**: 347-351.

Ghadiri, M. R., and Choi, C. (1990). *J. Am. Chem. Soc.* **112**: 1630-1632.

Gitay-Goren, H., Soker, S., Vlodavsky, I. and Neufeld, G. (1992). *J. Biol. Chem.* **267**: 6093-98.

Gospodarowicz, D., Ferrara, N., Schweigere, L., and neyfield, G. (1987). *Endocr. Rev.* **8**: 95.

Greiling, H., Stuhlsatz, W., Eberhard, T. and Eberhard, A. (1975). *Conn. Tissue Res.* **3**: 135-39.

Gross, M., Gosnell, M., Tsarbopoulos, A., and Hunziker, W. (1993). *J. Biol. Chem.* **268**: 20917-20922.

Guan, K., and Dixon, J. (1990). *Science*. **249**: 553-556.

Guimond, S., Maccarana, M., Olwin, B. B., Lindahl, U. and Rapraeger, A. C. (1993). *J. Biol. Chem.* **268**: 23906-14.

Haag, W. O., and Chen, N. Y. (1987). Catalyst Design—Progress and Perspectives.John Wiley & Sons.

Habuchi, H., Suzuki, S., Saito, T., Tamura, T., Harada, T., Yoshida, K. and Kimata, K. (1992). *Biochem. J.* **285**: 805-13.

Hahn, M., Olsen, O., Politz, O., Borriss, R., and Heinemann, U. (1995). *J. Biol. Chem.* **270**: 3081-3088.

Halasz, P., and Polgar, L. (1977). *Eur. J. Biochem.* **79**, 491-94

Halliwell, B., and Gutteridge, J. M. C. (1989). Free Radicals in Biology and Medicine. Oxford, Clarendon Press.

Hamai, A., Morikawa, K., Horie, K. and Tokuyasu, K. (1989). *Agric. Biol. Chem.* **53**: 2163-68.

Handel, T., and DeGrado, W. F. (1990). *J. Am. Chem. Soc.* **112**: 6710-11.

Hardingham, T. E. and Fosang, A. J. (1992). *FASEB J.* **6**: 861-70.

Hata, A., Ridinger, D., Sutherland, S., Emi, M., Shuhua, Z., Myers, R., Ren, K., Cheng, T., Inoue, I., Wilson, D., Iverius, P., and Lalouel, J. (1993). *J. Biol. Chem.* **268**: 8447-8457.

Hedman, K., Johansson, S., Vartio, T., Kjellen, L., Vaheri, A., and Hook, M. (1982). *Cell.* **28**: 663-671.

Hedstrom, L., Szilagyi, L., and Rutter, W. J. (1992). *Science*. **255**: 1249-1253.

Hedstrom, L., Perona, J. J., and Rutter, W. J. (1994). *Biochemistry*. **28**: 8757-8763.

Hegedus, L. L. (1987). Catalyst Design—Progress and Perspectives.

Higuchi, R. (1990). Recombinant PCR. PCR Protocols: A Guide to Methods and Applications. San Diego, CA, Academic Press, Inc.

Hinton, J. C. D., Sidebotham, D. R., Gill, G. P. and Solmond, C. (1989). *Mol. Microbiology.* **3**: 1785.

Hiyama, K. and Okada, S. (1975a). *J. Biochem.* **78**: 1183-90.

Hiyama, K. and Okada, S. (1975b). *J. Biol. Chem.* **250**: 1824-28.

Hiyama, K. and Okada, S. (1976). *J Biochem.* **80**: 1201-07.

Hiyama, K. and Okada, S. (1977). *J Biochem.* **82**: 429-36.

Hjalmarsson, K., Marklund, S. L., Engstrom, A., Edlund, T. (1987). *Proc. Natl. Acad. Sci. U.S.A.* **84**: 6340-6344.

Hoogewerf, A. J., Leone, J W., Reardon, I M., Howe, W J., Asa, D., Heinrikson, R L., and Ledbetter, S R. (1995). *J. Biol. Chem.* **270**: 3268-3277.

Hovingh, P., and Linker, A. (1970). *J. Biol. Chem.* **245**: 6170-6175.

Huang, L., van Halbeek, H., Eggimann, B. and Zimmermann, J. (1995). *Glycobiology.* **5**: 712.

Hunter, G. K., Wong, K. S. and Kim, J. J. (1988). *Arch. Biochem. Biophys.* **260**: 161-67.

Imai, K., Shiomi, T., Uchida, K., and Miya, M. (1986). *Biotechnol. Bioeng.* **28**: 198-203.

Imoto, T., Johnson, L. N., North, A. C., Phillips, D. C., and Rupley, J. A. (1972). The enzymes.

Isihara, M., Tyrrell, D. J., Stauber, G. B., Brown, S., Cousens, L. S. and Stack, R. J. (1993). *J. Biol. Chem.* **268**: 4675-83.

Itoh, K. and Sokol, S. Y. (1994). *Development.* **120**: 2703-11.

Iwakura, M., and Kokubu, T. (1993). *J. Biochem.* **114**: 339-343.

Jackson, R. L., Busch, S. J. and Cardin, A. D. (1991). *Physiological Reviews.* **71**: 481-539.

Johnson, L. N., Cheetham, J., McLaughlin, P. J., Acharya, K.R., Barford, D., and Phillips, D. C. (1988). *Curr. Top. Microbiol. Immunol.* **139**: 81.

Johnson, L. N. and. Phillips, D. C. (1965). *Nature.* **206**: 761.

Jordan, R. E., Kilpatrick, J., and Nelson, R. M. (1987). *Science.* **237**: 777.

Kamphius, I. G., Kalk, K. H., Swarte, M. B., and Drenth, J. (1984). *J. Mol. Biol.* **179**: 233-257.

Kellis, J. T., Todd, R. J., and Arnold, F. H. (1991). *Bio/Technology.* **9**: 994-95.

Kirby, A. J. (1987).*CRC Crit. Rev. Biochem.* **22**: 283-312.

Kjellèn, L. and Lindahl, U. (1991). *Ann. Rev. Biochem.* **60**: 443-75.

Klagsburn, M., and D' Amore, P. A. (1991). *Annu. Rev. Physiol.* **53**: 217-239.

Kojima, T., Shworak, N. W. and Rosenberg, R. D. (1992). *J Biol. Chem.* **267**: 4870-77.

Kolisis, F. N., and Thomas, D. (1987). *Biotechnol. Bioeng.* **30**: 160-163.

Kong, H. M., Kucera, R. B., and Jack, W. E. (1993). *J. Biol. Chem.* **268**: 1965-1975.

Kraut, J. (1977). *Annu. Rev Biochem.* **46**: 331-358.

Kretsinger, R. H., Tolbert, D., Nakayama, S. and Pearson, W. :

Kuramitsu, S., Ikeda, K., and Hamaguchi, K. (1977). *J. Biochem (Tokyo).* **82**: 585.

Kuroki, R., Yamada, H., Moriyama, T., and Imoto, T. (1986). *J. Biol. Chem.* **261**: 13571.

Laemmli, U. K. (1970). *Nature.* **227**: 680-85.

Langer, R., Linhardt, R. J., Hoffberg, S., Larsen, A. K., Cooney, C. L., Tapper, D. and Klein, M. (1982). *Science.* **217**: 261-63.

Larsen, B., Skjak-Braek, G., and Painter, T. (1986) *Carbohydr. Res.* **146**: 342-45.

Lattera, J., Ansbacher, R., and Culp, L. A. (1980). *Proc. Natl. Acad. Sci. U. S. A.* **77**: 6662-6666.

Leckband, D., & Langer, R. (1991). *Biotechnol. Bioeng.* **37**: 227-237.

Lee, M. K. and Lander, A. D. (1991). *Proc. Natl. Acad. Sci., USA.* **88**: 2768-72.

Lewis, S. D., Johnson, F. A., and Shafer, J. A. (1976). *Biochemistry.* **15**: 5009-5017.

Lewis, S. D., Johnson, F. A., and Shafer, J. A. (1981). *Biochemistry.* **20**: 48-51.

Liao, H., McKenzie, T., and Hageman, R. (1986). *Proc. Natl. Acad. Sci.* **83**: 576-580.

Lindahl, U., Feingold, D., and Roden, L. (1986). *Trends. Biochem. Sci.* **11**: 221-225.

Lindahl, U., Lidholt, K., Spillmann, D. and Kjellén, L. (1994). *Thrombosis Research.* **75**: 1-32.

Lindahl, U., Thinberg, L., Bäckstrom, G., Riesenfeld, J., Nordling, K. and Björk, I. (1984). *J. Biol. Chem.* **259**: 12368-76.

Linhardt, R. J., Fitzgerald, G. L., Cooney, C. L. and Langer, R. (1982). *Biochim. Biophys. Acta.* **702**: 197-203.

Linhardt, R. J., Cooney, C. L., Tapper, D., Zannetos, C. A., Larsen, A. K. and Langer, R. (1984). *App. Biochem. Biotech.* **9**: 41-55.

Linhardt, R. J., Gallilier, P. M. and Cooney, C. L. (1986). *Appl. Biochem. Biotech.* **12**: 135-75.

Linhardt, R. J., Rice, K. G., Kim, Y. S., Lohse, D. L., Wang, H. M. and Loganathan, D. (1988). *Biochem. J.* **254**: 781-87.

Linhardt, R. J., and Loganathan, D. (1990). Biomimetic Polymers. New York, Plenum Press.

Linhardt, R. J., Turnbull, J. E., Wang, H. M., Loganathan, D. and Gallagher, J. T. (1990). *Biochemistry.* **29**: 2611-17.

Linhardt, R. J., Ampofo, S. A., Fareed, J., Hoppensteadt, D., Mulliken, J. B. and Folkman, J. (1992a). *Biochemistry.* **31**: 12441-45.

Linhardt, R. J., Wang, H.M., Loganathan, D. and Bae, J.H. (1992b). *J. Biol. Chem.* **267**: 2380-87.

Linhardt, R. J., Wang, H. M. and Ampofo, S. A. (1992c). New methodologies in heparin structure analysis and the generation of LMW heparins. Heparin and related polysaccharides. New York, Plenum Press.

Linker, A., and Hovingh, P. (1965). *J. Biol. Chem.* **240**: 3724-3728.

Linker, A. and Hovingh, P. (1972). *Meth. Enzymol.* **28**: 902-11.

Lohse, D. L. and Linhardt, R. J. (1992). *J. Biol. Chem.* **267**: 23347-55.

Lyon, M., Deakin, J. A., Mizuno, K., Nakamura, T. and Gallagher, J. T. (1994). *J. Biol. Chem.* **269**: 11216-23.

Maccarana, M., Casu, B. and Lindahl, U. (1993). *J. Biol. Chem.* **268**: 23898-905.

Maimone, M. M. and Tollefsen, D. M. (1990). *J. Biol. Chem.* **265**: 18263-71.

Malcolm, B. A., Rosenberg, S., Corey, M. J., Allen, J. S., de Baetselier, A., and Kirsch, J. F. (1990). *Proc. Natl. Acad. Sci. U. S.A.* **86**: 133.

Marchuk, D., Drumm, M., Saulino, A. and Collins, F. S. (1991). *Nucl. Acids Res.* **19**: 1154.

Marcum, J. A. and Rosenberg, R. D. (1989a). Heparin - chemical and biological properties clinical applications. London, The Biochemical Society.

Marcum, J. A. and Rosenberg, R. D. (1989b). *Ann. N. Y. Acad. Sci.* **556**: 81-94.

Marsden, B., Shaw, G., and Sykes, B. (1990). *Biochem. Cell Biol.* **68**: 587-601.

Mathews, C. K., and van Holde, K. E. (1990). Biochemistry. Redwood City, CA, The Benjamin/Cummings Publishing Company.

McLean, M. W., Long, W. F. and Williamson, F. B. (1985). *Proc. 8th Int. Symp. Glycoconjugates 1.* : 73-74.

Merchant, Z. M., Erbe, E. E., Eddy, W. P., Patel, D., and Linhardt, R. J. (1986). *Atheroscelorosis.* **62**: 151-158.

Messing, R. A. (1975). Immobilized Enzymes for Industrial Reactors.

Michelacci, Y. M. and Dietrich, C. P. (1974). *Biochem. Biophys. res. Commun.* **56**: 973-980.

Miles, E. W. (1977). *Methods Enzymol.* **47**: 431-442.

Miran, S. G., Chang, S. H., and Raushel, F. M. (1991). *Biochemistry.* **30**: 7901-7907.

Moffat, C. F., McLean, M. W., Long, W. F. and Williamson, F. B. (1991a). *Eur. J. Biochem.* **197**: 449-59.

Moffat, C. F., McLean, M. W., Long, W. F. and Williamson, F. B. (1991b). *Eur. J. Biochem.* **202**: 531-41.

Mooser, G. (1992). The Enzymes. Academic Press.

Munier, H., Bouhss, A., Krin, E., Danchin, A., Gilles, A-M., Glaser, P., and Barzu, O. (1992). *J. Biol. Chem.* **267**: 9816-9820.

Muraki, M., Morikawa, M., Jigami, Y., and Tanaka, H. (1988). *Protein Engineering.* **2**: 49-54.

Nader, H. B., Porcionatto, A., Tersariol, M. A., Pinhal, A. S., Fernanda, W. O., Moraes, C. T. and Dietrich, C. P. (1990). *J. Biol. Chem.* **265**: 16807-13.

Nakajima, M., Irimura, T. and Nicolson, G. L. (1988). *J Cell. Biochem.* **36**: 157-67.

Neurath, H. (1985). Proteolytic Enzymes, Past and Present. Fed. Proc. **44:** 2907-2913.

Nickbarg, E. B., Davenport, R. C., Pestko, G. A., and Knowles, J. R. (1988). *Biochemistry.* **27**: 5948-5960.

Nieduszynski, I. A. (1985). Connective tissue polysaccharides. Polysaccharides - Topics in structure and morphology. Weinheim, Germany, VCH.

Nugent, M. A., & Edelman, E. R. (1992). *Biochemistry.* **260**: 8876-8883.

Oliphant, A. R., and Struhl, K. (1989). *Proc. Natl. Acad. Sci.* **86**: 9094-9098.

Ornitz, D. M., Yayon, A., Flanagan, J. G., Svahn, C. M., Levi, E. and Leder, P. (1992). *Mol. Cell. Biol.* **12**: 240-47.

Ototani, N., Kikuchi, M. and Yosizawa, Z. (1981). *Carb. Res.* **88**: 291-303.

Pantoliano, M. W., Whitlow, M., Wood, J. F., Rollence, M. L., Finzel, B. C., Gilliland, G., Poulos, T. L., and Bryan, P. N. (1988). *Biochemistry.* **27**: 8311-17.

Parsons, S. M., and Raftery, M. A. (1972). *Biochemistry.* **11**: 1630.

Parthasarathy, N., Goldberg, I. J., Sivaram, P., Mulloy, B., Flory, D. M. and Wagner, W. D. (1994). *J. Biol. Chem.* **269**: 22391-96.

Pejler, G., Bäckström, G., Lindahl, U., Paulsson, M., Dziadek, M., Fujiwara, S. and Timpl, R. (1987). *J. Biol. Chem.* **262**: 5036-43.

Pelton, P. D., and Ganzhorn, A. J. (1992). *J. Biol. Chem.* **267**: 5916-5920.

Perlin, A. S. (1975). Proceedings of the International Symposium of Macromolecules. Amsterdam, Elsevier.

Perona, J. J., and Craik, C. S. (1995). *Protein Science.* **4**: 337-360.

Phillips, D. C. (1966). *Sci. Am.* **215**: 78.

Pickersgill, R., Jenkins, J., Harris, G., Nasser, W. and Robert-Baudouy, J. (1994). *Structural Biology.* **1**: 717-23.

Plummer, J., Tarentino, A. L & Hauer, C. R. (1995). *J. Biol. Chem.* **270**: 13192-96.

Pot, D., Woodford, T., Remboutsika, E., Haun, R., and Dixon, J. (1991) *J. Biol. Chem.* **266** 19688-96

Rabin, B. R., and Watts, D. C. (1960). *Nature.* **188**: 1163-1165.

Rapraeger, A. C. (1993). *Curr. Op. Cell Biol.* **5**: 844-53.

Rapraeger, A. C., Krufka, A. and Olwin, B. B. (1991). *Science.* **252**: 1705-08.

Redini, F., Tixier, J., Petitou, M., Choay, J., Robert, L., and Horenbeck, W. (1988). *Biochem. J.* **252**: 515-519.

Rees, A. R., Sternberg, M. J. E., and Wetzel, R. (1992). Protein Engineering— A Practical Approach.

Rees, D. A., Morris, E. R., Thom, D. and Madden, J. K. (1982). Shapes and interactions of carbohydrate chains. The Polysaccharides. New York, Academic Press.

Rice, K. G., and Linhardt, R. J. (1989). *Carbohdr. Res.* **190**: 219-233.

Riordan, J. F., and Vallee, B. (1972). *Methods. Enzymol.* **25B**: 449-456.

Robinson, J. H., and Browne, M. J. (1991). *Trends in Biotechnol.* **9**: 86-90.

Rosenberg, R. D., and Damus, P. S. (1973). *J. Biol. Chem.* **248**: 6490.

Rosenberg, S., and Kirsch, J. F. (1981). *Biochemistry.* **20**: 3196.

Rosenfled, L., and Danishefsky, I. (1986). *Biochem. J.* **237**: 639-646.

Salyers, A. A., Vercellotti, J. R., West, S. E. H. and Wilkins, T. D. (1977). *Appl. Environ. Microbiol.* **33**: 319-22.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, 2 nd ED.,. Cold Spring Harbor, NY., Cold Spring Harbor Laboratory Press,.

San Antonio, J. D., Slover, J., Lawler, J., Karnovsky, M. J. and Lander, A. D. (1993). *Biochemistry*. **32**: 4746-55.

Sarma, R., and Bott, R. (1977). *J. Mol. Biol.* **113**: 555.

Sasisekharan, R. (1991). Cloning and biochemical characterization of heparinase from *Flavobacterium heparinum*. Ph.D thesis,

Sasisekharan, R., Bulmer, M., Moremen, K. W., Cooney, C. L. and Langer, R. (1993). *Proc. Natl. Acad. Sci., USA*. **90**: 3660-64.

Sasisekharan, R., Leckband, D., Godavarti, R., Venkataraman, G., Cooney, C. L. and Langer, R. (1995). *Biochemistry*. **34**: 14441-14448.

Sasisekharan, R., Moses, M. A., Nugent, M. A., Cooney, C. L. and Langer, R. (1994). *Proc. Natl. Acad. Sci., USA*. **91**: 1524-28.

Sasisekharan, R., Venkataraman, G., Godavarti, R., Ernst, S. E., Cooney, C. L. and Langer, R. (1996). *J. Biol Chem.,*. **271**: 3124-3131.

Sawaguchi, S., Yue, B., Yeh, P. and Tso, M. O. M. (1992). *Arch. Ophthalmol.* **110**: 110-17.

Scott, J. E. (1992). *FASEB J.* **6**: 2639-45.

Scott, J. E., Chen, Y. and Brass, A. (1992). *Eur. J. Biochem.* **209**: 675-80.

Scouten, W. H. (1987). *Methods. Enzymol.* **135**: 30-65.

Shihabi, D. S., Garwood, W. E., Chu, P., Miale, J. N., Lago, R. M., Chu, C. T.- W., and Chang, C. D. (1985). *J. Catal.* **93**: 471.

Shing, Y., Folkman, J., Sullivan, R., Butterfield, C., Murry, J., and Klagsburn, M. (1984). *Science*. **223**: 1296.

Silva, M. E., Dietrich, C. P. and Nader, H. B. (1976). *Biochim. Biophys. Acta.* **437**: 129-41.

Silverberg, I., Havsmark, B. and Fransson, L.-Å. (1985). *Carb. Res.* **137**: 227-38.

Skjak-Braek, G., and Larsen, B. (1985) *Carbohydr. Res.* **139**: 273-83.

Smith, M. C., Furman, T. C., Ingolia, T. D., and Pidgeon, C. (1988). *J. Biol. Chem.* **263**: 7211-7215.

Smith, J. W. and Knauer, D. J. (1987). . *J. Biol. Chem.* **262**: 11964-72.

Somorjai, G. A. (1987). Catalyst Design Progress and Perspectives. John Wiley and Sons.

Stayton, P. S., Olinger, J. M., Jiang, M.,Bohn, P. W., and Sligar, S. G. (1992). *J. Am. Chem. Soc.* **114**: 9298-9299.

Steyn, P. L., Pot, B., Segers, P., Kersters, K. and Joubert, J. J. (1992). *System. Appl. Microbiol.* **15**: 137-43.

Stuckey, J., Schubert., H., Fauman, E., Zhang, Z., Dixon, J., and Saper, M (1994) *Nature* **370**, 571-75

Sun, X., and Chang, J. (1990). *Biochemistry*. **29**: 8957.

Takegawa, K., Iwahara, K. and Iwahara, S. (1991). *J. Ferm. Bioeng.* **72**: 128-31.

Tam, Y.-C. and Chan, E. C. S. (1985). *Infect. Imm.* **47**: 508-13.

Tejidor, L., Oman, D., Zimmermann, J., Russo, A., Rose, M., Borzhemskaya, L. and Pelzer, H. (1993). *Thromb. Hemostasis*. **69**: 866.

Thomas, J. M. and. Klinowski., J. (1985). *Adv. Catal.* **33**: 200.

Thomas, P. G., Russell, A. J., and Fersht, A. R. (1985). *Nature*. **318**: 375-376.

Thompson, L. D., Pantoliano, M. W. and Springer, B. A. (1994). *Biochemistry*. **33**: 3831-40.

Torchinsky, Y. M. (1981). Sulfur in Proteins. NY, Permagon Press. Inc.

Umana, P., Kellis, J. T., Jr., and Arnold, F. H. (1993). *ACS. Symp. Ser.* **516**: 102-108.

Varki, A. (1993). *Glycobiology*. **3**: 97-130.

Venkataraman, G. (1992). Proteins at charged interfaces. Ph.D thesis,

Venkataraman, G., Sasisekharan, V., Herr, A., Ornitz, D., Waksman, G., Cooney, C. L. and Langer, R., Sasisekharan, R. (1996). *Proc. Natl. Acad. Sci., USA*. **93**: 845-850.

Von Heijne, G. (1988). *Biochim. Biophys. Acta.* **947**: 307-333.

Warshel, A., Naray-Sczabo, G., Sussman, F., and Hwang, J., -K. (1989). *Biochemistry*. **28**: 3631-3637.

Winkler, F. K., D' Arcy, A., and Hunziker, W. (1990). *Nature*. **343**: 771-774.

Wiseman, A. (1993). *J. Chem. Tech. Biotechnol.* **56**: 3-13.

Wu, H., Monroe, D., and Church, F. (1995). *Arch. Biochem. Biophy.* **317**: 85-92.

Yamada, S., Yoshida, K., Sugiura, M., Sugahara, K., Khoo, K. H., Morris, H. R. and Dell, A. (1993). *J Biol Chem*. **268**: 4780-87.

Yang, V. C., Linhardt, R. J., Bernstein, H., Cooney, C. L. and Langer, R. (1985). *J. Biol. Chem*. **260**: 1849-57.

Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P. and Ornitz, D. M. (1991). *Cell*. **64**: 841-48.

Yoder, M. D., Keen, N. T. and Jurnak, F. (1993a). *Science*. **260**: 1503-07.

Yoder, M. D., Lietzke, S. E. and Jurnak, F. (1993b). *Structure*. **1**: 241-51.

Yokoyama, Y., Tanioka, A., and Miyasaka, K. (1988). *J. Membr. Sci.* **38**: 223-236.

Yost, H. J. (1992). *Nature*. **357**: 158-61.

Zhang, Z., Wang, Y., Wu, L., Fauman, E., Stuckey, J., Schubert, H., Saper, M., and Dixon, J. (1994). *Biochemistry*. **33**: 15266-15270.

Zhang, Z., and Dixon, J. (1993). *Biochemistry*. **32**: 9340-9345.

Zhuang, P. and. Butterfield., D. A. (1992). *Biotechnol. Prog.* **8**: 204-210.

Zimmermann, J. J. (1984). *Ph.D Thesis, MIT, Cambridge, MA.*