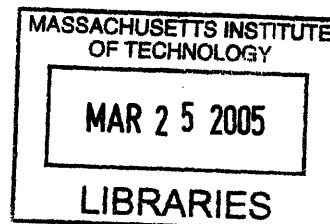# Nitric Oxide-Induced DNA Recombination
&
Glycosaminoglycan Mediated Differentiation in Stem Cells

by

Tanyel Kiziltepe

B.S., Bilkent University (1998)

Submitted to the Department of Chemistry
In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Biochemistry

at the

Massachusetts Institute of Technology

February, 2005

Signature of Author ....
Department of Chemistry
January 14, 2005

Certified by...
Bevin P. Engelward
Associate Professor of Biological Engineering
Thesis Co-Supervisor

Certified by...
Ram Sasisekharan
Professor of Biological Engineering
Thesis Co-Supervisor

Accepted by...
Robert W. Field
Haslam and Dewey Professor of Chemistry
Chairman, Department Committee on Graduate Students

This doctoral thesis has been examined by a committee of the Department of Chemistry as follows:

Professor John M. Essigmann..
Chairman

Professor Steven R. Tannenbaum...

Professor Gerald N. Wogan...

To my loving family

Preface

The work described in this dissertation is the result of collaborations with several members of the Sasisekharan, Engelward and Dedon Laboratories. I am extremely grateful to all those who have contributed to this thesis project.

Chapter 2 of this dissertation was produced under the supervision of Professor Bevin Engelward and is currently in press (*Chemistry & Biology*). The plasmids used were created by Erik Spek and Tet Matsuguchi. The ΔGF cell line was engineered and characterized by Vidya Jonnalagadda. The quantification of the DNA lesions was performed by Min Dong from the Dedon Laboratory. Amy Yan, who is a UROP in the Engelward Laboratory, worked on this project from the beginning to the end under my supervision. Laura Trudel from the Wogan Laboratory helped with the optimization of the nitric oxide exposure conditions for attached mammalian cells. I am grateful to all these individuals for their invaluable contributions.

Chapter 5 of this dissertation was produced under the supervision of Professor Ram Sasisekharan and is published in *Biochemistry* (June, 2002). The native enzyme was purified by Zachary Shriver. The molecular cloning and recombinant expression of this enzyme was performed in collaboration with James Myette. I am grateful to these individuals for their invaluable contributions.

Chapter 6 of this dissertation was produced under the supervisions of Professor Ram Sasisekharan. The confocal microscopy experiments were performed by Shiladitya Sengupta. The real time PCR experiments were performed in collaboration with David Eavarone. Aarthi Chandrasekaran did the enzymatic treatments of cells. Kris Holey assisted all the team

members with various aspects of this project. I am grateful to all these individuals for their

invaluable contributions.

Acknowledgements

Scientific research is a challenging pursuit and graduate students are at the bottom of the food chain. I have been lucky enough to have two great advisors, Professor Bevin Engelward and Professor Ram Sasisekharan, to guide me through this period of my life. It is with the greatest gratitude that I thank to my advisors for providing me with every opportunity, as well as challenge, for scientific and personal growth. I owe a great deal to Bevin for generously sparing her time for me, and for putting a special effort to train me to grow into a better scientist. I am grateful to Ram for giving me the freedom to grow into the scientist I became. His kindness, encouragement and support made a tremendous difference at very hard times. I am also grateful to my first advisor Professor Lawrence Stern, who was the first one to teach me how to critically think, design experiments and interpret data. Although, the research I pursued in his group is not a part of this thesis, I definitely transferred those skills to this thesis work. I have been very much inspired by all my advisors and I will carry what I learnt from all of them to my future life.

It is with the greatest respect, admiration and gratitude that I thank my thesis committee members Professor Steven Tannenbaum, Professor Gerald Wogan and Professor John Essigmann. I owe a great deal to these three individuals, more than I can express, for their unwavering support, encouragement, and wisdom. I have been the luckiest graduate student, because I had the best thesis committee a student could possibly have. In addition, I would like to thank Professor Peter Dedon for his genuine support and guidance. This journey would not have been possible without these individuals.

I am also grateful to the excellent colleagues and friends I met during my doctoral studies. I am thankful to all members of the Engelward and Sasisekharan Laboratories. I

would like to especially thank to Jack Manes, Shiladitya Sengupta, Vidya Jonnalagadda, David Eavarone, Min Dong, Laura Trudel, Chunqi Li and Susan Brighton, who made a big difference in my graduate student life. I am truly indebted to these friends for their sincere support and help.

Finally, I would like to extend the most special thanks to my family and my fiancé. This whole journey would not have been possible without their sacrifices, love and support. Words would not be enough to explain …

Nitric Oxide-Induced DNA Recombination
&
Glycosaminoglycan Mediated Differentiation in Stem Cells

by

Tanyel Kiziltepe

Submitted to the Department of Chemistry on January 14, 2005 in Partial Fulfillment of the Requirements for Degree of Doctor of Philosophy in Biochemistry

ABSTRACT

The new paradigm is that cancers may originate from stem cells, and that terminal differentiation of stem cells is a possible treatment for cancers. Understanding the origin of cancers requires elucidation of factors causing genetic rearrangements. The first part of this thesis explores the effects of $NO^{\bullet}$ on homologous recombination in embryonic stem cells. Using terminal differentiation of stem cells as a therapy for cancer necessitates an understanding of factors governing their differentiation. The second part of this thesis explores the effects of heparan sulfate glycosaminoglycans (HSGAGs) on stem cell differentiation.

Inflammation is increasingly recognized as an important risk factor for cancer. During inflammation, macrophages secrete $NO^{\bullet}$, which reacts with superoxide or oxygen to produce $ONOO^{-}$ or N_2O_3, respectively. Although $ONOO^{-}$ and N_2O_3 are potent DNA damaging agents, little was known about the ability of these agents to induce homologous recombination in mammalian cells. Homologous recombination events are a significant source of mutations that are likely to contribute to initiation and progression of some cancers. In the first part of this thesis, the recombinogenic potential of $ONOO^{-}$ and N_2O_3 was characterized by sister chromatid exchanges, chromosomal direct repeat substrate and interplasmid recombination assays. Our results show that on a per lesion basis, $ONOO^{-}$-induced oxidative base lesions and single strand breaks are more recombinogenic than N_2O_3-induced base deamination products. These results are in accordance with the model that $ONOO^{-}$-induced recombination may contribute to inflammation-induced cancer.

Directed differentiation of stem cells holds an immense potential for regenerative medicine as well as cancer therapy. Such therapeutic approaches require an understanding of the mechanisms regulating stem cell differentiation. The second part of this thesis investigates the role of HSGAGs in embryonic stem cell differentiation into endothelial cells. Differentiation of stem cells was accompanied by increases in the transcript levels of key HSGAG-biosynthetic enzymes, and the quantity of cell surface HSGAGs. Differentiation into endothelial cells was inhibited by ablation of the HSGAG-biosynthetic machinery by chlorate treatment, or by the enzymatic degradation of the HSGAGs. Exogenous addition of heparin to chlorate-treated cells partially restored differentiation into endothelial cells. These

8

effects were mirrored in phospho-ERK levels, suggesting the involvement of the MAPK pathway. These results suggest that stem cell differentiation can be regulated by modulating the HSGAG moiety and this opens up new treatment modalities for cancer therapy and regenerative medicine.

Thesis Co-Supervisor: Bevin P. Engelward
Title: Associate Professor of Biological Engineering

Thesis Co-Supervisor: Ram Sasisekharan
Title: Professor of Biological Engineering

Table of Contents

List of Figures

List of Tables

PART A

Effects of Nitric Oxide on Homologous Recombination in Mammalian Cells

CHAPTER 1

Introduction

CHAPTER 1

Introduction

1.1 Inflammation and Cancer

More than a century ago, in 1863, Rudolf Virchow for the first time made a connection between inflammation and cancer by hypothesizing that the origin of cancer was at sites of inflammation (*1*). Today, it is widely believed that increased risk of malignancy is associated with chronic inflammation (*2, 3*). It is estimated that ~15% of cancer cases worldwide are attributable to infectious diseases, many of which induce chronic inflammation (*4, 5*). Some examples of strong epidemiological associations of chronic inflammation and cancer include the association between colon carcinogenesis and inflammatory bowel disease, the association between liver carcinoma and Hepatitis C infection, and the association between bladder and colon carcinoma and schistosomiasis (*6*). Moreover, *Helicobacter pylori* infection is demonstrated to be a carcinogen of gastric cancer and it is known to be the leading cause of stomach cancer. More examples of epidemiological associations of chronic inflammation and cancer are presented in Table 1.1.

In inflamed tissues, a variety of immune system cells, including basophils, eosinophils, lymphocytes, neutrophils and macrophages are recruited and are activated to produce potent cytotoxic agents, primarily to destroy the invading pathogens and foreign bodies and tumor cells. These cytotoxic agents include various reactive oxygen and nitrogen species such as superoxide, hydrogen peroxide, nitric oxide (NO$^{\bullet}$), and the secondary products that result from the reaction of NO$^{\bullet}$ with oxygen and superoxide (discussed in detail below). Under some circumstances, inflammation lasts for prolonged periods of time,

17

extending to months and sometimes to years, during which time not only the invading pathogens and foreign bodies, but also the host cells are exposed to high levels of cytotoxic reactive oxygen and nitrogen species. These reactive oxygen and nitrogen species have been shown to induce DNA lesions and mutations under various experimental conditions (*e.g.*, see references (*7-13*). Although the underlying mechanism of inflammation-induced cancer is not yet fully understood, it has been proposed that the DNA damage and mutations produced by these reactive oxygen and nitrogen species are one of the main causes of initiation and progression of cancer (*14-17*). One example is gastric cancer caused by *Helicobacter pylori*, where DNA damage resulting from chronic inflammation is believed to be a major mechanism of induction and progression of disease (*18*). Much work has been focused on understanding $NO^{\bullet}$ induced DNA damage and mutations.

1.2 Nitric Oxide

$NO^{\bullet}$ is a free radical gas with multiple biological functions. With its electrically neutral charge and small size, it can freely diffuse through cell membranes and act as a signaling molecule in diverse biological processes. On the other hand, it is a radical with an unpaired electron, and it readily reacts with various molecules to form products that are toxic to a cell (for reviews see (*19-22*)). For example, $NO^{\bullet}$ can react with oxygen or superoxide to form nitrous anhydride (N_2O_3) or peroxynitrite ($ONOO^{-}$), respectively, which have been shown to be genotoxic agents (for reviews see references (*23,24*)). The fate of $NO^{\bullet}$ and its biological effects are determined by many factors, which include (1) the physicochemical factors such as the rate of $NO^{\bullet}$ production, the rate of $NO^{\bullet}$ diffusion and the distances between $NO^{\bullet}$ generator cells and the target cells; and (2) the chemical and biochemical

factors, such as the concentration of oxygen and superoxide, the concentration of catalase and superoxide dismutase, and the levels of antioxidants such as glutathione (*25, 26*). Depending on the combination of all these factors, NO$^\bullet$ acts either as a signaling molecule, facilitating inter-cellular and intra-cellular signaling events in various biological processes, or as a cytotoxic agent, used by the immune cells in host defense.

NO$^\bullet$ is synthesized from L-arginine by a family of isoenzymes called NO$^\bullet$ synthases (*27*). In mammals, there are three known types of nitric oxide synthases, nNOS (neuronal NOS), eNOS (endothelial NOS), and iNOS (inducible NOS). eNOS and nNOS are constitutively expressed enzymes, and the low concentrations of NO$^\bullet$ (~nM) released by these enzymes are known to play important roles in cellular signaling in cardiovascular and nervous systems, including processes such as vasodilation and neurotransmission and blood pressure regulation. In contrast to eNOS and nNOS, iNOS is an inducible enzyme. Under immunological stimuli, iNOS is induced, and NO$^\bullet$ is released for potentially long periods of time, resulting in high concentration of NO$^\bullet$ (steady state levels of ~1 μM) in the surrounding tissue (*28-30*). At these high concentrations, NO$^\bullet$ has been shown to be a cytotoxic agent.

iNOS was initially discovered in macrophages, which are among the cells of mammalian immune system. Consistently, NO$^\bullet$ plays a particularly important role in host defense. During an immune or inflammatory response, iNOS is induced by interferon-γ (*31*), leading to continuous production of NO$^\bullet$ by macrophages for days or weeks (*32-35*). In addition to NO$^\bullet$, macrophages also secrete various other species, such as reactive oxygen species (*i.e.*, superoxide and hydrogen peroxide), FAS ligand, tumor necrosis factor, and various cytokines and chemokines (*e.g.*, interleukins and interferon-γ) (*6, 17, 23, 26*). Among these species, NO$^\bullet$ is believed to be a key mediator of macrophage-induced cytotoxicity,

19

based on the observation that NO$^{\bullet}$ scavengers inhibit the cytotoxic effects of macrophages (*33, 36*). In some cases, inflammation continues for prolonged periods of time, extending to months or years. Under these circumstances, the host tissue inevitably gets exposed to cytotoxic levels of NO$^{\bullet}$ and reactive oxygen species. Under these circumstances, significant levels of tissue damage are generated by high levels of NO$^{\bullet}$ and reactive oxygen species. Interestingly, formation of high levels of NO$^{\bullet}$ in the presence of reactive oxygen species has been associated with a variety of human pathologies, including atherosclerosis, rheumatoid arthritis, neurodegenerative diseases and cancer (*37*).

1.3 Nitric Oxide induced DNA Damage and Mutagenesis

NO$^{\bullet}$ chemistry in biological systems is extremely complex due to the large number of the chemical species formed. Although NO$^{\bullet}$ itself is a reactive radical gas and it may be involved with Fenton chemistry and hydroxyl radical induced DNA damage (*24*), the genotoxicity of NO$^{\bullet}$ predominantly arises from its derivatives, nitrous anhydride (N_2O_3) and peroxynitrite ($ONOO^-$), which are the dominant reactive nitrogen species formed under physiological conditions (*23, 26*). The predominant types of DNA damages and point mutations produced by N_2O_3 and $ONOO^-$ (and $ONOOCO_2^-$) are summarized in Figure 1.1 and are explained in more detail below.

1.3.1 Nitrous Anhydride (N_2O_3)

N_2O_3 is formed from the oxidation of NO$^{\bullet}$ with molecular oxygen at physiological pH (*38*). The rate of formation of N_2O_3 has been shown to be second order in NO$^{\bullet}$ concentration and first order in oxygen, with a rate constant of 8.4×10^6 M^{-2} s^{-1} at 37°C (*28*). Thus, the half life of NO$^{\bullet}$ is inversely proportional to its concentration, such that as the NO$^{\bullet}$ concentration

increases, the rate of its turnover to N_2O_3 increases proportionally in a second order manner. Hence, under conditions such as chronic inflammation, where elevated level of $NO^•$ are secreted by the immune system cells for prolonged periods of time, $NO^•$ rapidly reacts with oxygen to form N_2O_3.

N_2O_3 is a potent DNA deaminating agent and a minor DNA crosslinking agent. At physiological pH, it is involved in conversion of cytosine to uracil, guanine to xanthine, adenine to hypoxanthine, 5-methylcytosine to thymine, and there is evidence that N_2O_3 generates low levels of G-G cross-links (*15, 26, 39-43*). Studies conducted under physiological conditions suggest that uracil, guanine to xanthine constitute most of the DNA damage induced by N_2O_3 (25-35% each) in an inflamed tissue. Only 4-6% of the total damage is made up of abasic sites and ~2% is made up of G-G cross-links (*7, 23, 43*).

The deamination of DNA bases with N_2O_3 has potential mutagenic consequences. For example, mispairing of xanthine can cause G:C → A:T transition mutation. Alternatively, xanthine can depurinate and can lead to non-informative mutagenic lesions. For example, an adenine may be inserted opposite to the abasic site, leading to a G:C → T:A transversion mutation. In addition, mispairing of uracil can cause a G:C → A:T transition mutation, and mispairing of hypoxanthine can cause A:T → G:C transition mutation. All these mutations have been detected in various mammalian cells exposed to conditions in which N_2O_3 was generated (*44-47*).

1.3.2 Peroxynitrite (ONOO⁻)

During inflammation, activated macrophages generate superoxide in addition to $NO^•$. These two radicals react rapidly with one another to generate $ONOO^-$. The rate constant of this reaction is known to be between $6.6\text{-}19 \times 10^9 \ M^{-1} \ s^{-1}$ at 37°C, which is in the range of

diffusion-controlled limit (*48-50*). In the presence of bicarbonate, such as in physiological systems, ONOO⁻ further reacts with bicarbonate, in a bimolecular reaction with a rate constant of $3-6 \times 10^4$ M^{-1} s^{-1}, to form nitrosoperoxycarbonate ($ONOOCO_2^-$) (*51-53*). Both ONOO⁻ and $ONOOCO_2^-$ are capable of conducting 1- and 2-electron chemistry, including oxidations and nitrations (*51, 52*).

ONOO⁻ is a potent DNA oxidizing agent, which is primarily involved in oxidation of guanine and oxidative breakdown of deoxyribose. The reaction of ONOO⁻ with guanine produces a variety of primary and secondary base lesions. The primarily lesions created include 8-oxoguanine (8-oxoG) and 8-nitroguanine (8-nitroG) (*15, 26, 54, 55*). Formation of oxazolone (*56*), 5-guanidino-4-nitroimidazole (*57*), and 4,5-dihydro-5-hydroxy-4-(nitrooxy)-2'-deoxyguanosine (*58*) has also been reported. It is has been shown that ONOO⁻ is at least 1000-fold more reactive with 8-oxoG than guanine itself (*61*), and further reaction of 8-oxoG with ONOO⁻ leads to the formation of a variety of secondary products, including spiroiminodihydantoin, guanidinohydantoin, 3a-hydroxy-5-imino-3,3a,4,5-tetrahydro-1H-imidazo[4,5-d]imidazol-2-one, 5-iminoimidazolidine-2,4-dione, 2,4,6-trioxo-[1,3,5]-triazinane-1-carboxamidine, parabinic acid, oxaluric acid cyanuric acid, and others (for reviews see (*23, 26*)). In addition to these primary and secondary oxidative base lesions, ONOO⁻ also induces direct single strand breaks in DNA via the oxidative breakdown of deoxyribose (*54, 59, 60*).

It has been shown that the proportions of base lesions to strand breaks induced by ONOO⁻ depends on the presence of bicarbonate (*60, 62*). For example, in the absence of bicarbonate, ONOO⁻ predominantly causes strand breaks. However in the presence of bicarbonate, ONOO⁻ is converted into $ONOOCO_2^-$, and $ONOOCO_2^-$ predominantly causes

base damage with a significant increase in 8-nitroG. Although, bicarbonate shifts the proportion of strand breaks to base damage, the total quantity of lesions remains constant.

ONOO⁻ has been shown to be mutagenic both in bacterial and mammalian systems (*9, 63*). In previous studies, most of the mutations have been shown to occur at the G:C base pairs, leading predominantly to G:C → T:A transversion mutations. However, G:C → C:G transversion mutations and G:C → A:T transition mutations have also been also detected. In addition to these point mutations, ONOO⁻ has also been shown to cause deletion and insertion type mutations (*63*).

In short, both N_2O_3 and ONOO⁻ have been shown to be mutagenic in mammalian cells, and the DNA damage and the resulting mutagenecity of reactive nitrogen species generated during inflammation has been proposed to be a mechanism of induction and progression of cancer. While NO·-induced point mutations have been studied extensively, as described above, very little is known about the ability of NO· to induce other classes of mutations, such as sequence rearrangements that are mediated by homologous recombination.

1.4 Homologous Recombination

Mitotic homologous recombination (HR) is one of the most important mechanisms that cells use to defend against DNA damage-induced toxicity. During HR, missing sequence information is extracted from a sister chromatid or from a homologous chromosome (*64-67*). HR is used to repair double-strand breaks (DSB), especially during the S and G2 phases of cell cycle (*68-70*). Importantly, it has been shown that in mammals, proteins that are key for homologous recombination are essential for life itself (*71-73*).

1.4.1. Classes and Sources of Lesions that Induce Homologous Recombination

Homologous recombination is best characterized for repair of DSB. A DSB in DNA may be formed via replication independent or dependent mechanisms.

Replication independent DSBs are two-ended (Figure 1.1A), such as those formed by ionizing radiation. Alternatively, replication-independent DSBs can be formed from proximity of two single strand breaks on the opposing strands of DNA. For example, single strand breaks are created enzymatically, as repair intermediates, during base excision repair (BER). Consequently, if DNA glycosylases initiate BER of closely opposed lesions, recombinogenic DSBs can be formed (*74, 75*) (Figure 1.1B).

Replication-dependent DSBs are one-ended. One of the mechanisms by which one-ended DSBs are formed is via collapse of a replication fork at a nick in the template DNA (Figure 1.1C). Nicks in the template DNA can be formed directly by strand break-inducing agents or as intermediates of DNA repair pathways. For example, during base excision repair (BER), single strand breaks may be created as repair intermediates, and when encountered by a replication fork, they may form highly recombinogenic one-ended DSBs. Such BER intermediates have been shown to be highly recombinogenic in *S.cerevisiae* (*76-78*) (Figure 1.1D). One-ended DSBs may also be formed when a replication fork encounters a blocking DNA lesion in the leading strand (Figure 1.1E). It has been shown in *E.coli* that inhibition of replication fork progression induces one-ended DSBs that are repaired by homologous recombination (*65, 79*). Alternatively, a blocking lesion could be formed in the lagging strand. It has been proposed that a blocking lesion in the lagging strand would result in the formation of a single strand region called a daughter strand gap (Figure 1.1F). Although not

completely understood, daughter strand gaps are thought to be repaired by homologous recombination.

1.4.2 Mechanisms of Homologous Recombination

Much of what we know about the mechanisms of recombination in eukaryotes comes from studies of *S. cerevisiae* (for extensive reviews see (*80-82*)). Mechanisms of homologous recombination are thought to be conserved from yeast to mammals (*69, 83*). Repair of DSBs is thought to happen through three classes of recombination events: single strand annealing (SSA), gene conversion (GC), and break-induced replication (BIR).

When a two-ended DSB is flanked by homologous sequences, it can be repaired via SSA. For SSA, initially the ends of the double strand breaks are resected 5'-to-3'. Following resection, the homologous 3' overhangs anneal and the gaps are filled (Figure 1.2A). This way, the DSB is repaired, albeit with the loss of significant genetic information, including one of the repeats (*80*).

Most homologous recombination events are GC events where information is transferred from one molecule to another in a non-reciprocal fashion without loss of sequence information (reviewed in (*68*)). In the prototypic recombinational repair model, as originally proposed by Szostak (*84*), initially the ends of the DBS are resected 5'-to-3' forming 3' overhangs. The 3' overhang then invades a homologous sequence (*e.g.*, from a sister chromatid or a homologous chromosome), and are extended by the initiation of new DNA synthesis (Figure 1.2B). This leads to the formation of two Holliday junctions, which can then be resolved in two different ways. Depending on which strands are cleaved, a GC event occurs either with crossing over or without crossing over of the flanking sequences.

25

This model predicts an equal number of crossing over and non-crossing over events. However, it is now known that most GC events occurs without crossing over events, which can not be explained with the traditional Szostak model (85). More recently, synthesis dependent strand annealing (SDSA), has been proposed to explain GC events that are not associated with crossing over of flanking sequences (80). SDSA is initiated similarly to the prototypic Szostak model. The main difference is that, in SDSA only one end of the resected double strand break end is required to invade the homologous DNA to form the D loop (Figure 1.2C). The second end does not necessarily bind to the D loop. Instead, once the missing sequence is restored, the Holliday junction is moved to release the first end, which then anneals to the 3' overhang of the non-invading double strand break end via a process similar to SSA. In this model there is no cleavage of the Holliday junction, thus crossing over events are prevented.

Both SSA and GC mechanisms are used to repair two-ended double strand breaks. A third pathway of homologous recombination exists, which is used to repair one-ended DSBs that are formed from broken replication forks. This mechanism is known as break-induced replication (BIR) (66). As is the case for SSA and GC mechanisms, in BIR, the initial step is again the 5'-to- 3' recession of the DSB. The 3' overhang then invades the homologous DNA, forming a Holliday junction (Figure 1.2D). This leading strand is then extended by synthesis of new DNA using the homologous sequence as a template. This is followed by the resolution of the Holliday junction to complete the process of restoring the broken replication fork.

Although homologous recombination takes place via all three pathways described above, it is noteworthy that in mammals, most two-ended DSBs are repaired by non-

homologous end joining (NHEJ). In NHEJ, the two broken ends are processed and rejoined to each other with some loss of sequence information (*86, 87*). Unlike two-ended DSBs, one-ended DSBs can only be repaired via the BIR pathway. Thus, homologous recombination is a particularly important and indispensable repair pathway for mammals, especially in the context of DSBs induced during replication.

1.4.3 Proteins involved in Homologous Recombination

Most of the proteins involved in homologous recombination were first identified by their requirement for the repair of ionizing-radiation-induced DNA damage in *Saccharomyces cerevisiae*. In yeast these proteins belong to the RAD52 group genes, which include, RAD50, RAD51, RAD52, RAD54, RDH54, RAD55, RAD57, RAD59, MRE11, and XRS2 (*88*). This repair pathway is thought to be evolutionarily conserved in higher organisms, as indicated by the fact that homologs of these genes can be found in mammals (for reviews see references (*67, 83*)). Mammalian homologues of essentially all these genes have been described (*89*), and these include Rad51, Rad52, Rad54, Rad50, Nbs1(Xrs2) and Mre11, as well as the five Rad51 paralogs, Xrcc2, Xrcc3, Rad51B/Rad51L1, Rad51C/Rad51L2, and Rad51D/Rad51L3 (*88, 90*).

The role of the Rad52 epistasis group proteins in homologous recombination has been reviewed in great detail (*68, 69, 88, 91, 92*) and the current understanding that has emerged is briefly summarized here. In all of the mechanisms of homologous recombination described above, repair of the double strand breaks is initiated with the re-dissection of the double strand break to generate 3' overhangs. In mammals, this process is initiated by the Mre11/ Rad50/ Nbs1 complex, and an as yet unidentified exonuclease ressects the break to generate

a single strand 3' overhang (*93, 94*). The single stranded DNA is then bound with Rad51 which forms proteo-filaments around the single-stranded DNA. The binding of Rad51 is stimulated by the initial binding of Replication Protein A (RPA) and Rad52 to the single stranded DNA (*95-98*). RPA is thought to stabilize the single stranded DNA, while Rad52 helps load Rad51 (*99*). Rad51 catalyzes homology searching, strand invasion and strand exchange (*100*). Strand exchange is enhanced by Rad54, which is a member of the Swi2/Snf2 family of chromatin-remodeling protein and is thought to move nucleosomes to assist in homology searching and strand displacement to form a D-loop (*101*). Although the Holliday junctions are thought to be cleaved by resolvases, the proteins involved in branch migration and resolution have yet not been identified in eukaryotes. However, a role for Mus81 endonuclease complex in cleaving Holliday junctions was proposed (*65*), and Rad51C was recently shown to play a role in junction resolution (*102*).

Studies of mouse models with defective recombination genes and studies of mammalian cell lines also support the significant role of the above mentioned proteins in mammals. For example, Rad51 appears to be one of the most critical protein in mammals, based on the observation that *Rad51$^{-/-}$* mutant mice die early during embryonic development (*71*). Moreover, Rad51 deficient cells accumulate chromosomal breaks when cultured and do not survive more than a few cell divisions (*103*). Similar to Rad51, Rad50 and Mre11 genes also seem to be critical for mammals, since *Rad50$^{-/-}$* is embryonic lethal in mice and *Rad50$^{-/-}$* *Mre11$^{-/-}$* mouse cells do not survive in culture (*69*). In *Rad54$^{-/-}$* mouse embryonic stem cells, a 30% decrease in sister chromatid recombination was observed showing that Rad 54 plays a significant role in recombination (*104*). Finally, it has been shown that cell lines that contain defects in the Rad51 paralogs, Xrcc2 and Xrcc3 have significantly lower levels of

recombination and a high level of genomic instability, showing the involvement of these proteins in homologous recombination and genomic stability (*105-108*).

1.5 Homologous Recombination and Cancer

Although homologous recombination is generally accurate and important for genomic stability, transfer of genetic material carries with it a certain amount of risk. Approximately one-third of the mammalian genome is composed of highly repeated DNA sequences (for review see (*109, 110*)), and recombination between misaligned sequences can lead to genetic rearrangements such as insertions, deletions, inversions and translocations (Figure 1.3A). In addition, exchanges between homologous chromosomes are responsible for causing most spontaneous loss of heterozygosity events in mammals (*111-115*) (Figure 1.3B). These genetic rearrangements may contribute to the development of diseases and put cells at risk of developing diseases such as cancer.

Several associations with cancer and homologous recombination have been made. Nearly all tumors show LOH and other types of sequence rearrangements that are likely to result from mitotic homologous recombination. For example, cancers including chronic myelogeneous leukemia, Ewing's sarcoma and breast cancer have been associated with mitotic recombination events involving misaligned *Alu* sequences (*116, 117*). In addition, it is known that one of the causes of acute myeloid leukemia is the partial duplication of the *ALL-1* gene that results from a misaligned recombination event between repetitive *Alu* sequences (*118*). Moreover, it has been shown, in yeast, human cells and mice that many known carcinogens are also recombinogens (for a review see references (*119, 120*)).

Consistent with the association between tumor cells and recombination events, patients whose cells have an increased frequency of recombination have an elevated frequency of cancer (reviewed in ref (*121, 122*)). For example, people with Bloom's syndrome are defective in the BLM helicase. A defect in BLM helicase results in elevated levels of sister chromatid exchanges, hyper recombination and chromosomal aberrations (*123*). These patients are prone to cancer and it has been indicated that one in nine Bloom's patients develop malignancies by the age of 20 (*124*). Another example can be provided by ataxia-telangectasia (AT) syndrome which is caused by a mutation in *ATM*. Patients of AT show a predisposition to cancer. Cells from these patients show an increased susceptibility to homologous recombination, which may contribute to their predisposition to cancer (*125, 126*). Lastly, Li-Fraumeni syndrome patients that carry a recessive mutation in TP53 have been shown to have an early onset of cancer such as carcinomas of breast, sarcomas, brain tumors, leukemia and lymphoma (*127*). Although the relationship between p53 and homologous recombination is yet not clearly understood, many studies show that cells lacking p53 have higher frequency of homologous recombination (*128-130*).

In summary, all the above examples suggest that, whether by exposure to endogenous and environmental recombinogens, or by inherited predisposition, conditions that lead to increased levels of homologous recombination are associated with an increased risk of cancer.

1.6 NO· and Homologous Recombination

At the time when this work was begun, there were almost no studies exploring the recombinogenicity of NO· in mammals. However, there were a few studies suggesting that

NO$^\bullet$ may induce recombination in mammalian cells. For example, patients who suffer from chronic inflammation associated with Crohn's Disease have increased levels of sister chromatid exchanges (SCEs) in their lymphocytes (*131*), though it is not known to what extent NO$^\bullet$ is responsible for this effect. In addition, two studies have shown that mammalian cells exposed to chemicals that give rise to NO$^\bullet$ ('NO$^\bullet$ donors') suffer increased levels of SCEs (*132, 133*), though the NO$^\bullet$ donors used in these studies also give rise to additional potentially recombinogenic radical species. The observation that mammalian cells exposed to NO$^\bullet$ have an increased susceptibility to LOH provides additional support for the possibility that NO$^\bullet$ induces recombination (*10*). Taken together, these observations suggest that NO$^\bullet$ may induce homologous recombination in mammalian cells.

Although double strand breaks are thought to be critical for inducing homologous recombination, *in vitro* studies using purified DNA have shown that neither N$_2$O$_3$, nor ONOO$^-$ efficiently creates double strand breaks by direct reaction with DNA (*59, 134, 135*). However, as described in the previous sections, base lesions, abasic sites, and single strand breaks can be converted into double strand breaks by enzymatic processing or when they are encountered by the replication fork. For example, during base excision repair (BER), single strand breaks are created as repair intermediates. Consequently, if DNA glycosylases initiate BER of closely opposed lesions, recombinogenic double strand breaks can be formed (*74, 75*). Alternatively, DNA lesions that inhibit replication fork progression, such as BER intermediates, are highly recombinogenic (*e.g.*, references (*76-78, 136*)). Indeed, previous work has demonstrated that NO$^\bullet$ induces recombination in *E. coli*, and DNA glycosylases promote this NO$^\bullet$-induced recombination in *E. coli*, presumably by converting base lesions into recombinogenic BER intermediates (*137, 138*). These results suggest that, as

demonstrated in *E. coli*, N$_2$O$_3$ and ONOO$^-$ created DNA lesions may also be converted into recombinogenic double strand breaks and cause induction of recombination in mammalian cells.

1.7 Specific Objectives

It is well established that there is an association between inflammation and increased risk of cancer, and it is proposed NO$^{\bullet}$ induced DNA damage plays a key role in initiation and progression of cancer. It is also well established that there is an association between elevated levels of homologous recombination and increased risk of cancer, and it has been shown that most known carcinogens are also recombinogens. While NO$^{\bullet}$-induced point mutations have been studied extensively, very little was known about the ability of NO$^{\bullet}$ to induce other classes of mutations, such as sequence rearrangements that are mediated by homologous recombination in mammalian cells. Thus, the aim of this work has been to reveal the relationship of NO$^{\bullet}$ and homologous recombination in mammalian cells. Specifically, tools were developed to study homologous recombination in mammalian cells and these tools were used to dissect the relative recombinogenicity of reactive nitrogen species. Furthermore, we set out to delineate the classes of recombinogenic DNA lesions. Our results show that NO$^{\bullet}$ generated peroxynitrite is a potent inducer of recombination in mammalian cells, which suggest that ONOO$^-$-induced recombinogenic DNA damage may play an important role in inflammation-induced cancers. The next chapter explains the strategy, results and the impacts of our findings in more detail.

1.8 References

1. Balkwill, F., and Mantovani, A. (2001) *Lancet 357*, 539-45.
2. Ekbom, A., Helmick, C., Zack, M., and Adami, H. O. (1990) *N Engl J Med 323*, 1228-33.
3. Gulumian, M. (1999) *Mol Cell Biochem 196*, 69-77.
4. Pisani, P., Parkin, D. M., Munoz, N., and Ferlay, J. (1997) *Cancer Epidemiol. Biomarkers Prev. 6*, 387-400.
5. Kuper, H., Adami, H. O., and Trichopoulos, D. (2000) *J. Intern. Med. 248*, 171-183.
6. Coussens, L. M., and Werb, Z. (2002) *Nature 420*, 860-867.
7. Dong, M., Wang, C., Deen, W. M., and Dedon, P. C. (2003) *Chem. Res. Toxicol. 16*, 1044-1055.
8. Tretyakova, N. Y., Niles, J. C., Burney, S., Wishnok, J. S., and Tannenbaum, S. R. (1999) *Chem. Res. Toxicol. 12*, 459-466.
9. Tretyakova, N. Y., Wishnok, J. S., and Tannenbaum, S. R. (2000) *Chem. Res. Toxicol. 13*, 658-664.
10. Li, C. Q., Trudel, L. J., and Wogan, G. N. (2002) *Proc. Natl. Acad. Sci. USA 99*, 10364-10369.
11. Li, C. Q., Trudel, L. J., and Wogan, G. N. (2002) *Chem. Res. Toxicol. 15*, 527-535.
12. Dizdaroglu, M. (1994) *Methods Enzymol 234*, 3-16.
13. Christen, S., Hagen, T.M., Shigenaga, M.K., Ames, M.K. (1999) *Microbes and Malignancy, Infection as a Cause of Human Cancers*, Oxford University Press.
14. Tamir, S., and Tannenbaum, S. R. (1996) *Biochim. Biophys. Acta. 1288*, F31-6.
15. deRojas-Walker, T., Tamir, S., Ji, H., Wishnok, J. S., and Tannenbaum, S. R. (1995) *Chem Res Toxicol 8*, 473-7.
16. Ohshima, H., and Bartsch, H. (1994) *Mutat. Res. 305*, 253-264.
17. Ohshima, H., Tatemichi, M., and Sawa, T. (2003) *Arch Biochem Biophys 417*, 3-11.
18. Ernst, P. B., and Gold, B. D. (2000) *Annu Rev Microbiol 54*, 615-40.
19. Mungrue, I. N., Bredt, D. S., Stewart, D. J., and Husain, M. (2003) *Acta Physiol Scand 179*, 123-35.
20. Moncada, S., and Higgs, E. A. (1991) *Eur. J. Clin. Invest. 21*, 361-374.
21. Kerwin, J. F., Jr., Lancaster, J. R., Jr., and Feldman, P. L. (1995) *J. Med. Chem. 38*, 4343-462.
22. Wiseman, H., and Halliwell, B. (1996) *Biochem. J. 313*, 17-29.
23. Dedon, P. C., and Tannenbaum, S. R. (2004) *Arch. Biochem. Biophys. 423*, 12-22.
24. Tannenbaum, S. R., Tamir, T., deRojas-Walker, J. S., and Wishnok, P. (1994) in *Nitrosaminess and related N-Nitrososcompounds* (Micheejda, C. J., Ed.) pp 120-135, American Chemical Society, Washington D.C.
25. Tamir, S., Lewis, R. S., de Rojas Walker, T., Deen, W. M., Wishnok, J. S., and Tannenbaum, S. R. (1993) *Chem. Res. Toxicol. 6*, 895-899.
26. Burney, S., Caulfield, J. L., Niles, J. C., Wishnok, J. S., and Tannenbaum, S. R. (1999) *Mutation Res. 424*, 37-49.
27. Knowles, R. G., and Moncada, S. (1994) *Biochem J 298 (Pt 2)*, 249-58.
28. Lewis, R. S., Tamir, S., Tannenbaum, S. R., and Deen, W. M. (1995) *J. Biol. Chem. 270*, 29350-29355.

29. Miwa, M., Stuehr, D. J., Marletta, M. A., Wishnok, J. S., and Tannenbaum, S. R. (1987) *Carcinogenesis 8*, 955-8.
30. Stuehr, D. J., and Marletta, M. A. (1987) *Cancer Res 47*, 5590-4.
31. Nathan, C. (1992) *Faseb J 6*, 3051-64.
32. Moncada, S., Palmer, R. M., and Higgs, E. A. (1991) *Pharmacol Rev 43*, 109-42.
33. Hibbs, J. B., Jr., Taintor, R. R., Vavrin, Z., and Rachlin, E. M. (1988) *Biochem Biophys Res Commun 157*, 87-94.
34. Feldman, P. L., Griffith, O. W., Hong, H., and Stuehr, D. J. (1993) *J Med Chem 36*, 491-6.
35. Marletta, M. A. (1988) *Chem Res Toxicol 1*, 249-57.
36. MacMicking, J., Xie, Q. W., and Nathan, C. (1997) *Annu Rev Immunol 15*, 323-50.
37. Gross, S. S., and Wolin, M. S. (1995) *Annu Rev Physiol 57*, 737-69.
38. Hughes, E. D., Ingold, C. K., and Ridd, J. H. (1958) *J. Chem. Soc.*, 58-98.
39. Caulfield, J. L., Wishnok, J. S., and Tannenbaum, S. R. (1998) *J. Biol. Chem. 273*, 12689-12695.
40. Nguyen, T., Brunson, D., Crespi, C. L., Penman, B. W., Wishnok, J. S., and Tannenbaum, S. R. (1992) *Proc. Natl. Acad. Sci. USA 89*, 3030-3034.
41. Suzuki, T., Kanaori, K., Tajima, K., and Makino, K. (1997) *Nucleic Acids Symp. Ser.*, 313-314.
42. Wink, D. A., Kasprzak, K. S., Maragos, C. M., Elespuru, R. K., Misra, M., Dunams, T. M., Cebula, T. A., Koch, W. H., Andrews, A. W., Allen, J. S., and Keefer, L. K. (1991) *Science 254*, 1001-1003.
43. Caulfield, J. L., Wishnok, J. S., and Tannenbaum, S. R. (2003) *Chem Res Toxicol 16*, 571-4.
44. Zhuang, J. C., Wright, T. L., deRojas-Walker, T., Tannenbaum, S. R., and Wogan, G. N. (2000) *Environ. Mol. Mutagen. 35*, 39-47.
45. Zhuang, J. C., Lin, C., Lin, D., and Wogan, G. N. (1998) *Proc Natl Acad Sci U S A 95*, 8286-91.
46. Routledge, M. N. (2000) *Mutat. Res. 450*, 95-105.
47. Routledge, M. N., Wink, D. A., Keefer, L. K., and Dipple, A. (1993) *Carcinogenesis 14*, 1251-1254.
48. Kissner, R., Nauser, T., Bugnon, P., Lye, P. G., and Koppenol, W. H. (1997) *Chem Res Toxicol 10*, 1285-92.
49. Kissner, R., Nauser, T., Bugnon, P., Lye, P. G., and Koppenol, W. H. (1998) *Chem Res Toxicol 11*, 557.
50. Huie, R. E., and Padmaja, S. (1993) *Free Radic Res Commun 18*, 195-9.
51. Lymar, S. V., and Hurst, J. K. (1995) *J. Am. Chem. Soc. 117*, 8867-8868.
52. Uppu, R. M., Squadrito, G. L., and Pryor, W. A. (1996) *Arch Biochem Biophys 327*, 335-43.
53. Denicola, A., Freeman, B. A., Trujillo, M., and Radi, R. (1996) *Arch Biochem Biophys 333*, 49-58.
54. Kennedy, L. J., Moore, K., Jr., Caulfield, J. L., Tannenbaum, S. R., and Dedon, P. C. (1997) *Chem. Res. Toxicol. 10*, 386-392.
55. Yermilov, V., Rubio, J., and Ohshima, H. (1995) *FEBS Lett. 376*, 207-210.
56. Stamler, J. S., Singel, D. J., and Loscalzo, J. (1992) *Science 258*, 1898-902.

57. Niles, J. C., Wishnok, J. S., and Tannenbaum, S. R. (2001) *J Am Chem Soc 123*, 12147-51.

58. Douki, T., Cadet, J., and Ames, B. N. (1996) *Chem Res Toxicol 9*, 3-7.

59. Salgo, M. G., Stone, K., Squadrito, G. L., Battista, J. R., and Pryor, W. A. (1995) *Biochem. Biophys. Res. Commun. 210*, 1025-1030.

60. Tretyakova, N. Y., Burney, S., Pamir, B., Wishnok, J. S., Dedon, P. C., Wogan, G. N., and Tannenbaum, S. R. (2000) *Mutat. Res. 447*, 287-303.

61. Uppu, R. M., Cueto, R., Squadrito, G. L., Salgo, M. G., and Pryor, W. A. (1996) *Free Radic Biol Med 21*, 407-11.

62. Yermilov, V., Rubio, J., Becchi, M., Friesen, M. D., Pignatelli, B., and Ohshima, H. (1995) *Carcinogenesis 16*, 2045-2050.

63. Juedes, M. J., and Wogan, G. N. (1996) *Mutat. Res. 349*, 51-61.

64. Helleday, T. (2003) *Mutat. Res. 532*, 103-115.

65. McGlynn, P., and Lloyd, R. G. (2002) *Nat. Rev. Mol. Cell. Biol. 3*, 859-870.

66. Kraus, E., Leung, W. Y., and Haber, J. E. (2001) *Proc. Natl. Acad. Sci. USA 98*, 8255-8262.

67. Haber, J. E. (2000) *Trends Genet. 16*, 259-264.

68. Johnson, R. D., and Jasin, M. (2001) *Biochem Soc Trans 29*, 196-201.

69. Karran, P. (2000) *Curr. Opin. Genet. Dev. 10*, 144-150.

70. Kowalczykowski, S. C. (2000) *Trends Biochem. Sci. 25*, 156-165.

71. Tsuzuki, T., Fujii, Y., Sakumi, K., Tominaga, Y., Nakao, K., Sekiguchi, M., Matsushiro, A., Yoshimura, Y., and MoritaT. (1996) *Proc. Natl. Acad. Sci. USA 93*, 6236-6240.

72. Pittman, D. L., and Schimenti, J. C. (2000) *Genesis 26*, 167-173.

73. Lim, D. S., and Hasty, P. (1996) *Mol. Cell. Biol. 16*, 7133-7143.

74. Yang, N., Galick, H., and Wallace, S. S. (2004) *DNA Repair (Amst) 3*, 1323-1334.

75. Blaisdell, J. O., and Wallace, S. S. (2001) *Proc. Natl. Acad. Sci. USA 98*, 7426-7430.

76. Swanson, R. L., Morey, N. J., Doetsch, P. W., and Jinks-Robertson, S. (1999) *Mol. Cell. Biol. 19*, 2929-2935.

77. Memisoglu, A., and Samson, L. (2000) *J. Bacteriol. 182*, 2104-2112.

78. Hendricks, C. A., Razlog, M., Matsuguchi, T., Goyal, A., Brock, A. L., and Engelward, B. P. (2002) *DNA Repair 1*, 645-659.

79. Michel, B., Flores, M. J., Viguera, E., Grompone, G., Seigneur, M., and Bidnenko, V. (2001) *Proc. Natl. Acad. Sci. USA 98*, 8181-8188.

80. Paques, F., and Haber, J. E. (1999) *Microbiol. Mol. Biol. Rev. 63*, 349-404.

81. Haber, J. E. (2000) *Mutat Res 451*, 53-69.

82. Haber, J. E. (2000) *Curr Opin Cell Biol 12*, 286-92.

83. Sonoda, E., Takata, M., Yamashita, Y. M., Morrison, C., and Takeda, S. (2001) *Proc Natl Acad Sci U S A 98*, 8388-94.

84. Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J., and Stahl, F. W. (1983) *Cell 33*, 25-35.

85. Stark, J. M., and Jasin, M. (2003) *Mol Cell Biol 23*, 733-43.

86. Weterings, E., and van Gent, D. C. (2004) *DNA Repair (Amst) 3*, 1425-35.

87. Iliakis, G., Wang, H., Perrault, A. R., Boecker, W., Rosidi, B., Windhofer, F., Wu, W., Guan, J., Terzoudi, G., and Pantelias, G. (2004) *Cytogenet Genome Res 104*, 14-20.

88. Symington, L. S. (2002) *Microbiol Mol Biol Rev 66*, 630-70, table of contents.

89. Wood, R. D., Mitchell, M., Sgouros, J., and Lindahl, T. (2001) *Science 291*, 1284-1289.

90. Thompson, L. H., and Schild, D. (2001) *Mutat. Res. 477*, 131-153.

91. Dudas, A., and Chovanec, M. (2004) *Mutat Res 566*, 131-67.

92. Krogh, B. O., and Symington, L. S. (2004) *Annu Rev Genet.*

93. D'Amours, D., and Jackson, S. P. (2002) *Nat. Rev. Mol. Cell. Biol. 3*, 317-327.

94. Tauchi, H., Kobayashi, J., Morishima, K., van Gent, D. C., Shiraishi, T., Verkaik, N. S., vanHeems, D., Ito, E., Nakamura, A., Sonoda, E., Takata, M., Takeda, S., Matsuura, S., and Komatsu, K. (2002) *Nature 420*, 93-98.

95. Baumann, P., Benson, F. E., and West, S. C. (1996) *Cell 87*, 757-766.

96. Milne, G. T., and Weaver, D. T. (1993) *Genes Dev. 7*, 1755-1765.

97. Shen, Z., Cloud, K. G., Chen, D. J., and Park, M. S. (1996) *J. Biol. Chem. 271*, 148-152.

98. Sigurdsson, S., Van Komen, S., Bussen, W., Schild, D., Albala, J. S., and Sung, P. (2001) *Genes Dev. 15*, 3308-3318.

99. Van Dyck, E., Stasiak, A. Z., Stasiak, A., and West, S. C. (1999) *Nature 398*, 728-31.

100. Baumann, P., and West, S. C. (1998) *Trends Biochem. Sci. 23*, 247-251.

101. Alexeev, A., Mazin, A., and Kowalczykowski, S. C. (2003) *Nat. Struct. Biol. 10*, 182-186.

102. Liu, Y., Masson, J. Y., Shah, R., O'Regan, P., and West, S. C. (2004) *Science 303*, 243-6.

103. Sonoda, E., Sasaki, M. S., Buerstedde, J. M., Bezzubova, O., Shinohara, A., Ogawa, H., Takata, M., Yamaguchi-Iwai, Y., and Takeda, S. (1998) *EMBO J. 17*, 598-608.

104. Dronkert, M. L., Beverloo, H. B., Johnson, R. D., Hoeijmakers, J. H., Jasin, M., and Kanaar, R. (2000) *Mol. Cell. Biol. 20*, 3147-3156.

105. Pierce, A. J., Johnson, R. D., Thompson, L. H., and Jasin, M. (1999) *Genes Dev. 13*, 2633-2638.

106. Liu, N., Lamerdin, J. E., Tucker, J. D., Zhou, Z. Q., Walter, C. A., Albala, J. S., Busch, D. B., and Thompson, L. H. (1997) *Proc Natl Acad Sci U S A 94*, 9232-7.

107. Johnson, R. D., Liu, N., and Jasin, M. (1999) *Nature 401*, 397-399.

108. Tebbs, R. S., Zhao, Y., Tucker, J. D., Scheerer, J. B., Siciliano, M. J., Hwang, M., Liu, N., Legerski, R. J., and Thompson, L. H. (1995) *Proc. Natl. Acad. Sci. USA 92*, 6354-6358.

109. Miki, Y. (1998) *J Hum Genet 43*, 77-84.

110. Deininger, P. L., Moran, J. V., Batzer, M. A., and Kazazian, H. H., Jr. (2003) *Curr Opin Genet Dev 13*, 651-8.

111. Gupta, P. K., Sahota, A., Boyadjiev, S. A., Bye, S., Shao, C., O'Neill, J. P., Hunter, T. C., Albertini, R. J., Stambrook, P. J., and Tischfield, J. A. (1997) *Cancer Res. 57*, 1188-1193.

112. Moynahan, M. E., and Jasin, M. (1997) *Proc. Natl. Acad. Sci. USA 94*, 8988-8993.

113. Morley, A. A., Grist, S. A., Turner, D. R., Kutlaca, A., and Bennett, G. (1990) *Cancer Res. 50*, 4584-4587.

114. Shao, C., Deng, L., Henegariu, O., Liang, L., Raikwar, N., Sahota, A., Stambrook, P. J., and Tischfield, J. A. (1999) *Proc. Natl. Acad. Sci. USA 96*, 9230-9235.

115. Zhu, X., Dunn, J. M., Goddard, A. D., Squire, J. A., Becker, A., Phillips, R. A., and Gallie, B. L. (1992) *Cytogenet. Cell. Genet. 59*, 248-252.
116. Deininger, P. L., and Batzer, M. A. (1999) *Mol Genet Metab 67*, 183-93.
117. Kolomietz, E., Meyn, M. S., Pandita, A., and Squire, J. A. (2002) *Genes Chromosomes Cancer 35*, 97-112.
118. Schichman, S. A., Caligiuri, M. A., Gu, Y., Strout, M. P., Canaani, E., Bloomfield, C. D., and Croce, C. M. (1994) *Proc Natl Acad Sci U S A 91*, 6236-9.
119. Bishop, A. J., and Schiestl, R. H. (2003) *Exp. Mol. Pathol. 74*, 94-105.
120. Bishop, A. J., and Schiestl, R. H. (2001) *Biochim. Biophys. Acta. 1471*, M109-M121.
121. Bishop, A. J., and Schiestl, R. H. (2000) *Hum. Mol. Genet. 9*, 2427-2334.
122. Thompson, L. H., and Schild, D. (2002) *Mutat. Res. 509*, 49-78.
123. Luo, G., Santoro, I. M., McDaniel, L. D., Nishijima, I., Mills, M., Youssoufian, H., Vogel, H., Schultz, R. A., and Bradley, A. (2000) *Nature Genet. 26*, 424-429.
124. Friedberg, E. C., Walker, G. C., and Siede, W. (1995) *DNA Repair and Mutagenesis*, ASM Press, Washington, D.C.
125. Luo, C. M., Tang, W., Mekeel, K. L., DeFrank, J. S., Anne, P. R., and Powell, S. N. (1996) *J Biol Chem 271*, 4497-503.
126. Meyn, M. S. (1993) *Science 260*, 1327-1330.
127. Hisada, M., Garber, J. E., Fung, C. Y., Fraumeni, J. F., Jr., and Li, F. P. (1998) *J Natl Cancer Inst 90*, 606-11.
128. Sturzbecher, H. W., Donzelmann, B., Henning, W., Knippschild, U., and Buchhop, S. (1996) *Embo J 15*, 1992-2002.
129. Willers, H., McCarthy, E. E., Alberti, W., Dahm-Daphi, J., and Powell, S. N. (2000) *Int. J. Radiat. Biol. 76*, 1055-1062.
130. Gebow, D., Miselis, N., and Liber, H. L. (2000) *Mol. Cell. Biol. 20*, 4028-4035.
131. Kang, M. H., Genser, D., and Elmadfa, I. (1997) *Mutat. Res. 381*, 141-148.
132. Donovan, P. J., Smith, G. T., Lawlor, T. E., Cifone, M. A., Murli, H., and Keefer, L. K. (1997) *Nitric Oxide 1*, 158-166.
133. Tanaka, R. (1997) *J. Toxicol. Sci. 22*, 199-205.
134. Burney, S., Niles, J. C., Dedon, P. C., and Tannenbaum, S. R. (1999) *Chem. Res. Toxicol. 12*, 513-520.
135. Tamir, S., deRojas-Walker, T., Wishnok, J. S., and Tannenbaum, S. R. (1996) *Methods Enzymol. 269*, 230-243.
136. Sobol, R. W., Kartalou, M., Almeida, K. H., Joyce, D. F., Engelward, B. P., Horton, J. K., Prasad, R., Samson, L. D., and Wilson, S. H. (2003) *J. Biol. Chem. 278*, 39951-39959.
137. Spek, E. J., Wright, T. L., Stitt, M. S., Taghizadeh, N. R., Tannenbaum, S. R., Marinus, M. G., and Engelward, B. P. (2001) *J Bacteriol 183*, 131-8.
138. Spek, E. J., Vuong, L. N., Matsuguchi, T., Marinus, M. G., and Engelward, B. P. (2002) *J. Bacteriol. 184*, 3501-3507.
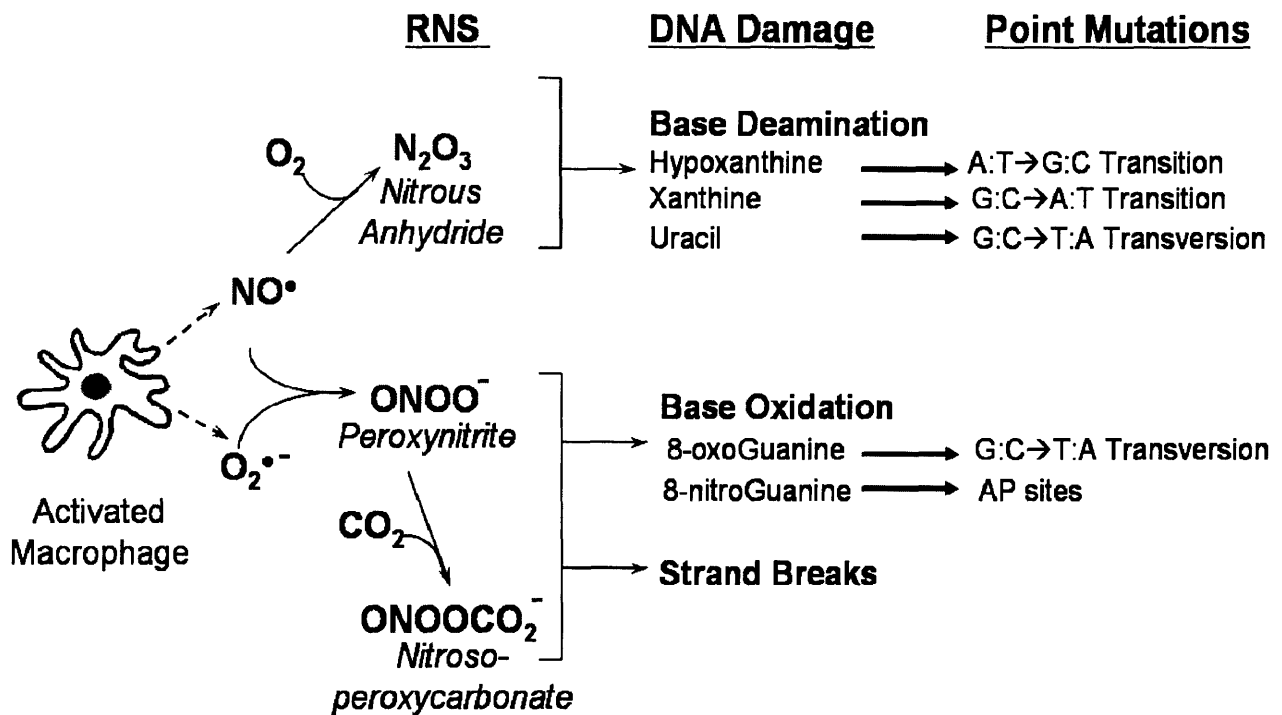
Figure 1.1. Schematics of the types of DNA lesions and point mutations induced by Nitric Oxide. Under physiological conditions Nitrous Anhydride and Peroxynitrite/ Nitroso-peroxycarbonate are the two predominant reactive nitrogen species (RNS) formed from NO•. The major classes of DNA lesions and the point mutations formed by these RNS are shown.
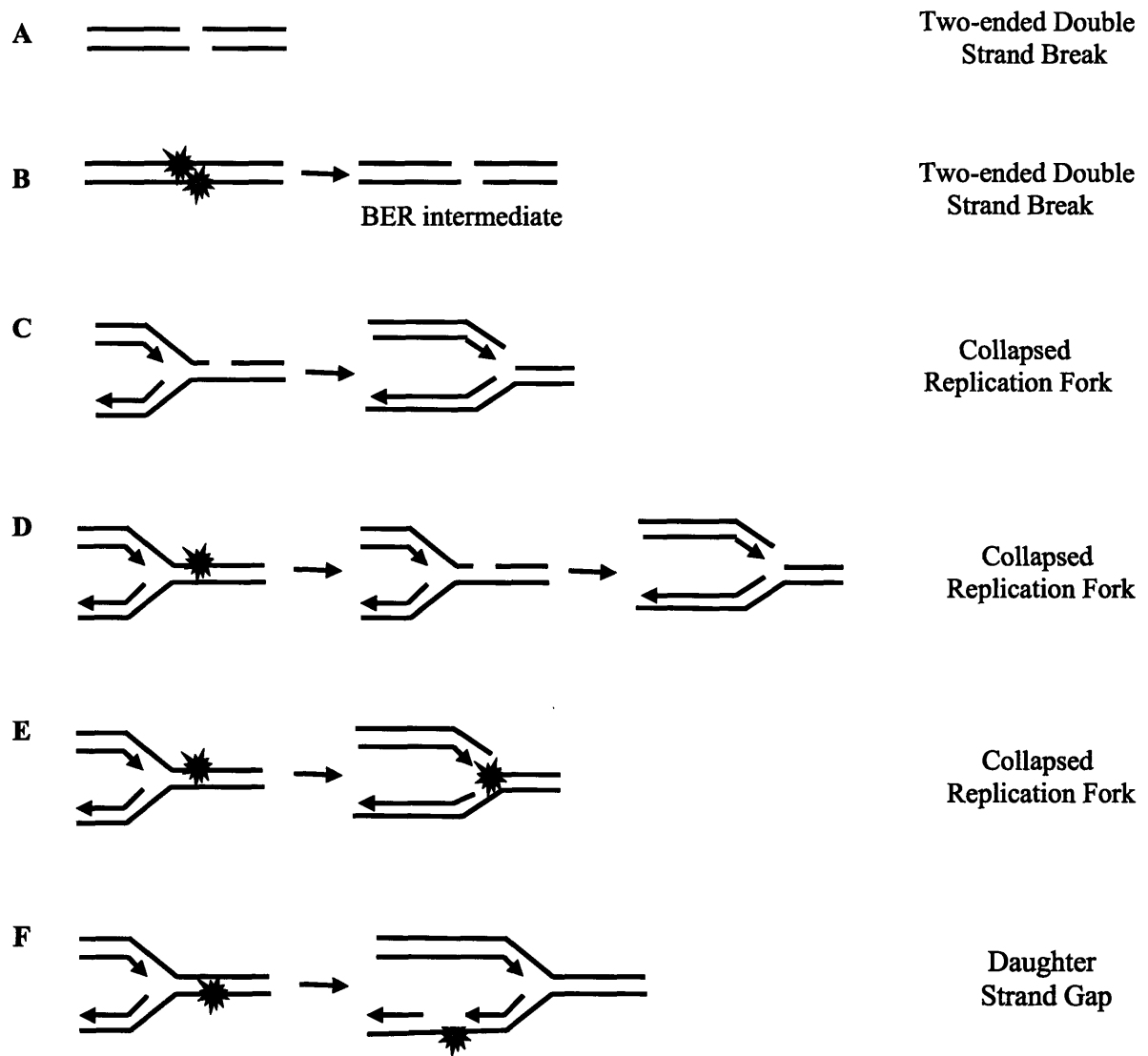
Figure 1.2. Classes of DNA lesions that induce homologous recombination. Replication independent mechanisms (A-B). Replication dependent mechanisms (C-F). (A) Two-ended double strand break can be directly formed, or can be formed from the proximity of two single strand breaks. (B) Two-ended double strand break can be formed from base excision repair intermediates. (C) Nicks can be converted to a one-ended double strand break when encountered by a replication fork. (D) Base damage can be converted into a nick as an intermediate of base excision repair, which can then be converted to a one-ended double strand break when encountered by a replication fork. (E) Inhibition of leading strand synthesis by a blocking lesion causes replication fork collapse forming a one-ended double strand break. (F) Inhibition of lagging strand synthesis by a blocking lesion leads to the formation of a daughter strand gap.
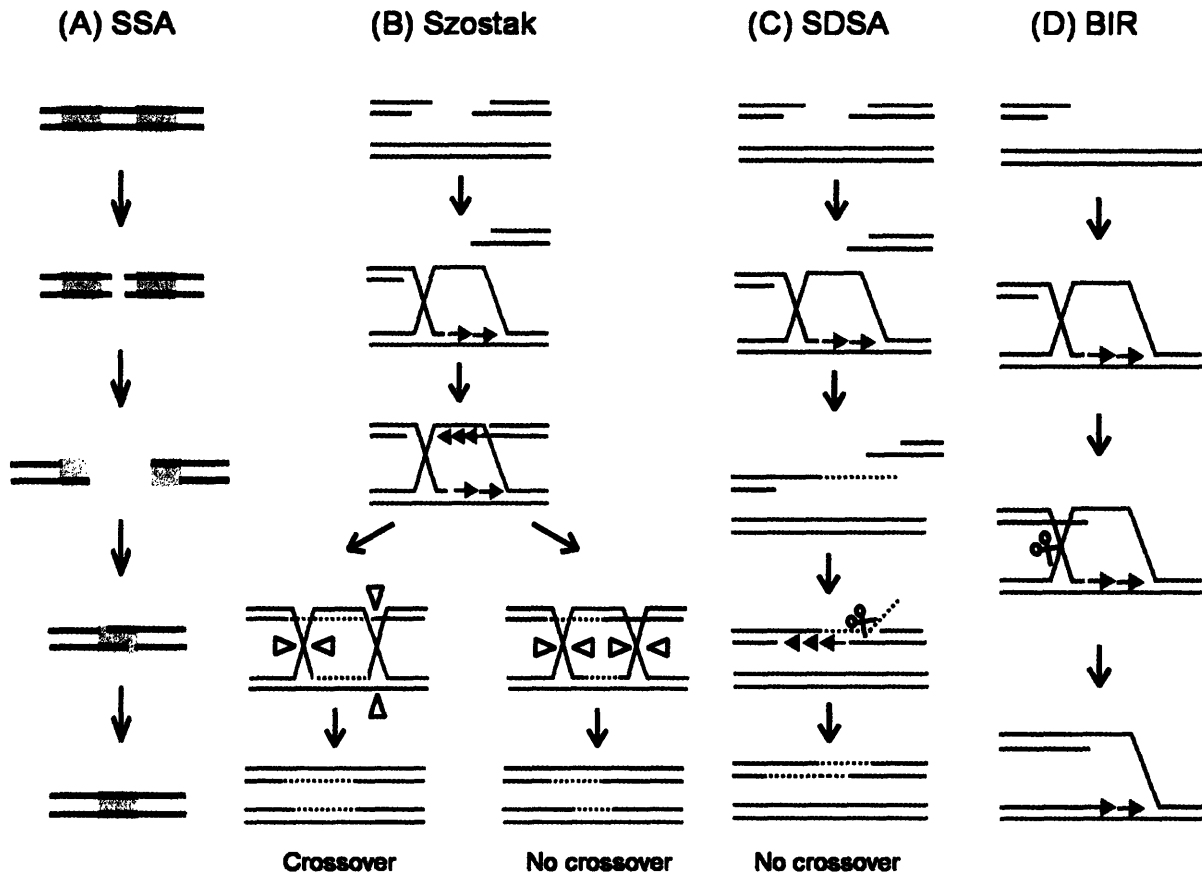
Figure 1.2. Mechanisms of Homologous Recombination. (A) Single Strand Annealing. The end of the double strand breaks are dissected till overhangs with significant sequence homology is reached. The homologous 3' overhangs are then annealed to restore intact duplex DNA. (B) Szostak Model. The ends of the DBS are resected forming 3' overhangs which then invade a homologous sequence. New DNA synthesis from the 3' overhangs leads to the formation of two Holliday junctions. Depending how these junctions are resolved, a GC event occurs either with crossing over or without crossing over. (C) Synthesis Dependent Strand Annealing (SDSA). The ends of the DBS are resected forming 3' overhangs. One of the overhangs invade the dublex initiating synthesis of new DNA. Once the missing information is restored, the invading strand is displaced and anneals to the 3' overhang of the other end enabling the synthesis of the missing information. No crossover events are associated with SDSA. (D) Break Induced Replication (BIR).The initial step is again the recession of the DSB end. The 3' overhang then invades the homologous DNA, forming a Holliday junction and enabling the synthesis of new DNA. The holliday junction is then resolved to complete the process. Figures are adapted from *(80,81)*.

40

A. BIR

Normal   Deletion   Insertion
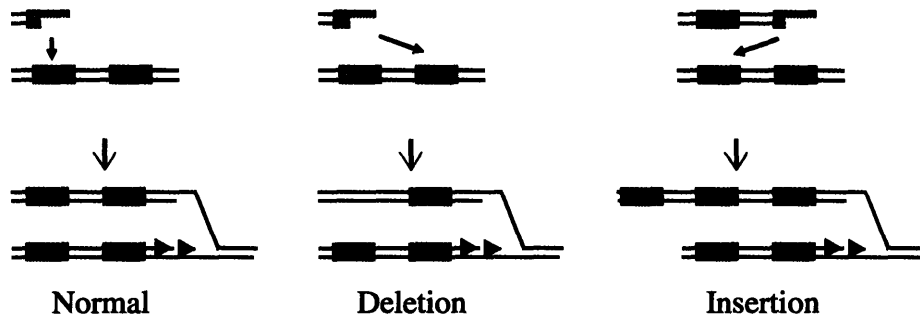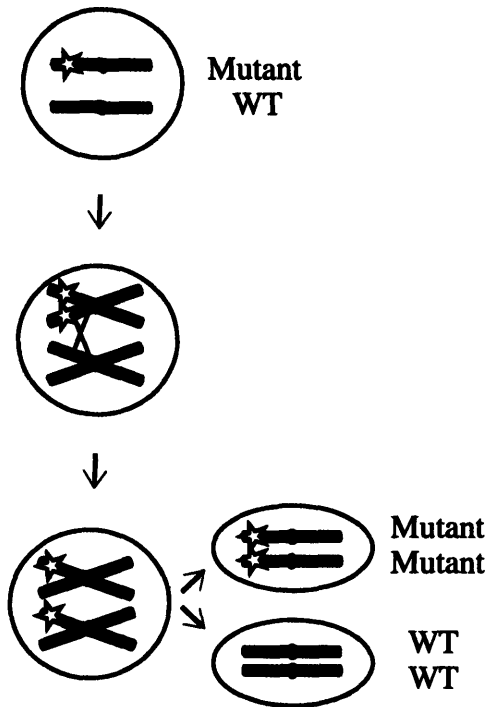
B. Crossover -LOH

Figure 1.3. Some examples of Gene Rearrangements that may result from homologous recombination. (A) Deletions and insertions may result from misaligned BIR pathway. (B) Schematics of how a crossover event during homologous recombination can lead to loss of heterozygosity (LOH). If the crossover event takes place between a mutant and a WT version of a gene, one of the daughter cells may end up with two copies of the mutated gene.

Neoplasm	Aetiologic agent
Bladder, liver, rectal carcinoma, follicular lymphoma of the Spleen	Schistosomiasis
Cervical carcinoma	Papillomavirus
Ovarian carcinoma	Pelvic inflammatory disease
Gastric adenocarcinoma, mucosa associated lymphoid tissue lymphoma	*Helicobacter pylori*
Mesothelioma carcinoma	Asbestos
Lung, bronchial carcinoma	Silica, asbestos, smoking
Kaposi's sarcoma	Human herpesvirus type 8
Colorectal carcinoma	Inflammatory bowel disease, Crohn's disease, chronic ulcerative colitis
Oesophageal carcinoma	Barret's metaplasia,
Hepatocellular carcinoma	Hepatitis virus B and /or C
Cholangiosarcoma, colon carcinoma	Liver flukes, bile acids

Modified from (1,6)

Table 1.1. Examples of associations between Infectious and Chronic Inflammatory conditions and Neoplasms

CHAPTER 2

Delineation of the Chemical Pathways Underlying

Nitric Oxide-Induced Homologous Recombination

in Mammalian Cells

Synopsis: Inflammation is an important risk factor for cancer. During inflammation, macrophages secrete $NO^\bullet$, which reacts with superoxide or oxygen to create $ONOO^-$ or N_2O_3, respectively. Although homologous recombination causes sequence rearrangements that promote cancer, little was known about the ability of $ONOO^-$ and N_2O_3 to induce recombination in mammalian cells. Here, we show that while $ONOO^-$ is a potent inducer of homologous recombination, N_2O_3 is at most only weakly recombinogenic. Furthermore, on a per lesion basis, $ONOO^-$-induced oxidative base lesions and single strand breaks are far more recombinogenic than N_2O_3-induced deamination products. Similar results were observed in mammalian cells from two different species, suggesting that the relative recombinogenicity of $ONOO^-$ and N_2O_3 is highly conserved in mammals. These results suggest that $ONOO^-$-induced recombination may be an important underlying mechanism of inflammation-induced cancer.

CHAPTER 2

Delineation of the Chemical Pathways Underlying

Nitric Oxide-Induced Homologous Recombination

in Mammalian Cells

2.1 Introduction

Nitric oxide (NO$^\bullet$) is a key mediator of diverse biological processes, including

vasodilation, neurotransmission, endotoxic shock, cerebral ischemia, and inflammation (*1-3*).

At low concentrations, NO$^\bullet$ is involved in intra- and inter-cellular signaling, while at high

concentrations NO$^\bullet$ is toxic (*4-6*). As a cytotoxic agent, NO$^\bullet$ is secreted by activated

macrophages, along with oxygen radicals and other cytotoxic chemicals during the

inflammatory response. Importantly, tissue inflammation associated with gastritis, hepatitis,

and colitis is an important risk factor for a variety of human cancers, such as gastric cancer,

liver cancer and cholangiocarcinoma (*7*). Furthermore, it is estimated that ~15% of cancer

cases worldwide are attributable to infectious diseases, many of which induce chronic

inflammation (*8*). Although the underlying mechanism of inflammation-induced cancer is not

yet fully understood, exposure of tissues to reactive oxygen and nitrogen species is thought to

cause mutations and genetic rearrangements that contribute to tumor initiation and

progression. While NO$^\bullet$-induced point mutations have been studied extensively (*e.g.*, see

references (*9-11*)), very little is known about the ability of NO$^\bullet$ to induce other classes of

mutations, such as sequence rearrangements that are mediated by homologous

recombination.

Mitotic homologous recombination allows cells to repair DNA double strand breaks by extracting missing sequence information from a sister chromatid (during S and G2 phases of cell cycle) or from a homologous chromosome. In addition, homologous recombination plays a critical role in restoring collapsed forks during DNA replication (for excellent reviews of homologous recombination, see references (*12-14*)). Although homologous recombination is generally highly accurate, transfer of genetic material carries with it a finite risk, since recombination between misaligned sequences can lead to insertions, deletions, inversions and translocations. In addition, exchanges between homologous chromosomes are responsible for causing most spontaneous loss of heterozygosity events in mammals (*15, 16*). Nearly all tumors carry sequence rearrangements that are likely the result of mitotic homologous recombination. Consequently, whether by exposure to endogenous and environmental recombinogens, or by inherited predisposition, conditions that lead to increased levels of homologous recombination are associated with an increased risk of cancer (*17-19*).

Very little is known about the recombinogenicity of NO$^{\bullet}$ and its derivative reactive nitrogen species in mammals. However, there are a few studies suggesting that NO$^{\bullet}$ may induce recombination in mammalian cells. For example, patients who suffer from chronic inflammation associated with Crohn's Disease have increased levels of sister chromatid exchanges (SCEs) in their lymphocytes (*20*), though it is not known to what extent NO$^{\bullet}$ is responsible for this effect. In addition, two studies have shown that mammalian cells exposed to chemicals that give rise to NO$^{\bullet}$ ('NO$^{\bullet}$ donors') suffer increased levels of SCEs (*21, 22*), though the NO$^{\bullet}$ donors used in these studies also give rise to additional potentially recombinogenic radical species. The observation that mammalian cells exposed to NO$^{\bullet}$ have

an increased susceptibility to loss of heterozygosity provides additional support for the possibility that NO• induces recombination (23). Taken together, these observations suggest that NO• may induce homologous recombination in mammalian cells. Here, we set out to define the potential of NO• to induce homologous recombination in mammalian cells, and to explore the underlying chemistry that might be responsible for its effects.

It is well established that the recombinogenicity of a DNA damaging agent is dependent on the types of DNA lesions that it induces. NO• reacts with oxygen and superoxide to form N_2O_3 and ONOO⁻, respectively, which are the dominant reactive nitrogen species formed under physiological conditions (Figure 2.1A). Although NO• does not directly damage DNA, N_2O_3 is a potent DNA deaminating agent, and ONOO⁻ is a potent DNA oxidizing agent. Most of the lesions created by N_2O_3 are base damages, while ONOO⁻ attacks both the base and sugar moities of DNA (11). For example, N_2O_3 deaminates DNA bases, creating lesions such as uracil, hypoxanthine, and xanthine (24-27), and ONOO⁻ oxidation of guanine leads to 8-oxoguanine and its secondary oxidation products, 8-nitroguanine, as well as abasic sites formed by spontaneous depurination of 8-nitroguanine (27-31). In addition to base lesions, ONOO⁻ also directly induces direct single strand breaks in DNA, by the oxidative breakdown of deoxyribose (29, 32, 33).

Although double strand breaks are thought to be critical for inducing homologous recombination, in vitro studies using purified DNA have shown that neither N_2O_3, nor ONOO⁻ efficiently creates double strand breaks by direct reaction with DNA (32, 34, 35). However, base lesions, abasic sites, and single strand breaks can be converted into double strand breaks by enzymatic processing (when in close proximity) or when they are encountered by the replication fork. For example, during base excision repair (BER), single

strand breaks may be created as repair intermediates by DNA glycosylases that have an associated AP lyase activity or by AP endonucleases. Consequently, if DNA glycosylases initiate BER of closely opposed lesions, double strand breaks can be formed *in vitro* (*36*). Alternatively, DNA lesions that inhibit replication fork progression, such as BER intermediates, are highly recombinogenic in mammalian cells (*e.g.*, see references (*37, 38*)). Indeed, previous work from this laboratory has demonstrated that DNA glycosylases promote NO$^{\bullet}$-induced recombination in *E. coli*, presumably by converting BER substrates into recombinogenic double strand breaks (*39*). Furthermore, studies in mammalian cells have recently shown that glycosylases promote radiation-induced strand breaks (*40*).

As a first step toward revealing the underlying mechanisms of NO$^{\bullet}$-induced sequence rearrangements in mammals, we set out to reveal the chemical basis for NO$^{\bullet}$-induced recombination in mammalian cells. Toward this end, we compared the recombinogenic effects of the two predominant reactive nitrogen species produced under physiological conditions: N_2O_3 and ONOO$^-$. For example, by exposing cells to NO$^{\bullet}$ and O_2 gases simultaneously in a specially designed NO$^{\bullet}$ gas delivery chamber (*41*), the major DNA damaging agent formed is N_2O_3. Alternatively, cells can be exposed to ONOO$^-$ using 3-morpholinosydnonimine (SIN-1), which produces equal amounts of superoxide and NO$^{\bullet}$ that rapidly react to form ONOO$^-$. However, in the case of the NO$^{\bullet}$ delivery chamber, due to the inevitable presence of intracellular superoxide, it is not possible to completely eliminate formation of some ONOO$^-$. Similarly, due to the presence O_2 during SIN-1 exposure, it is not possible to eliminate the possibility that some N_2O_3 is formed. Thus the NO$^{\bullet}$ delivery chamber and SIN-1, which are amenable for studies of cultured cells, provide effective strategies for creating conditions that favor the formation of ONOO$^-$ or N_2O_3, although not

with absolute purity. In contrast, *in vitro* conditions can readily be created in which plasmid DNA is virtually exclusively exposed to either N_2O_3 or $ONOO^-$. Therefore, the combined approaches of studying chromosomal recombination in cells exposed to conditions that favor formation of either N_2O_3 or $ONOO^-$, and studying plasmid recombination between DNA molecules exposed to conditions that give rise to exclusively N_2O_3 or $ONOO^-$, provide a powerful strategy for elucidating the chemical nature of $NO^{\bullet}$-induced recombination in mammalian cells.

Here, we show that exposure of mammalian cells to conditions that favor formation of N_2O_3 is relatively weakly recombinogenic, and that N_2O_3-induced DNA lesions fail to induce inter-plasmid recombination. In contrast, conditions that favored formation of $ONOO^-$ are highly recombinogenic, both when cells are exposed and when plasmid is exposed to $ONOO^-$ and subsequently transfected into cells. By quantifying strand breaks, abasic sites, and base lesions induced by N_2O_3 and $ONOO^-$ in plasmid DNA, we demonstrate that chemically-induced strand breaks and oxidative base lesions are far more recombinogenic than deaminated base lesions, even under conditions where the quantity of N_2O_3-induced lesions exceeds that of $ONOO^-$-induced lesions. The results of these studies contribute to our understanding of the biological consequences of chronic inflammation and provide a framework for delineating conditions that are most likely to lead to tumorigenic sequence rearrangements in people.

2.2 Experimental Procedures

Cell Culture

Mouse embryonic stem cells (J1) were maintained in DMEM containing 10% fetal bovine serum, L-glutamine, non-essential amino acids, β-mercaptoethanol and leukocyte inhibitory factor on gelatinized dishes without feeders. The COS-7L African Green Monkey kidney cell line (Invitrogen) was maintained in DMEM containing 10% fetal bovine serum and L-glutamine.

ΔGF Cells

Embryonic stem cells were engineered to carry a direct repeat recombination substrate integrated at the ROSA26 locus. The design of this substrate is similar to that of the FYDR substrate (42), except that reconstitution of full length coding sequences leads to expression of enhanced green fluorescent protein (rather than yellow fluorescent protein). This recombination detection system will be more fully described elsewhere (Jonnalagadda *et al.*, manuscript in preparation).

SIN-1 Treatment of Cells

Cells were plated in complete medium 24 h prior to treatment. Cells were exposed to the indicated doses of SIN-1 (Biomol Research Laboratories) in Dulbecco's PBS supplemented with 25 mM sodium bicarbonate, pH 7.4, at 37°C. Exposure of cultures to air was provided by shaking for 5 min every 10 min for 90 min, as described previously (43). After treatment, cultures were rinsed in PBS and incubated in fresh medium.

NO• Treatment of Cells

Cells were plated at a density of 0.5 - 2 x 10^6 cells per 60 mm dishes in complete medium. After 24 h, log phase cells were exposed to 10% NO• / 90% Ar gas mixture (BOC Gas Co.) which was passed through 7 cm of Silastic tubing in the NO• delivery chamber, as previously described (41). A 50% O_2 / 45% N_2 / 5% CO_2 gas mixture (BOC Gas Co.) was passed through a second 7 cm long tubing to maintain the O_2 level in the medium. These conditions produced steady-state levels of O_2 and NO• that are approximately those that are experienced under physiological conditions (NO• steady state levels in these experiments are 1.84 μM; physiological levels are on the order of ~1 μM) (4, 44, 45). The control cells were exposed to 100% Ar gas under the same conditions as NO• exposure. At the end of treatment, cells were washed and incubated in fresh complete medium.

Colony Forming Assay

Following treatment with SIN-1 and NO•/O_2, cells were harvested by trypsinization, counted and plated at a density of 100 cells/100-mm tissue culture dishes. After a week of culture, the colonies were fixed with ethanol, stained with trypan blue, and counted. Experiments were performed in triplicates and all experiments were performed three or more times.

SCE Assay

J1 cells were seeded at a density of 1 x 10^6 per 100 mm dish for SIN-1 treatments, and at 0.5 x 10^6 per 60-mm dish for NO•/O_2 treatment. After 24 h, log phase cultures were exposed to SIN-1 and NO•/O_2 as described above. After exposure, cells were washed and

incubated in fresh McCoy's 5A media (Invitrogen) supplemented with 10 µM BrdU. After 24

hours, Colcemid (0.1 µg/ml; Invitrogen) was added to the medium for an additional 3 h.

Metaphase spreads were prepared and sister chromatids were differentially stained as

previously described (*38*). Experiments were performed in duplicates, and twenty metaphase

spreads were counted per data point.

Chromosomal Direct Repeat Assay

Between 0.5 and 2.5 x 10^5 ΔGF cells were seeded onto 6 well dishes for SIN-1

treatments, and onto 60 mm tissue dishes for NO$^{\bullet}$/O$_2$ treatment. After 24 h, log phase cells

were exposed to the indicated doses of cytotoxic agents, as described above. After 72 h

(~3.25 population doublings in control cells), cells were harvested by trypsinization and

analyzed by flow cytometry to determine the frequency of fluorescent cells.

Plasmid Construction to Generate Recombination Substrates

The *Pst*I-*Bam*HI fragment in pCX-*EGFP* (gift of M. Okabe; (*46*)) was replaced with

a synthetic adaptor carrying *Nsi*I, *Not*I, and *Xho*I sites to create pCX-NNX-*EGFP*. Full

length and truncated coding sequences were PCR amplified using primers that carry synthetic

*Apo*I sites and subcloned between the *Eco*RI sites of pCX-NNX to create pCX-NNX-Δ5*egfp*

(abbreviated as pΔ5*egfp*) and (abbreviated as pΔ3*egfp*). The pΔ5*egfp* plasmid carries a 99 bp

deletion from the 5' end, and pΔ3*egfp* carried an 81 bp deletion from the 3' end of the open

reading frame. There are 540 bp of overlap between the deletions. Stop codons were placed

after and before these cassettes, respectively, to prevent read-through translation products.

Plasmid DNA was purified using the Qiagen plasmid purification kit. Restriction enzymes

were from New England Biolabs. Oligonucleotides were from Amitof, Inc. (Allston, MA). Oligonucleotide and plasmid sequences are available upon request.

ONOO⁻, SIN-1, and NO• treatment of plasmid DNA

ONOO⁻ was synthesized by ozonolysis of sodium azide (*47*) and the concentration of the stock solution was measured by UV immediately prior to treatment (ε_{302}=1670 M⁻¹ cm⁻¹). Bolus exposure to ONOO⁻ was achieved by placing a droplet of the ONOO⁻ stock solution on the side wall of a microcentrifuge tube containing plasmid DNA (2.5 μg in 50 μl) in 150 mM potassium phosphate / 20 mM sodium bicarbonate buffer, pH 7.4 followed by vigorous vortexing to mix the solutions. For SIN-1, plasmid DNA (5 μg in 100 μl) was treated in 150 mM potassium phosphate / 20 mM sodium bicarbonate buffer, pH 7.4, at 37°C for 90 min. For NO• treatment, plasmid DNA (20 μg/ml) was exposed in the NO• delivery chamber, under the same conditions as used for the cell exposures (see above). The NO• dose was controlled by removing samples at various times during exposure. For all treatments, after exposure, the DNA was recovered by ethanol precipitation, resuspended in PBS, and quantified by UV spectrometry.

Interplasmid Recombination Assay (extracellular DNA exposure)

J1 and COS-7L cells were plated at a density of 4 x 10⁴ cells per well in 24 well plates. After 24 hours, 0.1 μg pΔ5*egfp* (± exposure to DNA damage), and 1 μg pΔ3*egfp* plasmids were co-lipofected into the cells by using Lipofectamine 2000 (Invitrogen) in a total volume of 0.5 ml. To assay inhibition of *EGFP* expression, 0.1μg of p*EGFP* was treated with the indicated doses of these agents prior to lipofection. Transfection efficiency in each

experiment was measured by simultaneous lipofection of 0.1 μg p*EGFP* in three

independent samples. To keep DNA concentrations constant, 0.1 μg of damaged pLacZ

(Clontech) was used in lieu of damaged pΔ5*egfp*, and 1 μg of pLacZ was used in lieu of

pΔ3*egfp*. For all experiments, cells were trypsinized 48 h post lipofection, and analyzed by

flow cytometry to determine the frequency of fluorescent cells.

Interplasmid Recombination Assay (intracellular DNA exposure)

J1 and COS-7L cells were plated at a density of 4 x 10^4 cells per well in 6 well plates.

After 24 hours, 0.3 μg pΔ5*egfp*, and 0.3 μg pΔ3*egfp* plasmids were co-lipofected into the

cells in a total volume of 1 ml as described above. After 24 h, cells were treated with SIN-1

or $NO^{\bullet}/O_2$ as described above. Cells were analyzed by flow cytometry 48 hours post

treatment. The transfection efficiency was tested by simultaneous lipofection of 0.3 μg

p*EGFP* into three independent samples.

Flow Cytometry

Pelleted cells were resuspended in OptiMEM (GIBCO/BRL) and passed through a 70

μm filter (Falcon) prior to analysis on a Becton Dickinson FACScan flow cytometer

(excitation 488 nm, argon laser; emission 580/30). Live cells were gated according to

forward and side scatter.

Quantification of DNA damage by plasmid topoisomer analysis

A plasmid nicking assay was used to quantify the DNA deoxyribose and base

damage. For lesion quantification, similar results were obtained using either the pΔ5*egfp* or

the plasmid pSP189. The latter is a pBR-based, 4952 bp plasmid, which is similar in size to

p$\Delta 5egfp$ (4921 bp), which allows interchangeability of the two plasmids in the nicking assay.

After ONOO⁻, SIN-1 or NO˙/O_2 treatments, a portion (200 ng) of the DNA was treated with

putrescine (100 mM, pH 7.0, 1 h, 37°C) to convert all types of abasic sites to strand breaks

(*48, 49*). Another portion (200 ng) was kept on ice as a control for the direct strand breaks.

To quantify the oxidative base lesions formed by ONOO⁻ or SIN-1, a third portion of the

DNA sample was treated with *E. coli* formamidopyrimidine-DNA glycosylase (Fpg)

(Trevigen) to convert oxidized purines to strand breaks. More than 90% of the ONOO⁻

guanine oxidation products in DNA are recognized by Fpg (*49*). Fpg treatment was

performed in a volume of 10 μL containing 200 ng of ONOO⁻-treated DNA, 1 μL of Fpg

(~1.5 Units) and buffer containing 10 mM Tris-HCl (pH 7.5), 1 mM EDTA and 100 mM

NaCl at 37°C for 1 h. Fpg was then removed by phenol/chloroform extraction. Similarly, to

quantify nucleobase deamination products formed by NO˙ in the presence of O_2 (*e.g.*, N_2O_3),

plasmid DNA samples were incubated with *E. coli* Ung DNA glycosylase (Trevigen) and the

E. coli AlkA DNA glycosylase (Pharmingen), followed by putrescine treatment to convert

the resulting abasic sites to strand breaks. Both xanthine and hypoxanthine have been shown

to be recognized by AlkA ((*50*) and Dong *et al.*, manuscript in preparation). Cross reactions

of AlkA/Ung with ONOO⁻-treated DNA and Fpg with N_2O_3-treated DNA revealed <5%

additional strand breaks, which demonstrates the specificity of these enzymes for the specific

sets of lesions caused by N_2O_3 and ONOO⁻. The recovered DNA was redissolved in H_2O and

plasmid topoisomers were resolved by 1% agarose slab gel electrophoresis in the presence of

0.1 μg/mL ethidium bromide. The quantity of DNA in each band was determined by

fluorescence imaging (Ultra-Lum). For lesion quantification, similar results were obtained

using either the pΔ5*egfp* or the pSP189 plasmid (provided by Dr. M. Seidman; NIH,

Bethesda, MD), with the pSP189 plasmid used for the experiments presented in Figure 2.5.

Note that the base damage data for SIN-1 and ONOO⁻ has been previously reported (*51*).

2.3 Results

SIN-1, but not NO•/O2, induces homologous recombination at a chromosomally integrated direct repeat.

To study NO•/O2-induced homologous recombination in mammalian cells, we used mouse embryonic stem cells that carry a direct repeat substrate that yields a fluorescent phenotype specifically as a result of mitotic homologous recombination events (Jonnalagadda *et al.*, manuscript in preparation). The ΔGF cells carry expression cassettes for two different truncated *egfp* cDNAs that are arranged in a direct repeat (Figure 2.1B; essential coding sequences were deleted from either the 5' or the 3' end to create Δ5*egfp* and Δ3*egfp* respectively). In this system, recombination via several different mechanisms can reconstitute the full length *EGFP* sequence (*e.g.*, gene conversion, unequal sister chromatid exchanges and break-induced replication; (*14*)). Thus, recombination frequency can be readily estimated by quantifying green fluorescent recombinant cells by flow cytometry (*42, 52*).

To study the effects of ONOO⁻ and N2O3, ΔGF cells were exposed to various concentrations of SIN-1 (to deliver ONOO⁻) or to NO• and O2 gases delivered simultaneously via a Silastic tubing-based NO• delivery chamber to deliver N2O3 (referred to as 'NO•/O2' exposure below) (*41*). It is well established that for most DNA damaging agents, the frequency of recombination rises as the toxicity increases. Therefore, to compare the recombinogenic effects of ONOO⁻ and N2O3, we established conditions that induce similar levels of cytotoxicity (Figure 2.1C). Cells were then exposed to equitoxic doses of SIN-1 or NO•/O2 and were subsequently analyzed by flow cytometry to quantify fluorescent cells. As can be seen in Figure 2.1D, SIN-1 is clearly recombinogenic, inducing recombination up to

~3-fold. It is noteworthy that induction of recombination by SIN-1 initially increases and then decreases as a function of dose, an observation that has also been noted by others (*e.g.*, see reference (*53*)). In contrast to SIN-1, exposure to $NO^{\cdot}/O_2$ did not lead to any detectable increase in the frequency of recombinant cells (Figure 2.1D), even at doses that diminished survival to below 1% (data not shown). Thus, while SIN-1 induces recombination at a chromosomally integrated recombination substrate, $NO^{\cdot}/O_2$ does not.

SIN-1 is a more potent inducer of SCEs than $NO^{\cdot}/O_2$.

The chromosomal direct repeat assay offers a rapid and effective measure of recombination at a single locus. To assess recombination throughout the genome rather than a single locus, we measured SCEs, which are known to result from homologous recombination events (*54*). By culturing cells in the presence of BrdU for two cell cycles, exchanges between differentially stained sister chromatids can be visualized in metaphase spreads. For these experiments, cells were exposed to either SIN-1 or $NO^{\cdot}/O_2$, cultured in BrdU-containing media for two cell doublings, and analyzed as previously described (*38*). As can be seen in Figure 2.1E, SIN-1 significantly induces SCEs (~6-fold). In addition, although significantly less recombinogenic than SIN-1, $NO^{\cdot}/O_2$ exposure induces SCEs (~2-fold). Although these results suggest that N_2O_3 induces homologous recombination, it is also possible that $ONOO^-$ that is inevitably formed by reaction of $NO^{\cdot}$ with intracellular superoxide is responsible for this increase in SCE levels in the $NO^{\cdot}/O_2$-exposed cells.

Creation of an inter-plasmid recombination detection assay.

To further explore the relative recombinogenicity of ONOO⁻- or N_2O_3-induced DNA

lesions, we developed an inter-plasmid recombination assay in which plasmids can be

exposed to pure ONOO⁻ or to conditions that generate exclusively N_2O_3 prior to transfection

into the cells. For this purpose, we used plasmids that carry Δ3_egfp_ and Δ5_egfp_ (Figure

2.2A). A recombination event between these two plasmids can restore full length _EGFP_,

yielding a fluorescent phenotype. To test the efficacy of this approach, cells were transfected

with p_EGFP_, pΔ3_egfp_, pΔ5_egfp_, or they were cotransfected with both pΔ3_egfp_ and pΔ5_egfp_.

Whereas ~40% of the cells transfected with the positive control p_EGFP_ plasmid were

significantly fluorescent, transfection with either pΔ5_egfp_ or pΔ3_egfp_ alone did not lead to

any detectable _EGFP_ expression (Figure 2.2B). However, when cells were co-transfected

with pΔ5_egfp_ and pΔ3_egfp_ together, we observed a significant number of fluorescent cells,

indicative of inter-plasmid homologous recombination (related assays have previously been

described; _e.g._, see references (_55, 56_)).

A time course experiment revealed that the number fluorescent cells co-transfected

with pΔ5_egfp_ and pΔ3_egfp_ increases over time, reaching a maximal level by ~48 h, which

subsequently begins to decline sometime after 72h post-transfection. This rise and fall in the

frequency of fluorescent cells likely reflects the time required to achieve maximal expression

of the _EGFP_ protein, followed by the inevitable loss of plasmid DNA associated with

transient transfections. It is also noteworthy that the cells that carry a single copy of the

integrated substrate fluoresce following homologous recombination (which shows that a

single copy of an _EGFP_ expression cassette is sufficient to yield a detectable fluorescent

signal), however it is not known how many *EGFP* expression cassettes on extrachromosomal DNA are required for such a signal.

In order to use this inter-plasmid assay to study the relationship between DNA damage and recombination, it is critical to control for transfection efficiency, since cells that receive higher quantities of plasmid have an increased chance of undergoing inter-plasmid recombination. Under the conditions of these experiments, we found that the transfection efficiency was highly consistent from sample to sample (less than a 10% variation among samples; data not shown). To control for transfection efficiency, the frequencies of inter-plasmid recombinant cells were normalized according to the percentage of fluorescent cells transfected with the positive control vector (see Experimental Procedures).

Double strand breaks are known to be potent inducers of homologous recombination, and it has previously been shown that double strand breaks in plasmids can be repaired by recombining with homologous sequences available on another plasmid (*55*). To maximize the sensitivity of the inter-plasmid assay for detecting damage-induced recombination, we introduced a double strand break into pΔ5*egfp* using *BamH*I (Figure 2.2A), and measured double strand break-induced recombination under various conditions. Following extensive optimization, we found that we could increase the sensitivity of the assay to damage-induced recombination by an order of magnitude by co-transfecting damaged pΔ5*egfp* with a 10-fold molar excess of pΔ3*egfp* (Figure 2.2C) and by transfecting cells via lipofection, as opposed to using other transfection methods (data not shown). Given that recombination between two copies of pΔ5*egfp* can never reconstitute full length *EGFP* coding sequences, conditions in which there is a 10-fold molar excess of pΔ3*egfp* increase the likelihood that recombination events will be detectable by the appearance of a fluorescent signal.

59

Inter-plasmid homologous recombination is induced when plasmid DNA is exposed to ONOO⁻ and SIN-1, but not to $NO^•/O_2$.

To study the effects of ONOO⁻-induced DNA lesions, pΔ5*egfp* was exposed to SIN-1 or to purified ONOO⁻ (see Experimental Procedures). To study the effects of N_2O_3-induced DNA lesions, pΔ5*egfp* was exposed to N_2O_3 in the $NO^•/O_2$ delivery chamber, as described previously (*41*). The damaged pΔ5*egfp* was then mixed with undamaged pΔ3*egfp* and transfected into mouse ES cells. Although the frequency of recombination between plasmids could be assessed by quantifying fluorescent cells, if there are too many lesions, then even if there is a productive recombination event, *EGFP* will not be expressed (*e.g.*, due to inhibition of transcription, concurrent mutations, or degradation of damaged plasmid). To determine the optimal dose range, p*EGFP* was exposed to various concentrations of the damaging agents, transfected into mouse cells and the proportion of fluorescent cells was assessed. The highest dose at which no significant inhibition of *EGFP* expression was observed was then taken as the upper limit for studies of recombination (Figure 2.3A). It is noteworthy that even at very high levels of $NO^•/O_2$ exposure, we did not observe any significant inhibition of *EGFP* expression (Figure 2.3A, right). Therefore, we quantified the number of lesions created under these exposure conditions (described in detail below), and selected an $NO^•/O_2$ dose range that creates a comparable number of lesions per plasmid compared to SIN-1 and ONOO⁻. As can be seen in Figure 2.3B, DNA lesions induced by ONOO⁻ significantly induced homologous recombination (~2-fold) and DNA exposed to SIN-1 was similarly susceptible to damage-induced recombination. However, DNA lesions created by $NO^•/O_2$ exposure did not induce any increase in the frequency of recombination (Figure 2.3B, right).

60

Effects of ONOO⁻, SIN-1, and NO˙/O₂ on inter-plasmid homologous recombination are conserved in diverse cell types.

A major advantage to studying damage-induced recombination using extrachromosomal plasmids is that these experiments can be performed in different mammalian cell types. To determine if the observed effects of ONOO⁻, SIN-1 and NO˙/O₂ on inter-plasmid homologous recombination are similar in other mammalian cell types, we repeated these experiments using COS-7L African green monkey kidney cells. Despite the fact that these are very different cell types from two different species, we saw remarkably similar dose-response curves when we assayed for inhibition of *EGFP* expression (Figure 2.3C). Furthermore, although there are some subtle differences in the responses of these two cell types to damage-induced recombination (*e.g.*, the minimal dose required to induce recombination), overall the relative recombinogenicity of these exposures is highly similar (Figure 2.3D).

Exposure of mammalian cells to ONOO⁻ and SIN-1, but not to NO˙/O₂ induces extrachromosomal recombination.

One of the strengths of the inter-plasmid recombination assay is that by damaging the plasmid prior to transfection into cells, the effects of specific DNA lesions can be studied. Another advantage of this assay is that it can be readily adapted to many different cell types. Finally, in addition to introducing damage to the plasmids prior to transfection, the cells can also be exposed to damaging agents post transfection to explore the possibility that

concurrent cellular responses to damage-exposure significantly modify the effects of DNA lesions on recombination.

We have shown above that the relative recombinogenicity of DNA lesions created by $ONOO^-$ and SIN-1 versus $NO^\cdot/O_2$ are similar in both mouse ES and monkey COS 7L cells (Figure 2.3). We next asked if different cell types would respond similarly when the cells themselves, rather than plasmid DNA, are exposed to SIN-1 and $NO^\cdot/O_2$. To select the appropriate dose-range, we first established the dose dependent survival of COS 7L cells to SIN-1 or $NO^\cdot/O_2$ (Figure 2.4C; note that Figure 2. 4A is the same as Figure 2.1C, and is included here for ease of comparisons). COS 7L cells appear to be more resistant to the toxic effects of both SIN-1 and $NO^\cdot$ when compared to mouse ES cells (compare Figures 2. 4A and 2.4C; note the change in the dose range).

To study inter-plasmid recombination in exposed cells, we co-transfected cells with undamaged pΔ5egfp and pΔ3egfp and after 24 h (to allow time for recovery from transfection), we exposed the cells to increasing concentrations of SIN-1 or to $NO^\cdot/O_2$. After an additional 48 h, the cells were analyzed by flow cytometry (note that similar results were observed in pilot experiments assayed either 48 h or 72 h post damage-exposure). For both ES and COS 7L cells, exposure to SIN-1 after transfection caused a dose dependent induction of recombination (Figures 2.4B and 2.4D). Importantly, the same dose of SIN-1 resulted in a similar fold induction in both cell types. For example, exposure to 1 mM SIN-1 induces an ~2 fold increase in the frequency of recombination in both cell types, regardless of whether the exposure occurred before or after transfection (Figures 2.3B-2.3D). Therefore, it does not appear that cellular responses to damage-exposure dramatically alter the magnitude of SIN-1 induced inter-plasmid homologous recombination (despite the

differences in toxicity in these cell types at the selected doses). Consistent with the previous analysis of recombination between p$\Delta 5egfp$ and p$\Delta 3egfp$ (*e.g.*, Figure 2.1D and Figure 2.3), we did not observe any significant induction of recombination when cells were exposed to NO$^{\bullet}$/O$_2$. Thus, it again appears that DNA lesions induced by ONOO^{-} are significantly more recombinogenic than those induced by NO$^{\bullet}$/O$_2$.

ONOO^{-}-induced direct single strand breaks and oxidative base lesions are recombinogenic.

Reactive nitrogen species produce three major classes of lesions: base damage, abasic sites and single strand breaks. To gain a better understanding of how these different classes of lesions affect homologous recombination, we used a well established plasmid-nicking assay to quantify lesions generated when plasmid is exposed to ONOO^{-}, SIN-1, or N$_2$O$_3$ (*24, 33*). Topoisomer analysis was performed to quantify single strand breaks (by direct analysis of damaged plasmids), abasic sites (which were converted to strand breaks using putrescine), and base lesions (which were converted to strand breaks using a combination of purified DNA glycosylases and putrescine). The relative proportions of supercoiled, relaxed and linear plasmid molecules were then used to calculate the number of DNA lesions (see Experimental Procedures).

An example of an experiment in which the frequency of Fpg sensitive lesions was estimated is shown in Figure 2.5A. In the absence of Fpg, with increasing doses of ONOO^{-}, there is a gradual shift from Form I (supercoiled DNA) to Form II (nicked closed-circular DNA), which reflects the ability of ONOO^{-} to directly induce single strand breaks. Plasmid was then digested with purified Fpg glycosylase, which both removes damaged bases and cleaves the backbone via its associated lyase activity (*57*). In the case of control undamaged

plasmid DNA that was analyzed before and after incubation with Fpg glycosylase (Figure 2.5A, lane 1), there is very little change in the ratio of supercoiled (Form I) and nicked (Form II) following digestion with Fpg (Fpg produced a small amount of the plasmid nicking , most likely as a consequence of cutting at background damage). In contrast, at the highest concentration of ONOO⁻ (5 mM, far right lane), digestion with Fpg eliminates nearly all of the supercoiled Form I DNA, indicating that essentially all of the plasmids had at least one Fpg sensitive site. Importantly, it also appears that Fpg induces a small but detectable level of double strand breaks, as indicated by the appearance of linear Form III DNA, made more apparent by the adjustment in contrast shown in Figure 2.5B.

The ratio of Forms I and II can be used to estimate the number of lesions per million nucleotides (although the total amount of DNA per lane can vary, depending on loading, the ratio remains constant). Using this topoisomer approach, we estimated the levels of chemically-induced single strand breaks, abasic sites, and base lesions. The distinction between guanine oxidation products arising from ONOO⁻ (*49*) and the nucleobase deamination products arising from N_2O_3 (*24*) was achieved by differential recognition of the lesions by Fpg and Ung/AlkA, respectively (see Experimental Procedures). We found that ONOO⁻ creates primarily base damage and single strand breaks, whereas SIN-1 creates primarily single strand breaks (Figure 2.5C). In contrast, NO• exposure creates almost exclusively base lesions, with a small number of abasic sites and an undetectable level of single strand breaks, (Figure 2.5C), which is consistent with our published studies of N_2O_3-induced DNA damage spectrum (*24*). It is noteworthy that base lesions and abasic sites are masked by the presence of single strand breaks in the topoisomer assay. Thus, above 60 lesions per million nucleotides, there is more than one damage event per plasmid molecule

on average (according to a Poisson distribution), so that the calculated levels of damage underestimate the true level of damage (this is likely the reason for the appearance of a plateau in base damage levels with doses of ONOO⁻ above 0.01 mM; Figure 2.5C). This behavior is independent of the identity of the DNA damaging agent, however, so corrective measures are not necessary when making comparisons of different agents, as in the present studies.

In order to gain a better understanding of the classes of lesions that lead to recombination, we compared the quantity and quality of lesions created by the lowest recombinogenic dose of ONOO⁻ (0.02 mM) to those created by the highest dose of NO˙ (2.7 mM·min) used for the interplasmid recombination assay. Although 2.7 mM·min NO˙ creates more abasic sites and many more base lesions than 0.02 mM ONOO⁻ (Figure 2.5D), NO˙ was not recombinogenic whereas ONOO⁻ was highly recombinogenic. Importantly, NO˙ does not create any detectable single strand breaks, whereas significant levels of single strand breaks were detected at recombinogenic doses ONOO⁻. Note that the masking effect of single strand breaks likely results in only a minor underestimate of the number of base lesions at 0.02 mM ONOO⁻, so on a per lesion basis, ONOO⁻ is more recombinogenic than NO˙. We therefore conclude that single strand breaks induced by deoxyribose oxidation and oxidatively damaged bases are significantly more recombinogenic than nucleobase deamination products in mammalian cells.

It is well established that double strand breaks induce homologous recombination in mammalian cells (*58, 59*). Although DNA exposed to 0.2 mM ONOO⁻ was highly recombinogenic, none of the plasmid DNA appeared to have been linearized by exposure to 0.2 mM ONOO⁻ (this dose falls between lanes 8 and 9 in the top gel in Figure 2.5A).

However, following digestion with Fpg *in vitro*, double strand breaks can be detected by the appearance of linear DNA (see lanes 8 and 9 in the lower gel of Figure 2.5A, and in 2.5B). Intriguingly, at 1-5 mM ONOO⁻, digestion with Fpg essentially eliminates the supercoiled Form I DNA, and it is around this dose range that the frequency of induced recombination appears to decline (compare 0.2 and 2 mM ONOO⁻ in Figure 2.5D, left), and the percentage of cells able to express *EGFP* from the ONOO⁻ treated positive control pEGFP vector starts to plummet (Figure 2.3A and 2.3C). When DNA is transfected into mammalian cells, most of the DNA is rapidly degraded by cytosolic nucleases (*60*). Furthermore, it has been shown that whereas supercoiled plasmid yields high levels of expression from a marker transgene 72 h post lipofection, linearized DNA yields no detectable expression (*61*). Given that mammalian cells harbor multiple DNA glycosylases that act on a broad range of substrates (*57, 62*), it is likely that a much greater proportion of the damaged plasmid is linearized by mammalian glycosylases following transfection, than by Fpg digestion *in vitro*, rendering the DNA vulnerable to nucleolytic degradation. Taken together, these results suggest that following transfection, mammalian DNA glycosylases create double strand breaks that both induce homologous recombination, and render the DNA vulnerable to degradation by exonucleases.

2.4 Discussion

It has long been known that chronic inflammation contributes to cancer. Although it is well established that NO$^{\bullet}$ creates reactive nitrogen species that form mutagenic DNA lesions (*9, 10, 35*), very few studies had investigated the possibility that NO$^{\bullet}$ induces homologous recombination events, even though homologous recombination causes sequence rearrangements that are known to contribute to cancer (*16, 18, 63*). Indeed, dozens of known carcinogens are recombinogens (*17, 19*). Furthermore, people who are born predisposed to spontaneously high levels of recombination are prone to cancer (*18, 19*). Thus, whether caused by exposure or by an inherited predisposition, homologous recombination events are a form of genetic instability that can lead to cancer.

Here, we have shown that ONOO^{-}, the product of the reaction of NO$^{\bullet}$ with superoxide, causes highly recombinogenic DNA lesions. In contrast, when NO$^{\bullet}$ reacts with oxygen to form N_2O_3, the resulting DNA lesions did not induce a detectable increase in interplasmid recombination. Therefore, NO$^{\bullet}$ becomes particularly recombinogenic specifically under conditions where it reacts with superoxide to create ONOO^{-}. Importantly, NO$^{\bullet}$ mediates both cell signaling and inflammation, but only under conditions of inflammation does the body simultaneously create high levels of both NO$^{\bullet}$ and superoxide. Therefore, it is likely that NO$^{\bullet}$ does not significantly induce recombination when produced at low concentrations during cell-signaling, whereas the results of the present studies show that following reaction with superoxide, it becomes highly recombinogenic and can thus contribute to tumorigenic sequence rearrangements.

We are primarily interested in the effects of inflammatory chemicals on genomic stability under physiologically relevant exposure conditions. In the case of NO$^{\bullet}$/O_2, it has

been estimated that the steady state concentration of NO$^{\bullet}$ at sites of inflammation is $\leq$ 1 μM

(*4, 45, 64, 65*). In the studies described here, cells were exposed to 1.3 μM NO$^{\bullet}$ (steady

state), which is a reasonable approximation of the level expected to be present within

inflamed tissues. In terms of ONOO^{-}, we elected not to expose cells to pure ONOO^{-}, because

if added to the media, the result would be a bolus exposure that does not accurately reflect

conditions during inflammation (the half life of ONOO^{-} is ~ 50 ms (*66*)). In contrast to

ONOO^{-}, SIN-1 makes it possible to achieve longer term exposures, since SIN-1 decomposes

to form ONOO^{-} at a fairly steady rate (at least during the first 90 minutes of exposure; (*67*)).

At 1 mM SIN-1, ONOO^{-} is produced at a rate of ~0.2 μM/sec (*67*), and in the studies

presented here, 0.5 mM SIN-1 significantly induced homologous recombination at the

integrated direct repeat substrate (note that this exposure was also relatively non-toxic; see

Figs. 1C-1D). Although the exact concentration of ONOO^{-} at sites of inflammation are not

yet known, the conditions used in these experiments are on par with the expected nM-μM

range of ONOO^{-} concentrations thought to be present during inflammation (*4*).

Comparisons of ONOO^{-}-induced recombination (reported here) and ONOO^{-}-induced

mutations (reported in the literature) suggest that ONOO^{-} is similarly potent as both a

mutagen and a recombinogen. For example, at a dose of SIN-1 that kills ~75% of TK6 cells,

mutations at HPRT and TK are induced by ~2 fold (*23, 43*), whereas in the studies presented

here, a dose of SIN-1 that kills ~25% of the cells induced a ~3 fold increase in recombination

a single locus in the chromosomally integrated recombination substrate. These results suggest

that ONOO^{-} may be as recombinogenic as it is mutagenic, although further studies are

required, since toxicity does not necessarily directly reflect the levels of damage. Comparing

the studies presented here to those in which exposed plasmid DNA has been used to study

mutagenesis (thus avoiding the complexities of assessing damage levels in cells), we again found similarities in the mutagenic and recombinogenic effects of ONOO⁻. Specifically, in previous studies using extrachromosomal plasmids, 2-2.5 mM ONOO⁻ induced a 4-9-fold increase in point mutations, respectively (*9, 33*). Here, we found that 0.02 mM ONOO⁻ induced a 2-fold increase in interplasmid recombination in mouse ES cells, and that 2 mM ONOO⁻ induced a 4-fold increase in interplasmid recombination in monkey kidney cells. Although further studies are necessary to control for cell-type effects, these comparisons suggest that the mutagenic potential of ONOO⁻ is on par with its recombinogenic potential.

Although double strand breaks are thought to be critical for inducing homologous recombination, ONOO⁻ creates oxidative base lesions, abasic sites, and single strand breaks (*4*). Consistent with previous studies, in the topoisomer analysis performed here, we were unable to detect the presence of any ONOO⁻ or SIN-1-induced double strand breaks at doses that were highly recombinogenic. However, we have shown that Fpg can introduce double strand breaks in ONOO⁻-treated plasmids (Figure 2.5A) and in SIN-1 treated plasmids (data not shown). In contrast, treatment with DNA glycosylases did not induce double strand breaks in NO˙/O₂-treated plasmid (data not shown). Thus, following transfection, closely opposed lesions on opposite strands could potentially be converted into double strand breaks by DNA glycosylases, as was originally demonstrated for oxidative damage *in vitro*, and has been recently demonstrated for radiation damage in mammalian cells (*36, 40*). Alternatively, abasic sites and single strand breaks may be converted into double strand breaks during replication, which is likely to contribute to SIN-1 induced recombination at the integrated direct repeat. However, given that the plasmids used in these experiments do not have an

origin of replication, the most likely cause of interplasmid recombination induced by ONOO⁻ are double strand breaks created by enzymatic processing in the recipient cells.

The results of these studies show that conditions that favored formation of N_2O_3 did not induce recombination at any of the recombination substrates, although there was an small but significant increase in sister chromatid exchanges in cells exposed to $NO^•/O_2$. It is certainly possible that some of the lesions induced by N_2O_3 induce SCEs that were not detected using the engineered substrates. However, due to the inevitable presence of superoxide in normal cells, it is also possible that $NO^•$ reacts with superoxide to form ONOO in these cells. Nevertheless, it is clearly the case that ONOO⁻ is significantly more recombinogenic than N_2O_3 in all of the recombination assays, and N_2O_3 lesions failed to induce any detectable increase in interplasmid recombination. These results were somewhat unexpected, given that lesions created by both N_2O_3 and ONOO⁻ are substrates for the base excision repair pathway, and that base excision repair intermediates have been shown to be highly recombinogenic in mammalian cells (e.g., (*37, 38*)).

One major difference between ONOO⁻ and N_2O_3 is that ONOO⁻ can directly react with deoxyribose to induce single strand breaks. Therefore, it is possible that chemically-induced single strand breaks are converted into recombinogenic double strand breaks when another single strand break is enzymatically introduced on the opposite strand by base excision repair enzyme(s). Given that glycosylases rapidly find their targets ($t_{1/2}$ = minutes for the Aag DNA glycosylase, for example (*68*)), it is not difficult to imagine that a glycosylase might find its target before a chemically-induced strand break on the opposite strand has been repaired. Furthermore, different types of glycosylases act on ONOO⁻-induced guanine oxidation and nitration products (*e.g.*, 8-nitroguanine, 8-oxoguanine and its

70

secondary oxidative products) compared to N_2O_3-induced nucleobase deamination products (uracil, xanthine and hypoxanthine) (*49, 50*) and Dong *et al.*, manuscript in preparation). The vast majority of ONOO⁻-induced lesions are repaired by bifunctional glycosylases (Fpg and Nth, and their mammalian homologs), and these enzymes have the ability to both to remove the base lesions and to cleave the sugar-phosphate backbone (*e.g.*, via their associated lyase activity) (*69, 70*). In contrast, those glycosylases known to act on N_2O_3-induced lesions are monofunctional (*e.g.*, Ung and AlkA, and their mammalian homologs; (*50, 71*) and Dong *et al.*, manuscript in preparation), so these enzymes remove the base lesions but cannot create single strand breaks. Taken together, the observation that ONOO⁻ is more recombinogenic than N_2O_3 may be because bifunctional glycosylases can create recombinogenic double strand breaks by cleaving across from a pre-existing chemically-induced strand break and by enzymatic cleavage at two closely opposed base lesions.

Although deaminated bases are known to be mutagenic (*27*), we did not find that *EGFP* expression was inhibited by exposure to NO˙/O_2. At the highest dose of NO˙/O_2, we estimated that there are ~2 lesions/plasmid. Given that only ~15% of the plasmid sequences code for *EGFP*, it is likely that copies of the plasmid that harbor undamaged coding sequences are responsible for the observed *EGFP* expression. Unexpectedly, we found that an even lower number of ONOO⁻ lesions/plasmid inhibited *EGFP* expression. Following the same logic, it seems unlikely that *EGFP* expression is suppressed by mutagenic damage to coding sequences. Given that linearized plasmid DNA is highly vulnerable to degradation by exonucleases (*60, 61*), one possibility is that *EGFP* expression is suppressed due to plasmid loss, rather than to mutations. Thus, the observations that suppression of *EGFP* and induction of recombination go hand in hand for ONOO⁻ and SIN-1, and that NO˙/O_2 neither suppresses

EGFP nor induces recombination, are consistent with a model wherein glycosylases introduce double strand breaks that both increase the vulnerability of the DNA to degradation, and induce homologous recombination.

One interesting result of these studies was that the relative recombinogenicity of ONOO⁻ and N_2O_3 is very similar in mouse embryonic cells and monkey kidney cells, regardless of whether the DNA was exposed prior to transfection, or cells were exposed after transfection. However, there are also some subtle differences. For example, there is a significant difference in the doses of ONOO⁻ required to achieve maximal induction of recombination in embryonic cells versus kidney cells (Figure 2.3). Although there are many possible explanations for this observation, one possibility is that embryonic stem cells have relatively higher levels of DNA glycosylase that induce double strand breaks. If this were the case, then glycosylase-induced double strand breaks could potentially be induced at lower levels of adducts. Although ONOO⁻-induced recombination is a general feature of mammalian cells, these results show that the relative sensitivity of cells to ONOO⁻-induced recombination clearly is variable among different cell types.

The magnitude of ONOO⁻-induced recombination is significant. Mammalian cells subjected to conditions that favor formation of ONOO⁻ are susceptible to ~3-fold increase in recombination at the integrated ΔGF recombination substrate, and to ~6-fold increase in SCEs. For comparison, it has been reported that phenytoin, an oxidizing agent, induces a 2- to 3-fold increase in recombination at an analogous direct repeat substrate in CHO cells (*72*). Indeed, in the SCE literature, most of the agents that induce such high levels of SCEs are chemicals that most of us will never experience (*73*). Given that inflammation is unavoidable in the lifetime of most people, and that ONOO⁻ is highly recombinogenic, endogenously

produced $ONOO^-$ may be one of the most significant recombinogens to which most people are exposed. As such, $ONOO^-$-induced recombination may be an important risk factor for tumorigenesis.

2.5 References

1. Moncada, S., and Higgs, E. A. (1991) *Eur. J. Clin. Invest. 21*, 361-374.
2. Kerwin, J. F., Jr., Lancaster, J. R., Jr., and Feldman, P. L. (1995) *J. Med. Chem. 38*, 4343-462.
3. Wiseman, H., and Halliwell, B. (1996) *Biochem. J. 313*, 17-29.
4. Dedon, P. C., and Tannenbaum, S. R. (2004) *Arch. Biochem. Biophys. 423*, 12-22.
5. Burney, S., Tamir, S., Gal, A., and Tannenbaum, S. R. (1997) *Nitric Oxide 1*, 130-144.
6. Marletta, M. A., and Spiering, M. M. (2003) *J Nutr 133*, 1431S-3S.
7. Ohshima, H., and Bartsch, H. (1994) *Mutat. Res. 305*, 253-264.
8. Kuper, H., Adami, H. O., and Trichopoulos, D. (2000) *J. Intern. Med. 248*, 171-183.
9. Juedes, M. J., and Wogan, G. N. (1996) *Mutat. Res. 349*, 51-61.
10. Routledge, M. N. (2000) *Mutat. Res. 450*, 95-105.
11. Tamir, S., Burney, S., and Tannenbaum, S. R. (1996) *Chem. Res. Toxicol. 9*, 821-827.
12. West, S. C. (2003) *Nat. Rev. Mol. Cell. Biol. 4*, 435-445.
13. McGlynn, P., and Lloyd, R. G. (2002) *Nat. Rev. Mol. Cell. Biol. 3*, 859-870.
14. Paques, F., and Haber, J. E. (1999) *Microbiol. Mol. Biol. Rev. 63*, 349-404.
15. Morley, A. A., Grist, S. A., Turner, D. R., Kutlaca, A., and Bennett, G. (1990) *Cancer Res. 50*, 4584-4587.
16. Shao, C., Deng, L., Henegariu, O., Liang, L., Raikwar, N., Sahota, A., Stambrook, P. J., and Tischfield, J. A. (1999) *Proc. Natl. Acad. Sci. USA 96*, 9230-9235.
17. Aubrecht, J., Rugo, R., and Schiestl, R. H. (1995) *Carcinogenesis 16*, 2841-2846.
18. Thompson, L. H., and Schild, D. (2002) *Mutat. Res. 509*, 49-78.
19. Bishop, A. J., and Schiestl, R. H. (2001) *Biochim. Biophys. Acta. 1471*, M109-M121.
20. Kang, M. H., Genser, D., and Elmadfa, I. (1997) *Mutat. Res. 381*, 141-148.
21. Donovan, P. J., Smith, G. T., Lawlor, T. E., Cifone, M. A., Murli, H., and Keefer, L. K. (1997) *Nitric Oxide 1*, 158-166.
22. Tanaka, R. (1997) *J. Toxicol. Sci. 22*, 199-205.
23. Li, C. Q., Trudel, L. J., and Wogan, G. N. (2002) *Proc. Natl. Acad. Sci. USA 99*, 10364-10369.
24. Dong, M., Wang, C., Deen, W. M., and Dedon, P. C. (2003) *Chem. Res. Toxicol. 16*, 1044-1055.
25. Caulfield, J. L., Wishnok, J. S., and Tannenbaum, S. R. (1998) *J. Biol. Chem. 273*, 12689-12695.
26. Nguyen, T., Brunson, D., Crespi, C. L., Penman, B. W., Wishnok, J. S., and Tannenbaum, S. R. (1992) *Proc. Natl. Acad. Sci. USA 89*, 3030-3034.
27. Burney, S., Caulfield, J. L., Niles, J. C., Wishnok, J. S., and Tannenbaum, S. R. (1999) *Mutation Res. 424*, 37-49.
28. deRojas-Walker, T., Tamir, S., Ji, H., Wishnok, J. S., and Tannenbaum, S. R. (1995) *Chem Res Toxicol 8*, 473-7.
29. Kennedy, L. J., Moore, K., Jr., Caulfield, J. L., Tannenbaum, S. R., and Dedon, P. C. (1997) *Chem. Res. Toxicol. 10*, 386-392.
30. Yermilov, V., Rubio, J., and Ohshima, H. (1995) *FEBS Lett. 376*, 207-210.
31. Lee, J. C., Niles, J. S., Wishnok, J. S., and Tannenbaum, S. R. (2002) *Chem. Res. Toxicol. 15*, 7-14.

32. Salgo, M. G., Stone, K., Squadrito, G. L., Battista, J. R., and Pryor, W. A. (1995) *Biochem. Biophys. Res. Commun. 210*, 1025-1030.

33. Tretyakova, N. Y., Burney, S., Pamir, B., Wishnok, J. S., Dedon, P. C., Wogan, G. N., and Tannenbaum, S. R. (2000) *Mutat. Res. 447*, 287-303.

34. Burney, S., Niles, J. C., Dedon, P. C., and Tannenbaum, S. R. (1999) *Chem. Res. Toxicol. 12*, 513-520.

35. Tamir, S., deRojas-Walker, T., Wishnok, J. S., and Tannenbaum, S. R. (1996) *Methods Enzymol. 269*, 230-243.

36. Harrison, L., Hatahet, Z., Purmal, A. A., and Wallace, S. S. (1998) *Nucleic Acids Res 26*, 932-41.

37. soCoquerelle, T., Dosch, J., and Kaina, B. (1995) *Mutat. Res. 336*, 9-17.

38. Sobol, R. W., Kartalou, M., Almeida, K. H., Joyce, D. F., Engelward, B. P., Horton, J. K., Prasad, R., Samson, L. D., and Wilson, S. H. (2003) *J. Biol. Chem. 278*, 39951-39959.

39. Spek, E. J., Vuong, L. N., Matsuguchi, T., Marinus, M. G., and Engelward, B. P. (2002) *J. Bacteriol. 184*, 3501-3507.

40. Yang, N., Galick, H., and Wallace, S. S. (2004) *DNA Repair (Amst) 3*, 1323-1334.

41. Wang, C., and Deen, W. M. (2003) *Ann. Biomed. Eng. 31*, 65-79.

42. Hendricks, C. A., Almeida, K. H., Stitt, M. S., Jonnalagadda, V. S., Rugo, R. E., Kerrison, G. F., and Engelward, B. P. (2003) *Proc. Natl. Acad. Sci. USA 100*, 6325-6330.

43. Li, C. Q., Trudel, L. J., and Wogan, G. N. (2002) *Chem. Res. Toxicol. 15*, 527-535.

44. Miwa, M., Stuehr, D. J., Marletta, M. A., Wishnok, J. S., and Tannenbaum, S. R. (1987) *Carcinogenesis 8*, 955-8.

45. Lewis, R. S., Tamir, S., Tannenbaum, S. R., and Deen, W. M. (1995) *J. Biol. Chem. 270*, 29350-29355.

46. Okabe, M., Ikawa, M., Kominami, K., Nakanishi, T., and Nishimune, Y. (1997) *FEBS Lett 407*, 313-9.

47. Pryor, W. A., Cueto, R., Jin, X., Koppenol, W. H., Ngu-Schwemlein, M., Squadrito, G. L., Uppu, P. L., and Uppu, R. M. (1995) *Free Radic. Biol. Med. 18*, 75-83.

48. Lindahl, T., and Andersson, A. (1972) *Biochemistry 11*, 3618-3623.

49. Tretyakova, N. Y., Wishnok, J. S., and Tannenbaum, S. R. (2000) *Chem. Res. Toxicol. 13*, 658-664.

50. Wuenschell, G. E., O'Connor, T. R., and Termini, J. (2003) *Biochemistry 42*, 3608-3616.

51. Arai, K., Morishita, K., Shinmura, K., Kohno, T., Kim, S. R., Nohmi, T., Taniwaki, M., Ohwada, S., and Yokota, J. (1997) *Oncogene 14*, 2857-2861.

52. Jonnalagadda, V., and Engelward, B. P. (in preparation).

53. Bill, C. A., and Nickoloff, J. A. (2001) *Mutat. Res. 487*, 41-50.

54. Sonoda, E., Sasaki, M. S., Morrison, C., Yamaguchi-Iwai, Y., Takata, M., and Takeda, S. (1999) *Mol. Cell. Biol. 19*, 5166-5169.

55. Mudgett, J. S., and Taylor, W. D. (1990) *Mol. Cell. Biol. 10*, 37-46.

56. Slebos, R. J., and Taylor, J. A. (2001) *Biochem. Biophys. Res. Commun. 281*, 212-219.

57. David, S. S., and Williams, S. D. (1998) *Chem Rev 98*, 1221-1262.

58. Taghian, D. G., and Nickoloff, J. A. (1997) *Mol. Cell. Biol. 17*, 6386-6393.

59.   Sargent, R. G., Brenneman, M. A., and Wilson, J. H. (1997) *Mol. Cell. Biol. 17*, 267-277.

60.   Lechardeur, D., Sohn, K. J., Haardt, M., Joshi, P. B., Monck, M., Graham, R. W., Beatty, B., Squire, J., O'Brodovich, H., and Lukacs, G. L. (1999) *Gene Ther 6*, 482-97.

61.   Tanswell, A. K., Staub, O., Iles, R., Belcastro, R., Cabacungan, J., Sedlackova, L., Steer, B., Wen, Y., Hu, J., and O'Brodovich, H. (1998) *Am J Physiol 275*, L452-60.

62.   Parsons, J. L., and Elder, R. H. (2003) *Mutat Res 531*, 165-75.

63.   Bishop, A. J., and Schiestl, R. H. (2003) *Exp. Mol. Pathol. 74*, 94-105.

64.   Stuehr, D. J., and Marletta, M. A. (1987) *J. Immunol. 139*, 518-525.

65.   Stuehr, D. J., and Marletta, M. A. (1987) *Cancer Res 47*, 5590-4.

66.   Lymar, S. V., and Hurst, J. K. (1996) *Chem Res Toxicol 9*, 845-50.

67.   Doulias, P. T., Barbouti, A., Galaris, D., and Ischiropoulos, H. (2001) *Free Radic Biol Med 30*, 679-85.

68.   Smith, S. A., and Engelward, B. P. (2000) *Nucleic Acids Res. 28*, 3294-3300.

69.   Wilson, D. M., 3rd, Sofinowski, T. M., and McNeill, D. R. (2003) *Front Biosci. 8*, 963-981.

70.   Fromme, J. C., and Verdine, G. L. (2004) *Adv Protein Chem 69*, 1-41.

71.   Wilson III, D. M., Engelward, B. P., and Samson, L. (1998) in *DNA Damage and Repair: Biochemistry, Genetics, and Cell Biology* (Nickoloff, J. A., and Hoekstra, M. F., Eds.) pp 29-64, Humana Press Inc., Totowa, NJ.

72.   Winn, L. M., Kim, P. M., and Nickoloff, J. A. (2003) *J. Pharmacol. Exp. Ther. 2*, 523-527.

73.   Sandberg, A. A. (1982) *Sister Chromatid Exchange*, Vol. II, Alan R. Liss, Inc., New York, NY.
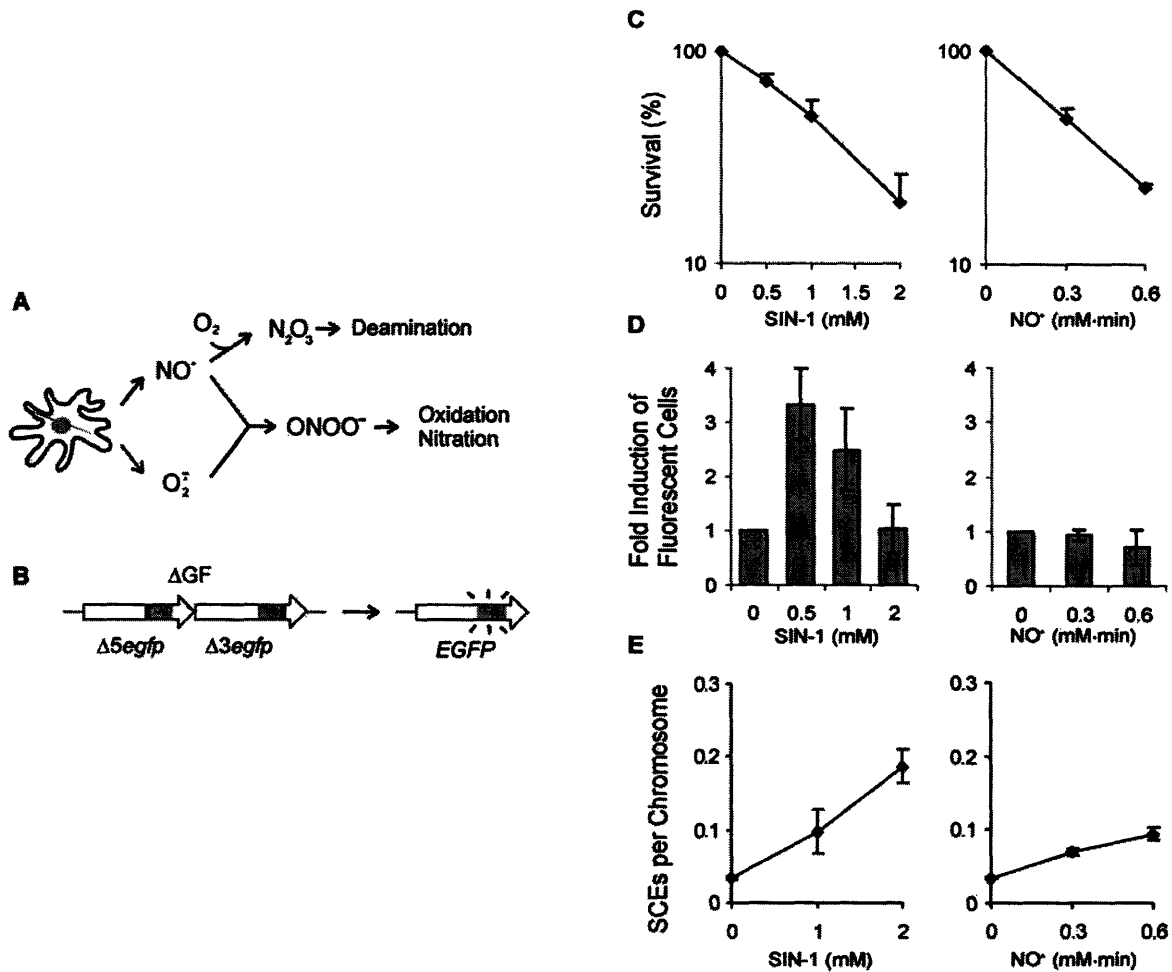
Figure 2.1. NO·/O₂- or SIN-1-induced toxicity and homologous recombination in mouse embryonic stem cells. (A) Schematic depicting the major reactions of NO· secreted by activated macrophages. (B) The chromosomal direct repeat substrate. Expression cassettes are represented by large arrows with emphasis on the coding sequences (gray) and deleted regions (black). Δ5egfp and Δ3egfp expression are driven by the chicken beta-actin promoter and cytomegalovirus enhancer. Homologous recombination can lead to the restoration of full-length EGFP coding sequence and fluorescence expression. (C) Relative toxicity induced by exposure to the indicated doses of NO· delivered in the gas delivery chamber or SIN-1 as determined by colony forming assay. For NO· exposure, Ar gas was used as a negative control and relative survival was normalized accordingly. Values are the mean of three samples ± SEM. (D) Relative levels of damage-induced recombination at a chromosomally integrated direct repeat. Fold induction of fluorescent recombinant cells is shown. Following exposure to NO·/O₂ or SIN-1 (72 h), cells were analyzed by flow cytometry. Values are the mean of three samples ± SEM. (E) SCEs induced by SIN-1 and NO·/O₂. SCEs per chromosome were detected by analyzing 20 metaphase chromosomes at each dose. The results shown are mean ± SD for duplicate experiments.
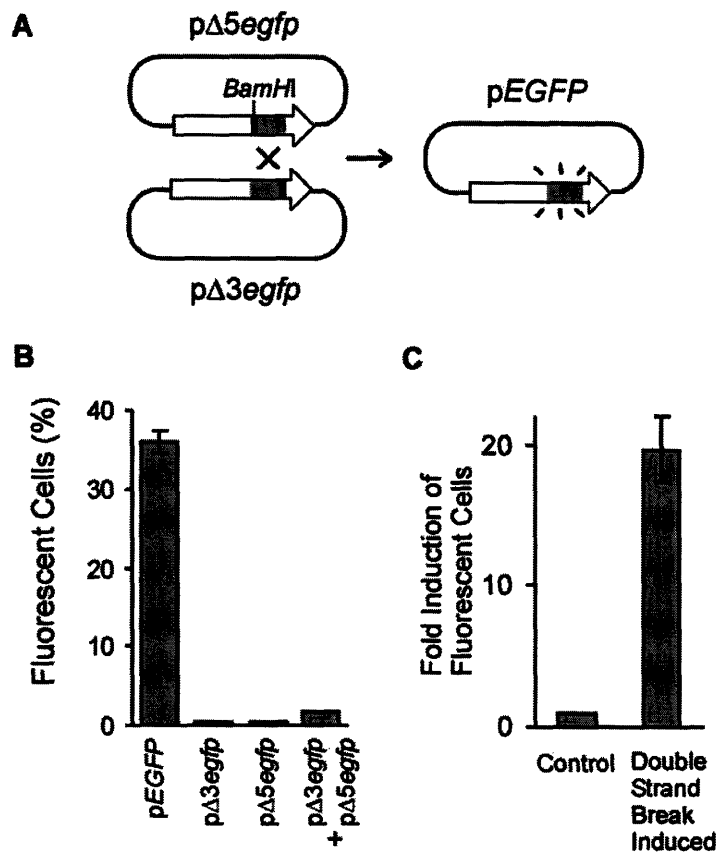
Figure 2.2. Inter-plasmid recombination assay in mouse ES cells and in COS 7L Monkey Kidney Cells. (A) Schematic depicting the interplasmid recombination assay. Expression cassettes are as described in Figure 2. 1B. (B) Percent fluorescent cells following transfection with the indicated plasmids. ES cells were analyzed by flow cytometry 48 hours after transfection. The p*EGFP* plasmid was used as a positive control. (C) Double strand break-induced interplasmid recombination. The pΔ5*egfp* plasmid was linearized at the *BamH*I site indicated in part A. Fold induction relative to control plasmid (uncut pΔ5*egfp*) is shown.

Extracellular DNA Damage
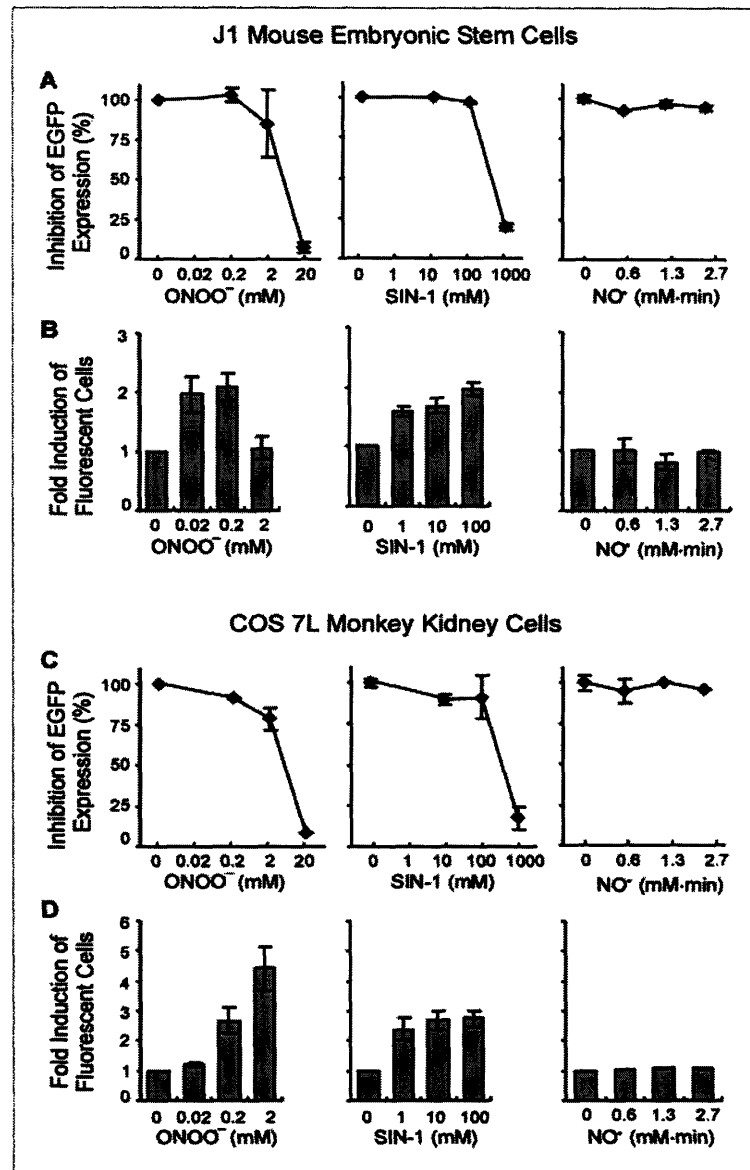
Figure 2.3. Interplasmid recombination induced by ONOO⁻, SIN-1 or NO˙/O_2-exposed plasmid DNA in Mouse ES cells (A and B) and in COS 7L Monkey Kidney Cells (C and D). (A, C) Inhibition of *EGFP* expression caused by p*EGFP* exposure to ONOO⁻, SIN-1 or NO˙. The frequency of fluorescent cells was determined by flow cytometry 48 hours after transfection. Values indicate the averages of two independent experiments (+ SEM). (B, D) Interplasmid recombination induced by ONOO⁻, SIN-1 or NO˙/O_2-exposed pΔ5*egfp*. The pΔ5*egfp* plasmid was exposed to the indicated doses, mixed with undamaged pΔ3*egfp*, and transfected into cells. After 48 h, the frequency of fluorescent cells were analyzed by flow cytometry. In all the experiments, frequencies were normalized according to the lipofection efficiency. Values indicate two independent experiments, each performed with duplicate samples (+ SEM).

Intracellular DNA Damage
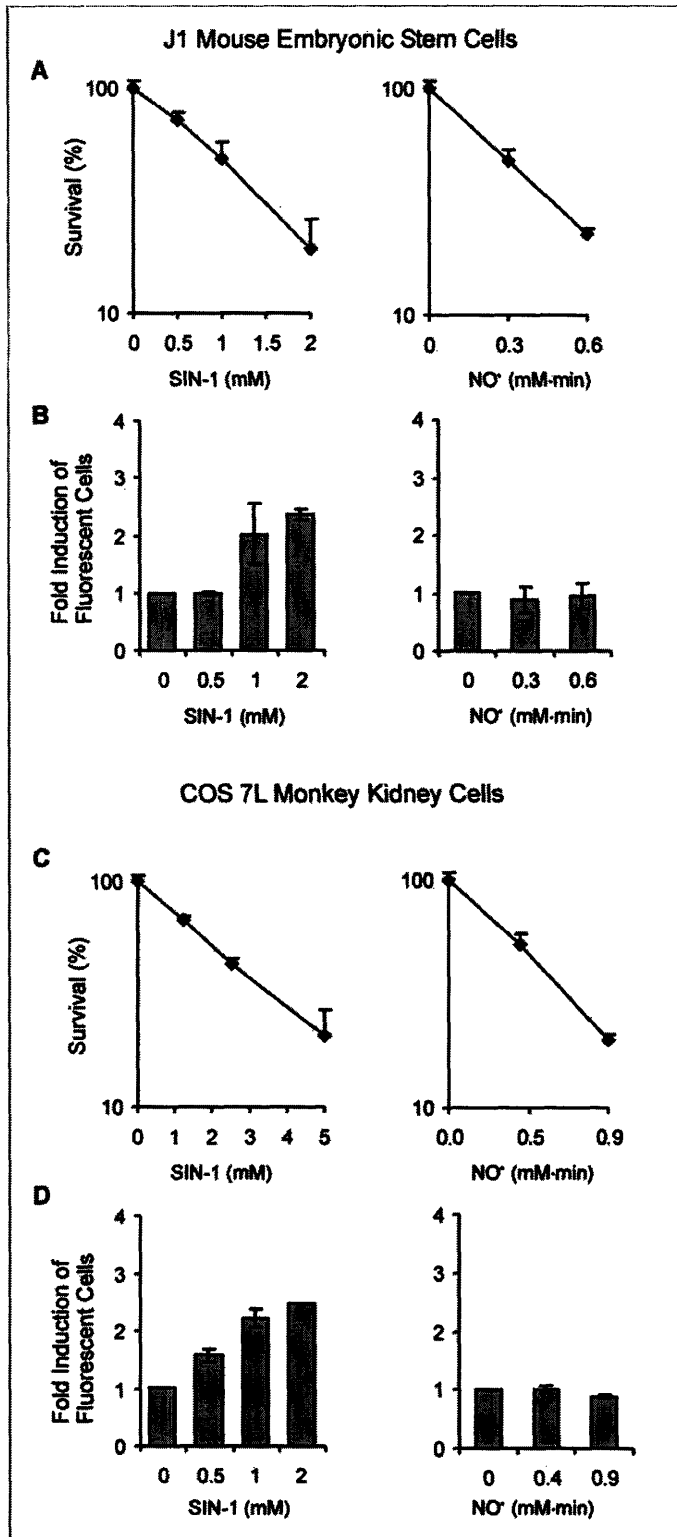
Figure 2.4. Interplasmid homologous recombination induced by exposure of cells to SIN-1 or NO•/O₂. Results for Mouse ES cells (A and B) and COS 7L Monkey Kidney Cells (C and D) are shown. (A, C) Relative toxicity induced by exposure to the indicated doses of NO• or SIN-1 as determined by colony forming assay. For NO•/O₂ exposure, Ar gas was used as a negative control and relative survival was normalized accordingly. Values are the mean of three samples ± SEM. (B, D) Interplasmid recombination induced by SIN-1 or NO•/O₂ in cells that carry pΔ5egfp and pΔ3egfp. Cells were co-lipofected with pΔ5egfp and pΔ3egfp 24 h prior to exposure to the indicated doses of SIN-1 or NO•. After 48 hours, the frequency of fluorescent cells was quantified by flow cytometry. In all the experiments, frequencies were normalized according to the lipofection efficiency. Values indicate two independent experiments, each performed with duplicate samples (± SEM).
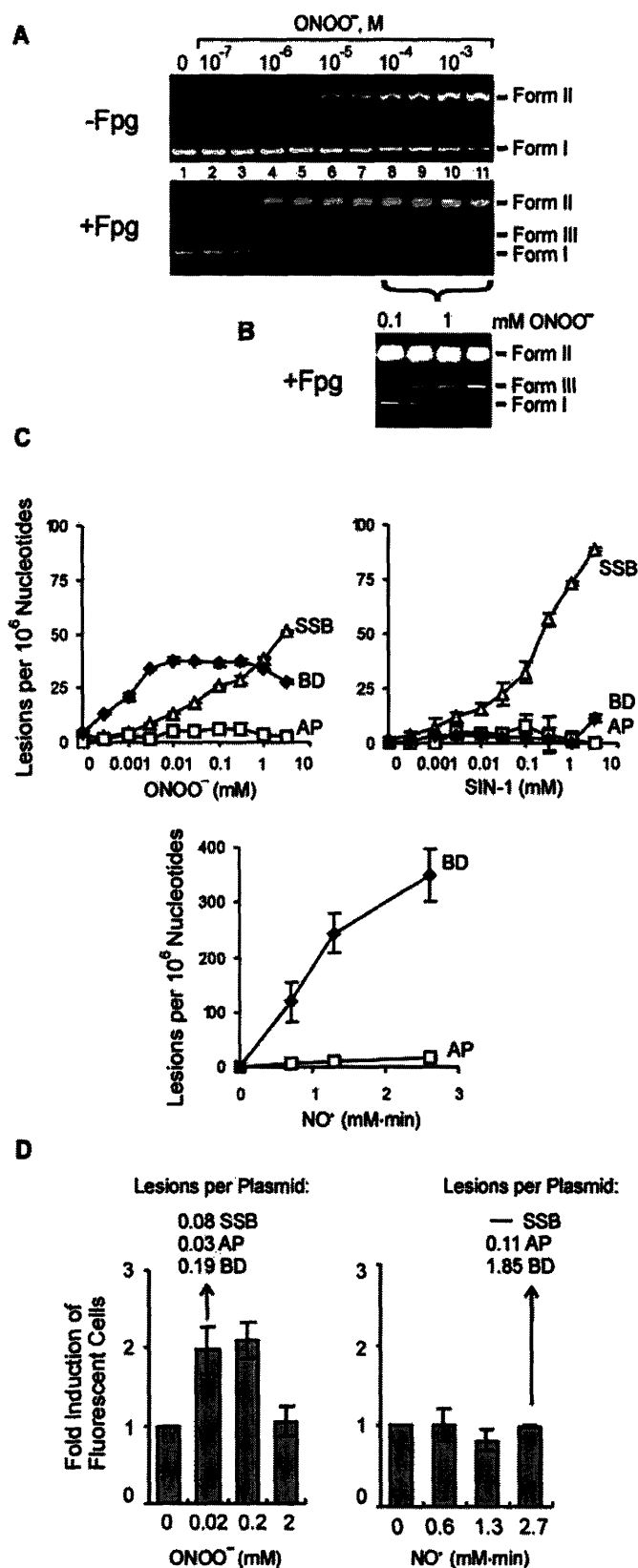
Figure 2.5. Quantification of the major classes of DNA lesions induced by ONOO⁻, SIN-1 and NO˙/O₂ *in vitro* and comparison of lesion levels to recombination levels. (A) Example of gel electrophoretic analysis in the plasmid nicking assay used to quantify base and deoxyribose lesions. For technical reasons, the plasmid nicking assays were performed with pSP189; this plasmid is similar in size to pΔ5egfp (both are ~ 5kb). Upper panel: direct single strand breaks produced in plasmid pSP189 by ONOO- cause a dose-dependent shift from Form I (supercoiled) to Form II (nicked, closed-circular plasmid). Lower panel: ONOO⁻-treated plasmid DNA was reacted with Fpg to convert base lesions and abasic sites to strand breaks. Fpg incubation causes the appearance of linear plasmid (Form III). ONOO⁻ concentrations are in half-log steps. (B) The brightness/contrast of lanes 8-11 in the bottom panel of 'A' were adjusted to emphasize the appearance of Form III DNA. (C) Quantity of base damage (BD), abasic sites (AP) and single strand breaks (SSB) per 10^6 nucleotides as determined by plasmid topoisomer analysis. Levels of base adducts were estimated by converting base lesions to single strand breaks using purified Fpg glycosylase for ONOO⁻ and SIN-1, or purified Ung and AlkA glycosylases followed by putrescine cleavage of abasic sites for NO˙/O₂). Mean ± SD for n = 4. (D) A comparison of the quality and quantity of lesions created by the lowest recombinogenic dose of ONOO⁻ and the highest dose of NO˙/O₂. Graphs are identical to those in Figure 2. 3B and are shown here to facilitate comparison to damage levels.

81

Mitomycin C (MMC) is a well studied, highly recombinogenic DNA damaging agent. Experiments conducted to measure $NO^{\bullet}/O_2$- or SIN-1-induced toxicity and homologous recombination in mouse embryonic stem cells (Figure 2.1) were conducted in parallel with MMC as a positive control. Although this data was not included in the manuscript (Chapter 2), it provides valuable information about the relative recombinogenicity of $NO^{\bullet}/O_2$- or SIN-1 when compared to MMC. As evident from Supp. Figure 2.1, MMC is significantly more recombinogenic than both $NO^{\bullet}/O_2$- or SIN-1 at equitoxic doses.
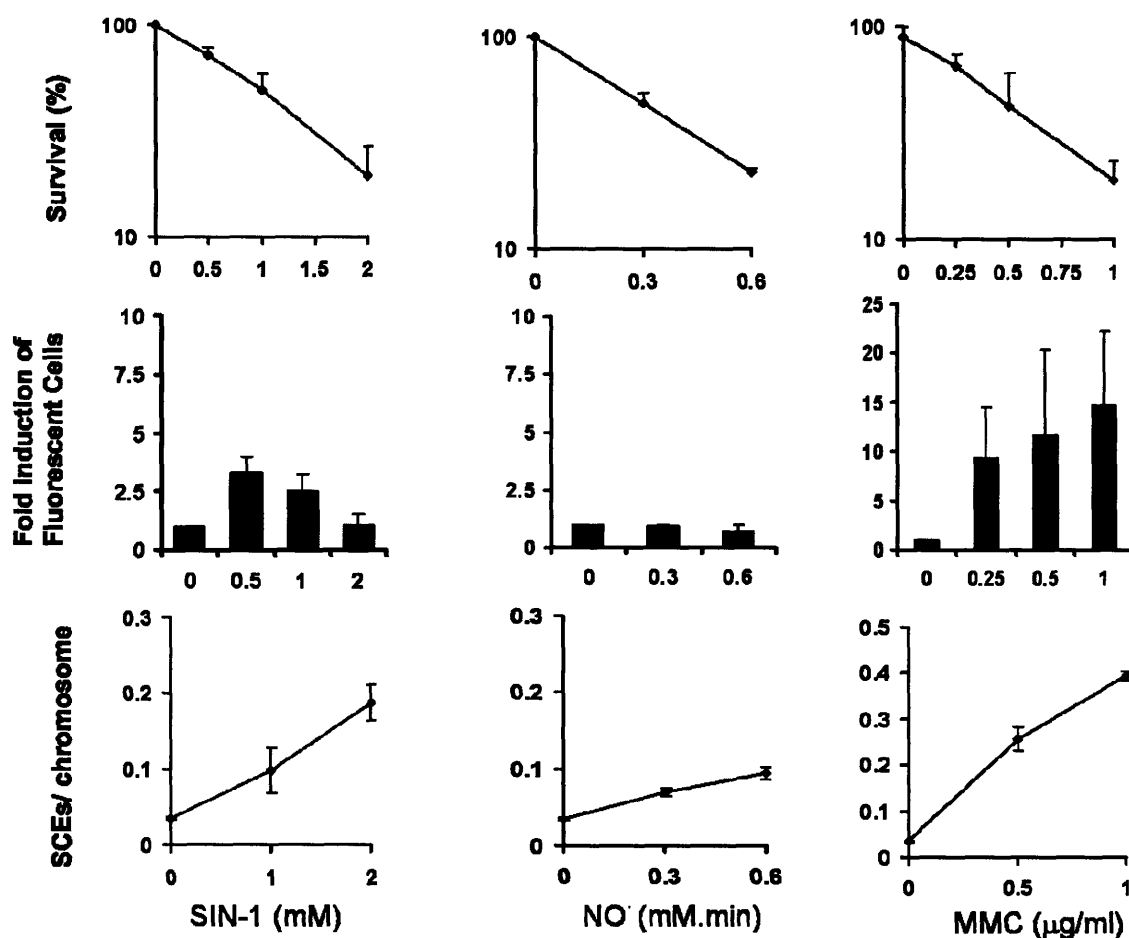
Supp. Figure 2.1. $NO^{\bullet}/O_2$- or SIN-1-induced homologous recombination in comparison to Mitomycin C (MMC)-induced homologous recombination at equitoxic doses in mouse embryonic stem cells. First row shows the survival curves, second row shows damage induced recombination at a chromosomally integrated direct repeat, and the third row shows damage induced SCEs.

CHAPTER 3

Conclusion and Future Directions

CHAPTER 3

Conclusion and Future Directions

Homologous recombination events are an important source of mutations that are likely to be the underlying cause of one or more mutations in every tumor. Nevertheless, we do not yet have an extensive understanding of what causes recombination in people. Although exposure to toxic levels of most carcinogens is recombinogenic, most humans are is not likely to be exposed to high concentrations of these agents. In contrast, inflammation is a widespread pathophysiological response in humans, and it is known to contribute to cancer. Indeed, during inflammation, the high levels of reactive oxygen and nitrogen species that are produced by activated macrophages are highly toxic to neighboring normal cells. The major physiologically relevant reactive nitrogen species are N_2O_3 and $ONOO^-$, which are formed when $NO^\bullet$ reacts with either oxygen or superoxide, respectively. This study has shown, by sister chromatid exchange, chromosomal direct repeat, and interplasmid recombination assays, that $ONOO^-$ is a strong recombinogen, whereas N_2O_3 is relatively weakly recombinogenic. Consequently, $NO^\bullet$ has the potential of being highly recombinogenic specifically under conditions when superoxide is present, as is the case during inflammation. These results suggest that $ONOO^-$-induced recombination may contribute to tumorigenic sequence rearrangements and may play an important role in inflammation-induced cancers.

Double strand breaks (DSBs) are thought to be critical for inducing homologous recombination. In our studies, we observed that $ONOO^-$ created DNA damage induced significant levels of inter-plasmid recombination, whereas no significant levels of inter-plasmid recombination was detected with N_2O_3. Interestingly, neither $ONOO^-$ nor N_2O_3 directly induce double strand breaks. Exposure of plasmid DNA to $ONOO^-$ created

predominantly base oxidation products and single strand breaks, while exposure of plasmid DNA to N_2O_3 created base deamination products and abasic sites. Together these results suggest that, on a per lesion basis, $ONOO^-$-induced oxidative base lesions and single strand breaks are significantly more recombinogenic than N_2O_3-induced base deamination products. Since DSBs are thought to induce recombination, we hypothesize that $ONOO^-$ can induce formation of double strand breaks via (1) two closely opposed chemically induced single strand breaks; (2) a closely opposed chemically induced single strand break and an oxidative base lesion; (3) two closely opposed oxidative base lesions (Figure 3.1). These closely opposed oxidative base lesions could potentially be converted into a single strand break by DNA glycosylases, leading to the formation of recombinogenic DNA double strand breaks.

These possible mechanisms proposed for the differential recombinogenicity observed for $ONOO^-$ and N_2O_3 require further testing. For example, our studies suggest that base oxidation products are more recombinogenic than base deamination products on a per lesion basis. Given that both of these lesions are repaired by base excision repair (BER) pathway, indeed, both types of lesions could enzymatically be converted into strand breaks as intermediates of BER pathway. This raises an interesting question; what is the key factor that accounts for the differential recombinogenicity of oxidation and deamination base products? One major difference is the type of glycosylases that acts on $ONOO^-$-induced base oxidation products (*e.g.*, 8-nitroguanine, 8-oxoguanine and its secondary oxidative products) and the type of glycosylases that acts on N_2O_3-induced base deamination products (uracil, xanthine and hypoxanthine) (*1, 2*) and Dong *et al.*, manuscript in preparation). The vast majority of $ONOO^-$-induced lesions are repaired by bifunctional glycosylases (Fpg and Nth, and their mammalian homologs). These enzymes have the ability to both to remove the base lesions

and simultaneously cleave the sugar-phosphate backbone (*e.g.*, via their associated lyase activity) (Figure 3.2) (*3*). In contrast, glycosylases known to act on N_2O_3-induced lesions are monofunctional (*e.g.*, Ung and AlkA, and their mammalian homologs; (*1, 4*) and Dong *et al.*, manuscript in preparation), so these enzymes remove the base lesions, but cannot create single strand breaks (Figure 3.2). In the case of monofunctional glycosylases, generation of nicks requires an additional step catalyzed by AP endonucleases. This difference in the action of monofunctional versus bifunctional glycosylases could be the key factor leading to the difference in recombinogenicity observed with $ONOO^-$-induced base oxidation products versus N_2O_3-induced base deamination products. This hypothesis needs further testing and one approach to test it may be possible by studying the effects of over-expression as well as knocking down of some BER pathway enzymes that are known to play a role in the repair of base oxidation products and base deamination products.

Replication is believed to play an important role in converting various types of template damage into recombinogenic double strand breaks (Figure1.1). Nevertheless, the plasmids that were used in the extrachromosomal recombination studies (Figure 2.3 and Figure 2.4) lacked mammalian origin of replication. To address the possible effects of replication on the recombinogenicity of $ONOO^-$ or N_2O_3-induced DNA lesions, we repeated the set of experiments presented in Figure 2.3 and Figure 2.4 with plasmids carrying SV40 origin of replication. Although SV40 origin of replication replicates efficiently in SV40 transformed monkey kidney cells (COS 7L), we did not detect any effect of replication on inter-plasmid recombination induced by $ONOO^-$ or N_2O_3-induced DNA lesions (data not shown). These results are controversial to what is traditionally believed in regarding the role

of replication in generation of recombinogenic DNA lesions, and suggests that the extrachromosomal plasmid assay is not an ideal system to study the effects of replication.

Another interesting question that needs to be further addressed is the physiological relevance of our findings. In this study, to dissect the effects of ONOO⁻ and N_2O_3 on homologous recombination in mammalian cells, we exposed cells and plasmids to each of these chemicals. We have shown that ONOO⁻ is significantly recombinogenic, and based on this result we have hypothesized that during inflammation ONOO⁻ induces recombination. The relevance of our findings to inflammation needs to be supported with further studies. One way to mimic inflammation in cell culture experiments would be to perform co-culture experiments by using macrophages. Macrophages can be cultured *in vitro* and be stimulated to secrete reactive oxygen and nitrogen species by using interferon-γ (*5*). Thus, by co-culturing mammalian cells with stimulated macrophages, we can ask if during conditions that mimic inflammation, homologous recombination is induced in mammalian cells. Lastly, the physiological relevance of our findings can be tested by using various mouse models for inflammation and analyzing SCEs in the inflamed tissues, or by using the FYDR mouse that was developed to detect recombination (*6*).

Given that inflammation is unavoidable in the lifetime of most people, and that ONOO⁻ is highly recombinogenic, this study suggests that endogenously produced ONOO⁻ may be one of the most significant recombinogens to which most people are exposed. Thus, ONOO⁻-induced recombination may be an important risk factor for tumorigenesis in an inflamed tissue.

References

1. Wuenschell, G. E., O'Connor, T. R., and Termini, J. (2003) *Biochemistry 42*, 3608-3616.

2. Tretyakova, N. Y., Wishnok, J. S., and Tannenbaum, S. R. (2000) *Chem. Res. Toxicol. 13*, 658-664.

3. Wilson, D. M., 3rd, Sofinowski, T. M., and McNeill, D. R. (2003) *Front Biosci. 8*, 963-981.

4. Wilson III, D. M., Engelward, B. P., and Samson, L. (1998) in *DNA Damage and Repair: Biochemistry, Genetics, and Cell Biology* (Nickoloff, J. A., and Hoekstra, M. F., Eds.) pp 29-64, Humana Press Inc., Totowa, NJ.

5. Zhuang, J. C., Lin, C., Lin, D., and Wogan, G. N. (1998) *Proc Natl Acad Sci U S A 95*, 8286-91.

6. Hendricks, C. A., Almeida, K. H., Stitt, M. S., Jonnalagadda, V. S., Rugo, R. E., Kerrison, G. F., and Engelward, B. P. (2003) *Proc. Natl. Acad. Sci. USA 100*, 6325-6330.

7. Friedberg, E. C., Walker, G. C., and Siede, W. (1995) *DNA Repair and Mutagenesis*, ASM Press, Washington, D.C.
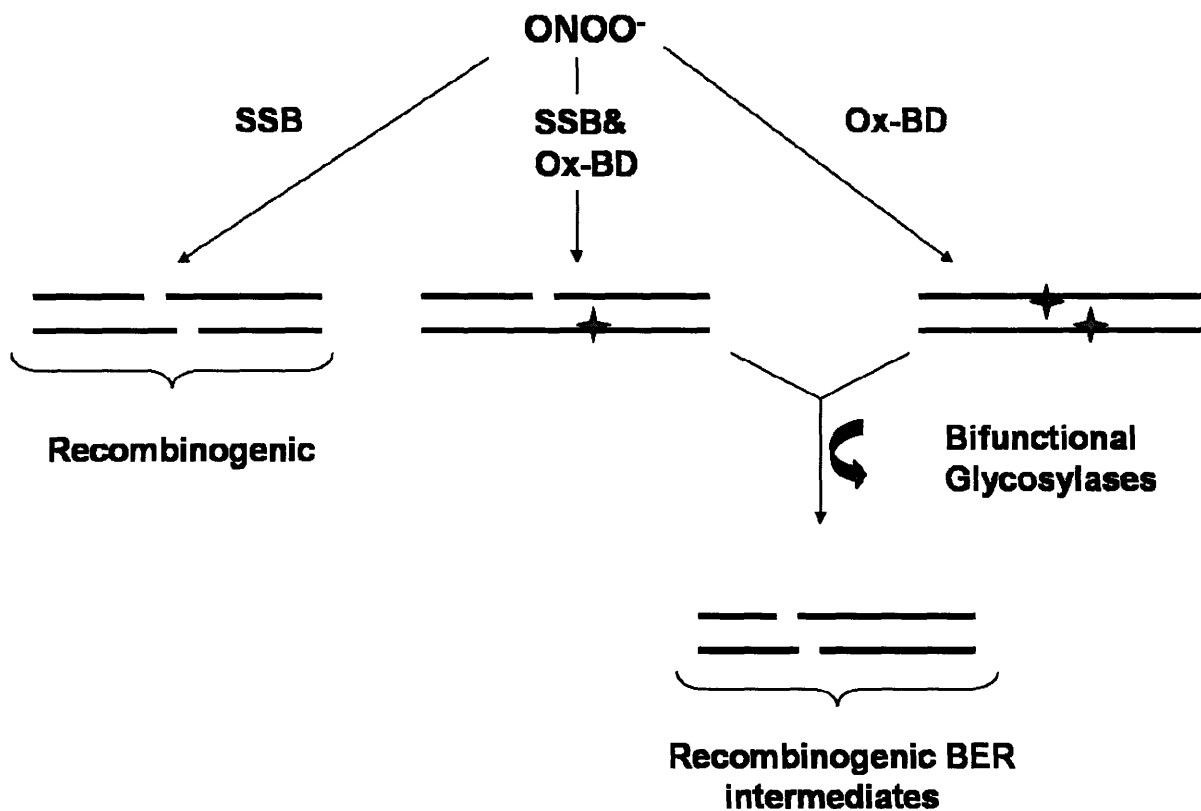
Figure 3.1. Possible mechanisms for the formation of recombinogenic double strand breaks with ONOO⁻ treatment. Double strand breaks may be formed via closely opposed single strand breaks (SSB) and oxidative base lesions (Ox-BD) (denoted by a star) on opposite strands, which could potentially be converted into double strand breaks by DNA glycosylases.
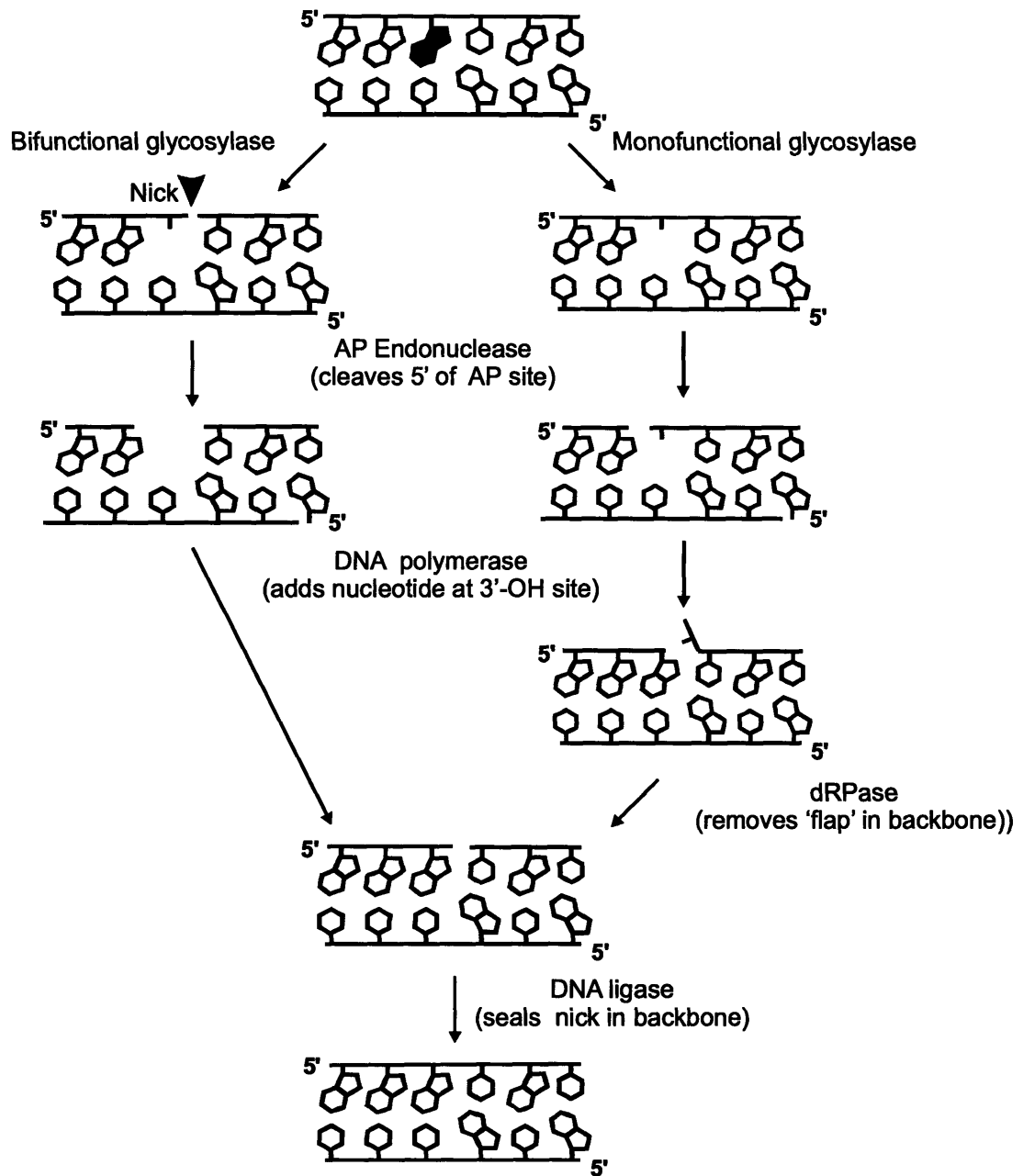
Figure 3.2. Schematics of Base Excision Repair (BER) Pathway. Bifunctional glycosylases remove the damaged base and subsequently cleave the sugar phosphate backbone 3' to the abasic site (shown with an arrow head). Monofunctional glycosylases remove the damaged base creating an abasic site (no subsequent cleavage of the sugar phosphate backbone takes place). AP endonucleases cleave the sugar phosphate backbone 5' of the abasic site. The DNA polymerase adds the missing nucleotide. In case of monofunctional glycosaylases, the abasic ribose is removed by deoxyribosephosphodiesterase (dRpase). The repair is completed by the action of ligase (7).

90

PART B

Dissecting Heparan Sulfate Glycosaminoglycan Structure

& its Effects on Stem Cell Differentiation

CHAPTER 4

Introduction

CHAPTER 4

Introduction

4.1 Extracellular Matrix and its Components

The extracellular matrix (ECM) in which cells reside determines many of the biological properties of the cells and the characteristics of the tissues formed (*1*). Although ECM was historically thought as an inert scaffold for cell and tissues, in recent years ECM has emerged as an important modulator of inter-cellular and intra-cellular signaling events (*e.g.*, see references (*2-6*)). The currently emerging view is one in which there is an active interplay between the cell and ECM. According to this view, cells synthesize their ECM in a highly regulated and specific fashion. In turn the ECM serves to modulate the signaling events in a structurally specific manner and impinges on the functioning of the cell.

The ECM is composed of two major classes of molecules – proteins and polysaccharides. The protein components include fibrous proteins (*e.g.*, collagens and elastin), which mainly provide tensile strength, adhesive glycoproteins (*e.g.*, fibronectin, laminin, tenascin), and proteoglycans- the core proteins with attached polysaccharides. While glycoproteins and proteoglycans provide a hydrated gel which resists compressive forces, they also play a major role in cellular signaling events (*7*). Complex polysaccharides called glycosaminoglycans (GAGs) constitute the second major component of the ECM. GAGs are linear polysaccharides made up of disaccharide building blocks consisting of an amino sugar (either a glucosamine or galactosamine) and a uronic acid (either a glucuronic acid or iduronic acid) (*8*). In addition to these resident components, the ECM also contains cytokines, chemokines and growth factors (*e.g.*, EGF, FGF, TNF, IL-2) which are involved

in signaling events.

There are 4 main classes of GAG family of polysaccharides - hyaluronan, chondroitin/dermatan sulfate, keratan sulfate, and heparin/heparan sulfate (7, 9). The disaccharide repeat units, number of disaccharides that make up a chain and examples of tissue distribution of these 4 classes of GAGs are shown in Table 4.1. It is noteworthy that hyaluronan is considered to be the simplest class of GAGs (10). In contrast, keratan sulfate, chondroitin/dermatan sulfate and heparin/heparan sulfate may have sulfation at various sites of their disaccharide repeat units which adds complexity to their structure and enables them to have structural specificities (7).

Main Classes of GAGs	Disaccharide Repeat Unit	Number of Dissacchari-des	Examples of Tissue Distribution in Mammals	Refs.
Hyaluronan	Glucosamine linked to glucuronic acid	250 to 25,000	Skin, skeletal tissues, synovidal fluid and the vitreous humor of the eye	(10)
Chondroitin/ Dermatan Sulfate	Galactosamine linked to uronic acid	Average of 40	Cartilage, brain, connective tissue, fibroblasts, neural cells, endothelial cells, lymphocytes and myeloid cells	(11, 12)
Keratan Sulfate	Glucosamine linked to galactose	≤ 50	Bone and cartilage	(13)
Heparin/ Heparan Sulfate	Glucosamine linked to uronic acid	20 to 200	Mast cells (heparin), every type of cells (heparan sulfate)	(14, 15)

Table 4.1. The disaccharide repeat units, number of disaccharides that make up a chain, and examples of tissue distribution of the 4 main classes of GAGs.

Although hyaluronan is found as a free polysaccharide in the ECM, the other GAGs are found attached to proteoglycan cores (16). Most proteoglycans carry between 1-3 GAG

94

chains. Numerous growth factors and cytokines are known to bind to the GAG components

of proteoglycans (9, 17). Accordingly, GAGs on a cell's surface are critical modulators of the

signals received from the extracellular milieu. Furthermore, it is increasingly becoming clear

that the fine structure of GAGs play an important role in their binding to the signaling

molecules (18).

4.2 Heparan Sulfate Glycosaminoglycans (HSGAGs)

The most complex member of the GAG family is HSGAGs, and this results from

their structural diversity. The acronym HSGAG includes both heparin (a highly sulfated

polysaccharide commonly isolated from mast cell or mucosa) and heparan sulfate that is

present at the cell surface and the ECM. HSGAGs are characterized by a linear chain of 20–

200 disaccharide units of uronic acid linked 1,4 to an α-D-glucosamine (Figure 4.1) (9, 17,

19). HSGAGs can vary both in terms of the number of disaccharide repeat units and the

chemical modifications internal to each repeat unit. The chemical modifications that may

take place include the O-sulfation at the second carbon of the uronic acid (2-O) and the O-

sulfation at the third and sixth carbons of the glucosamine (3-O and 6-O), as well as N-

sulfation (N-S) or N-acetylation (N-Ac) of the second carbon of the glucosamine (Figure 4.1)

(20). Together, the four different modification sites give rise to 24 different possible

structures for a disaccharide repeat for a particular uronic acid isomer. Furthermore, there are

two variations of the uronic acid component: α-L-iduronic acid or β-D-glucuronic acid.

Altogether, these possible modifications combined with the two possible uronic acid epimers

yield 48 possible disaccharide, which makes HSGAG the most information dense and diverse

biopolymers when compared to DNA (made up of 4 bases), and proteins (made up of 20

amino acids). For example, an HSGAG polymer made up of 5 disaccharide units can have a total of 48^5 (~250 million possible sequences, whereas a penta-peptide made of 5 amino acids can have 20^5 (3.2 million) different sequences or a DNA strand of 5 nucleotides can only have 4^5 (1024) different sequences.

Due to this high sequence diversity of HSGAGs, understanding how the assembly of 48 building blocks is regulated by a cell to provide a functional polysaccharide and the biological functions mediated by these specific HSGAG structures has been one of the most challenging areas of research.

4.3 Biosynthesis and Degradation of HSGAGs

Virtually all cells express proteoglycans that carry HSGAG chains (7). Much of the key advances in dissecting the biosynthetic pathway for HSGAGs has been made possible through targeting the enzymes involved in HSGAG biosynthesis at a whole-organism level (*21-29*).

Unlike other structurally complex charged biopolymers such as DNA, RNA or proteins, HSGAGs do not undergo template-based biosynthesis. Instead, their biosynthesis in mammals is regulated by a complex series of enzymatic steps initiated within the golgi apparatus. HSGAG biosynthesis occurs in 4 key steps in mammalian cells. These steps are (1) synthesis of the proteoglycan core; (2) synthesis of the tetrasaccharide linkage sequence; (3) extension of the HSGAG chain by copolymerization of glucuronic acid and N-acetylglucosamine; and (4) modification of the nascent chain by a complex series of enzymatic steps. These steps are explained in more detail below.

The synthesis of the proteoglycan core is the first step in HSGAG biosynthesis. Some

examples of HSGAG proteoglycans are perlecan, agrin, syndecans, betaglycans, glypicans and serglycin (30-32). The core proteins of cell surface proteoglycans may be transmembrane (e.g., syndecan) or GPI-anchored (e.g., glypican). These proteins range in size from 10kDa (serglycin) to 400 kDa (perlecan). However the number of HSGAG chains attached to these core proteins does not necessarily parallel their size. For example, while 1-3 HSGAG chains are attached to perlecan, serglycin may have as high as 10-15 heparin chains. Except for serglycin, most proteoglycans, irrespective of their size carry 1-3 HSGAG chains. Proteoglycans are synthesized in endoplasmic reticulum of cells and are transported to Golgi for subsequent polysaccharide addition (7).

HSGAGs are attached to their core proteoglycan at a serine residue via a tetrasaccharide linkage regions consisting of serine–xylose-galactose-galactose-glucuronic acid. Thus, the synthesis of the proteoglycan is followed by the synthesis of the tetrasaccharide linkage sequence (Figure 4.2), which constitutes the second step in HSGAG biosynthesis (12, 33). Xylotransferase initiates this process by using UDP-xylose as a donor. Although many studies have been pursued to find a consensus sequence of amino acids determining the attachment site, a perfect consensus sequence has not been found. However, it is noteworthy to mention that a glycine residue was always found next to the carboxyl-terminal side of the serine (34).

After the synthesis of the tetrasaccharide linkage sequence, the HSGAG chain is elongated by alternating additions of glucuronic acid and N-acetyl-glucosamine to form repeating $\beta1,4$-linked disaccharide units (Figure 4.2). This polymerization constitutes the third step of HSGAG biosynthesis and is performed by two glycosyl transferases, EXT1 and EXT2, which belong to the exostosin family of putative tumor suppressor genes. Together,

the protein products of these genes form a stable heterodimeric complex in the Golgi and extend the HSGAG chain (*35, 36*).

The last step in HSGAG biosynthesis is the modification of the nascent HSGAG chain. The nascent chain is modified either concomitantly or independently by the sequential activity of a series of enzymes which leads to N-deacetylation-sulfation, O-sulfation and epimerization of some of the glucuronic acid to iduronic acid (Figure 4.3) (*14, 37*). The HSGAG modifying enzymes, N-deacetylase-N-sulfotransferase, the 2-O, 3-O, and 6-O heparan sulfate sulfotransferases, and epimerase play a key role in this modification process. In the recent years, several tissue-specific isoforms of the biosynthetic enzymes have been identified, giving this system an additional level of complexity (*38*). For example, five isoforms of heparan sulfate 3-O sulfotransferase and three isoforms of heparan sulfate 6-O sulfotransferase have been identified to date (*39, 40*). Therefore, cell or tissue-specific HSGAG biosynthesis might be a result of tissue-specific expression of certain enzyme isoforms.

As a result of the four biosynthetic steps described above, a highly heterogeneous HSGAG chain is formed. These HSGAG chains are either transported to the cell surface or secreted into the extracellular environment. Although the regulatory factors behind the location of the HSGAGs are not completely understood, it is believed that the proteoglycan core is the determining factor (*16*). For example, while syndecan is often transported to the cell surface, perlecan is found in the extracellular matrix.

The synthesized proteoglycans are also subjected to continuous turnover (*15*). While some of these proteoglycans are shed via proteolytic degradation, others can be internalized by endocytosis. A series of enzymes are involved in the chemical degradation of the

endocytosed HSGAG chains in the lyzozymes (*41*). The long carbohydrate chain is first cleaved into smaller polysaccharide fragments by endoglycosidases termed heparanases (*42, 43*). The remaining polysaccharide fragments are sequentially degraded by cleavage at the terminal end via an endoglucuronidase (either α-iduronidase or β-glucuronidase) followed by desulfation of this uronic acid residue via an epimer-specific sulfatase. The resultant terminal glucosamine is next cleaved by α -N-acetylglucoaminidase, followed by desulfation and N-acetylation of this residue by a combination of sulfatases and acetyl transferases of which heparan-N-sulfatase, acetyl-coA:N-acetyltransferase, N-acetylglucosamine-6-sulfatase, and N-acetylglucosamine-3-sulfatase are key players (*46, 47*). Interestingly, heparanase expression has also been implicated in a variety of pathological processes including cancer progression, angiogenesis, and development (*44, 45*).

4.4 Biological Functions of HSGAGs

In recent years, there has been a significant increase in the number of proteins identified that interact with HSGAGs (*17, 48*). These proteins include growth factors, cytokines, morphogens, enzymes, and modulate a variety of signaling pathways that impinge on biological processes such as embryogenesis (*49, 50*), hemostasis (*51*), tumor-progression, -metastasis (*52, 53*), and angiogenesis (*54*). Some known examples of HSGAG binding proteins and their biological functions are listed in Table 4.2.

Area	Protein	Effect of Interaction with HLGAGs	References
Hemostasis	Antithrombin III	Conformational change	(55)
Growth Factors	FGF HGF	Dimerization, receptor specificity Growth Factor sequestering	(56, 57)
Morphogens	WNT	Gradient formation	(29, 58)
Lipoprotein Clearance	Lipoprotein lipase	Stabilization of lipoprotein lipase, enhanced uptake	(59)
Cell-Cell Adhesion	P-selectin	Modulation of leukocyte homing	(60, 61)
Cell-Matrix Adhesion	Laminin Fibronectin	Adhesion to ECM Formation of Focal Adhesion points	(62) (63)
Angiogenesis	VEGF	Modulation of receptor-ligand interactions	(64)
Alzheimer	β-amyloid	Inhibition of clearance	(65)
Enzymes	Heparanase, Heparinase, Sulfotase	Substrate	(66) (67) (68)

Table 4.2. Examples of HSGAG binding proteins and their biological functions.

4.4.1 Mechanisms of HSGAG Action

It is becoming increasingly clear that HSGAGs at the cell surface and in the ECM interact with proteins primarily to (1) facilitate receptor-ligand interactions, (2) modulate their diffusion and effective concentration and (3) act as a reservoir for signaling molecules (18, 69). These mechanisms are described below in more detail.

HSGAGs interact with proteins at the cell surface and ECM to mediate receptor-ligand interactions. Several crystal structure analysis of HSGAG chains with proteins revealed the mechanism of how HSGAGs bind to proteins to mediate receptor-ligand interactions (55). These studies revealed that (1) proteins bind to specific units on HSGAG chains, (2) HSGAGs may act as a platform that enables oligomerization of proteins, and (3) HSGAGs may cause conformational changes of proteins upon binding. For example, it has

been shown that FGF binds to a specific hexasaccharide unit within an HSGAG chain (*70-73*). Thus, long HSGAG chains that carry this specific hexasaccharide unit in tandem may act as a platform and enable oligomerization of FGF ligands. Oligomerization of the ligands affects the kinetics and thermodynamics of receptor-ligand interactions, and is the major mechanism of some signaling cascades. For example, it has been shown that HSGAGs mediate FGF2 oligomerization and lead to increased receptor phosphorylation. The increased receptor phosphorylation in turn leads to increased signaling cascades (*74, 75*).

HSGAGs have also been shown to cause conformational changes in the proteins upon binding. This conformational change in turn converts the inactive form of the protein to its active form. For example, HSGAGs bind to antithrombin III and induce a conformational change to convert it into an active inhibitor of thrombin and factor Xa serine proteases. This process is a critical mediator of the coagulation cascade (*55*) (described in more detail in section 4.4.2).

Understanding the structure-function relationships of HSGAG–protein interactions has been the focus of research for many years. The conformational flexibility of the polysaccharide backbone provided by the iduronate residue was proposed to be central to the specific binding of HSGAG oligosaccharides to a given protein (*76*). Unlike most other pyranose rings, iduronate is able to adopt a range of different ring conformations such as chair, $^1C_4$ or $^4C_1$, or skew-boat, $^2S_0$, thus providing the polysaccharide chain with conformational flexibility (*77, 78*). The conformational changes in HSGAG that are induced upon protein binding are reminiscent of those occurring in DNA, where, as a rule, protein binding causes kinked DNA or DNA bending for maximum contact and specificity (*17, 18, 79*).

HSGAGs modulate the diffusion and effective concentration of proteins. HSGAGs

mediate diffusion of proteins and form a gradient from the ECM to the cell surface by

providing passive resistance to free diffusion as well as by dictating diffusion and gradient

formation in a sequence specific fashion. This is best exemplified by the developmental

studies done on the common fruit fly – *Drosophila* (*29, 50, 80*). These studies have shown

that a major mechanism by which cells create a gradient for a morphogen is by expressing

specific HSGAGs on the ECM, which bind and modulate the movement of a morphogen. For

example, Wnt family of morphogens control cell proliferation and differentiation during

development, and it was found that HSGAGs have tissue-specific effects on Wnt signaling,

modulating both short and long-range functioning of the morphogen (*81, 82*). It was also

shown that alteration of the fine structure of the HSGAGs involved in this process abolished

gradient formation and arrested development (*29, 80*).

HSGAGs act as a reservoir of signaling molecules at the cell surface and ECM. An

example for the role of HSGAGs in sequestering proteins is provided by FGF. It has been

shown that HSGAGs in the ECM, specifically in the basement membrane, bind and sequester

FGF molecules in an inactive form (*69, 75*). Upon external stimuli, these FGF proteins are

rapidly released by controlled digestion of the ECM, providing a rapid response without the

need for *de novo* protein synthesis (*83, 84*).

4.4.2 Effects of HSGAGs on Physiological and Pathophysiological Processes

The interaction of HSGAGs with a variety of proteins impinges on many

physiological and pathophysiological processes. Some of the well studied processes are

hemostasis, morphogenesis, tumor-progression, tumor-metastasis and angiogenesis. These

processes and how HSGAG sequences impinge their roles in these processes are discussed below.

HSGAGs in Hemostasis. Hemostasis is the maintenance of the hemodynamic state of blood. In humans, a complex coagulation cascade exists to stop excessive bleeding from ruptured blood vessels to maintain the hemodynamic state of blood. One of the key steps of this coagulation cascade is the activation of prothrombin (factor II) to thrombin (factor IIa), which then acts on fibrin to initiate the mesh network responsible for clot formation *(69)*. To maintain hemostasis, it is as important to inhibit the coagulation cascade once the bleeding stops, as it is to initiate it. One way this can be achieved is via the inhibition of Factor IIa. Factor IIa can be inhibited when bound by inhibitory AT-III. The inhibitory effect of AT-III is mediated by a specific HSGAG sequence, $H_{NAc/S,6S}GH_{NS,3S,6S}I_{2S}H_{NS,6S}$ *(85)*. Once bound to this highly sulfated pentasaccharide, AT-III goes through a conformational change and inhibits many of the coagulation proteases including Factor IIa *(55, 85, 86, 87)*. As a summary, specific HSGAG sequences inhibit the coagulation cascade by binding to AT-III, which subsequently inhibits factor IIa. The role of HSGAGs in modulation of coagulation cascade exemplifies how specific HSGAG sequences impinge on important biological functions. Importantly, highly sulfated heparin is currently being used clinically as a major anticoagulant for a variety of pathophysiological conditions *(85)*.

HSGAGs in Morphogenesis. Various genetic studies done with Drosophila have shown the significance of HSGAGs in body patterning in vivo. Genetic screens for mutations that lead to abnormal morphogenesis and body development uncovered a number of genes involved in HSGAG biosynthesis (see Table 4.3). The normal phenotypes were recovered by ectopic expression of the wild type HSGAG biosynthetic enzymes. Furthermore, all of these

mutants were shown to have defective signaling pathways that are known to be critical in

development, such as Wingless, Hedgehog, and FGF signaling. These studies, combined with

cell biological and biochemical studies done to elucidate the mechanism of Wingless,

Hedgehog and FGF signaling provide convincing evidence that impaired HSGAG

biosynthesis compromises morphogenesis (*88-91*). Indeed, impaired GlcNAc/GlcA

biosynthesis was found to be related to Hereditary Multiples Exostoses in humans, which is a

pathologic condition that is associated with abnormal skeletal development (*92*).

HSGAG Biosynthetic Enzyme	Mutated Drosophila Genes	Defective Pathways	Human Homologue	Related Pathology	Refs.
UDP-glucose dehydogenase	*sugarless*	Wingless	UDP-Glc-DH	None	(*81*) (*93*) (*94*)
N-Deacetylase-Sulfotransferase	*sulfateless*	Wingless FGF	NDST-1/2	None	(*80*)
GlcNAc/GlcA Polymerase	*tout-velu*	Hedgehog	EXT-1/2/3 EXL-1/2/3	Hereditary Multiples Exotoses	(*95*) (*92*)
2-O sulfotransferase	*pipe*	Dorsal-ventral patterning	HS2TS HS3TS HS6TS	None	(*50*) (*96*)

Adapted from (*97*)

Table 4.3. HSGAG biosynthetic enzymes that are involved in morphogenesis and the signaling pathways impinged on.

HSGAGs in Tumor Progression and Metastasis. HSGAGs have been implicated to

play a critical role in tumor onset, progression, and metastasis. It has been proposed that

HSGAGs in the basement membrane constitute a barrier to infiltration of tumor cells into

normal tissue, and that the metastatic and invasive properties of tumors are associated with

the degradation of HSGAGs *(98)*. Consistent with this idea, it was found that many tumors

over-express an extracellular mammalian HSGAG degrading enzyme, heparanase *(66, 99)*. It

also appears that the metastatic potential of some tumors correlates with the expression levels of heparanase *(98)* and suppression of heparanase activity has been shown to inhibit tumor metastasis in vivo *(100)*. It has also been proposed that HSGAGs modulate tumor cell adhesion to the basement membrane and prevent its infiltration, and it has been shown that exogenous addition of highly sulfated HSGAGs interfere with tumor cell adhesion to the basement membrane and leads to tumor metastasis *(101, 102)*. In extensive studies done in our laboratory, tumor transplants treated with bacterial heparinase I and heparinase III, which have distinct substrate specificities, revealed that HSGAGs play a critical role in a structurally specific manner, in tumor onset, progression, and metastasis . While heparinase I cleaves at the highly sulfated regions, heparinase III cleaves at the undersulfated regions *(8, 103, 104)*. In our studies heparinase III inhibited tumor growth and suppressed the development of lung metastasis while heparinase I promoted tumor growth and was ineffective in preventing metastasis. . It is now established that there are specific structural changes in tumor cell surface HSGAGs as a function of the oncogenic process *(105, 106)*. For example, elevated levels of 6-O-sulfation and reduced levels of 2-O- and N-sulfation were found during progression of human colon adenoma to carcinoma (105). Upregulation of 6-O sulfation was also found in Lewis lung carcinoma cells *(107)*. On the contrary, 6-O-sulfation was downregulated during malignant transformation of mouse mammary epithelial cells *(108)*. These studies again suggest a role for specific HSGAG structures in tumor progression and metastasis.

HSGAGs in Angiogenesis. Angiogenesis is the process of generating new capillary blood vessels. It is controlled by a fine balance between pro- and anti-angiogenic factors, and is a critical process for a developing tumor to grow beyond a certain size (~3mm3) *(109)*.

Various studies have implicated the role of HSGAGs in the modulation of angiogenesis and it has been shown that this modulation occurs through the structural-interactions between HSGAGs and the angiogenic growth factors or the negative regulators of angiogenesis (*54, 64, 109, 110*). For example, it was shown in a chorioallantoic membrane assay that heparinases inhibited angiogenesis (*54*). In another study, endostatin, a potent anti-angiogenic molecule, was found to be an HSGAG binding protein (*111*). The emerging view from these studies is that HSGAGs are potent mediators of pro- and anti-angiogenic signalling pathways and a change in the HSGAG structure can tip the angiogenic balance to either promote or inhibit angiogenesis.

4.5 Role of HSGAG Degrading Enzymes in Structure-Function Studies

Understanding specific structure-function relationship of HSGAGs has been a very challenging task due to their highly heterogeneous structure and their low availability on the cell surface and ECM. Furthermore, unlike other biopolymers (*e.g.* polynucleotides or polypeptides), tools that can be used to amplify HSGAGs have yet not been developed because of their complex biosynthesis.

Tools for the structural analysis of HSGAGs are just coming of age. One of the main focuses of our lab has been to develop technologies based on recombinant HSGAG degrading enzymes from *Flavobacterium heparinum*. Flavobacteria synthesize many glycosaminoglycan degrading enzymes as part of their catabolic life cycle (*112, 113*). Among these enzymes are heparinases, specifically, heparinase-I, heparinase-II, heparinase-III, Δ4,5-glycuronidase, 2-O-sulfatase, 3-O-sultafase and N-sulfatase. Each of these enzymes has unique substrate specificities and act on the HSGAG chain in a sequential order (*8, 114*).

The reactions catalyzed by these enzymes and the order they act on the HSGAG chain is shown in Figure 4.4. Thus, with their substrate specificities, these enzymes provide us with invaluable tools that can be used in determination of HSGAG sequences, similar to the way proteases are used to obtain sequence information about proteins, and restriction enzymes are used to obtain sequence information about DNA.

The first step in the catabolic cycle of *Flavobacteria* is the depolymerization of the HSGAG chains. This is catalyzed by heparinases, which cleavage the glycosidic bond by β-elimination (*8*). There are three types of heparinases, -I, -II, and –III, and all three heparinases have distinct substrate specificities and cleave HSGAG chains at distinct sites (*115, 116*). Heparinase I cleaves HSGAGs at sites with an *O*-sulfated L-iduronic linkage, heparinase III cleaves at unsulfated uronic acid moiety, and heparinase II cleaves both at sulfated iduronic and unsulfated uronic acid regions of HSGAGs. Exhaustive digestion of an HSGAG chain with a combination of heparinases yields the disaccharide building blocks that make up the HSGAG chain. The next step in the catabolic cycle of *Flavobacteria* is the removal of the 2-O-sulfate from the disaccharide by the action of 2-O sulfatase, followed by the hydrolysis of the disaccharide into its monosaccharide units by Δ4,5-glycuronidase. This is followed by the removal of the 6-O, 3-O and N-sulfates from the glucosamine moiety by the actions of 6-O, 3-O and N-sulfatases respectively. It is yet not known whether 3-O sulfatase acts on the glucosamine before N-sulfatase, or vice versa (*8*). In summary, all these enzymes have unique substrate specificities and can be used in obtaining HSGAG sequence information.

In addition to their role in HSGAG sequencing, HSGAG degrading enzymes also provide us with invaluable tools that can be used to modify cell surface and ECM HSGAG

107

structures *in vitro* and *in vivo* to reveal specific structure-function relationships. For instance, heparinases from *Flavobacteria* has been used to reveal the role of HSGAGs in tumor progression and metastasis, as well as in angiogenesis (*53, 54*).

4.6 Specific Objectives

The emerging view is that specific HSGAG sequences bind to proteins, impinge on critical signaling pathways and modulate fundamental biological processes. It is becoming increasingly clear that glycomics impinges on proteomics, and by doing so, it impinges on transcriptomics and even genomics (*18*). Thus, a thorough understanding of the role that HSGAGs play is a very important dimension in our understanding of biological processes in the post-genomic era.

One of the major focuses of our lab has been to use recombinant HSGAG degrading enzymes from Flavobacteria to (1) determine specific sequence information of HSGAG chains, (2) to reveal the role HSGAGs play in biological processes.

Chapter 5 in this thesis describes the molecular cloning, recombinant expression and biochemical characterization of an HSGAG degrading enzyme, Δ4,5 glycuronidase from *Flavobacterium heparinum*. This enzyme, with its unique substrate specificity is an invaluable addition to the repertoire of recombinant HSGAGs degrading enzymes that can be used for sequencing of HSGAG chains.

In Chapter 6 of this thesis, the studies done to elucidate the role of HSGAGs in differentiation of embryonic stem cells into endothelial cells are described. These studies were conducted by using some of the HSGAG degrading enzymes described above, in combination with other tools, and have significant implications in regenerative therapies. The

HSGAG paradigm we discovered in ES cell differentiation opens new approaches and strategies for tissue engineering and will enable the development of novel molecular therapeutics.

4.7 References

1. Boudreau, N. J., and Jones, P. L. (1999) *Biochem J 339 (Pt 3)*, 481-8.
2. Zhan, M., Zhao, H., and Han, Z. C. (2004) *Histol Histopathol 19*, 973-83.
3. Cordes, N., and Beinke, C. (2004) *Cancer Biol Ther 3*, 47-53.
4. Cordes, N., and Meineke, V. (2004) *J Mol Histol 35*, 327-37.
5. Humphries, M. J., Travis, M. A., Clark, K., and Mould, A. P. (2004) *Biochem Soc Trans 32*, 822-5.
6. Piez, K. A. (1997) *Matrix Biol 16*, 85-92.
7. Varki A, C. R., Esko J, Freeze H, Hart G, Marth J (1999) *Essentials of Glycobiology*, Cold Spring Harbor, New York.
8. Ernst, S., Langer, R., Cooney, C. L., and Sasisekharan, R. (1995) *Crit Rev Biochem Mol Biol 30*, 387-444.
9. Lindahl, U., and Hook, M. (1978) *Annu Rev Biochem 47*, 385-417.
10. Fraser, J. R., Laurent, T. C., and Laurent, U. B. (1997) *J Intern Med 242*, 27-33.
11. Sundblad, G., Kajiji, S., Quaranta, V., Freeze, H. H., and Varki, A. (1988) *J Biol Chem 263*, 8897-903.
12. Sundblad, G., Holojda, S., Roux, L., Varki, A., and Freeze, H. H. (1988) *J Biol Chem 263*, 8890-6.
13. Greiling, H. (1994) *Exs 70*, 101-22.
14. Salmivirta, M., Lidholt, K., and Lindahl, U. (1996) *Faseb J 10*, 1270-9.
15. Williams, K. J., and Fuki, I. V. (1997) *Curr Opin Lipidol 8*, 253-62.
16. Bernfield, M., Gotte, M., Park, P. W., Reizes, O., Fitzgerald, M. L., Lincecum, J., and Zako, M. (1999) *Annu Rev Biochem 68*, 729-77.
17. Conrad, H. E. (1998) *Heparin-Binding Proteins*, Academic Press, San Diego.
18. Sasisekharan, R., and Venkataraman, G. (2000) *Curr Opin Chem Biol 4*, 626-31.
19. Tumova, S., Woods, A., and Couchman, J. R. (2000) *Int J Biochem Cell Biol 32*, 269-88.
20. Venkataraman, G., Shriver, Z., Raman, R., and Sasisekharan, R. (1999) *Science 286*, 537-42.
21. van den Born, J., Pikas, D. S., Pisa, B. J., Eriksson, I., Kjellen, L., and Berden, J. H. (2003) *Glycobiology 13*, 1-10.
22. Pikas, D. S., Eriksson, I., and Kjellen, L. (2000) *Biochemistry 39*, 4552-8.
23. Sugahara, K., and Kitagawa, H. (2000) *Curr Opin Struct Biol 10*, 518-27.
24. Grobe, K., Ledin, J., Ringvall, M., Holmborn, K., Forsberg, E., Esko, J. D., and Kjellen, L. (2002) *Biochim Biophys Acta 1573*, 209-15.
25. Ringvall, M., Ledin, J., Holmborn, K., van Kuppevelt, T., Ellin, F., Eriksson, I., Olofsson, A. M., Kjellen, L., and Forsberg, E. (2000) *J Biol Chem 275*, 25926-30.
26. Merry, C. L., and Gallagher, J. T. (2002) *Biochem Soc Symp*, 47-57.
27. Merry, C. L., Bullock, S. L., Swan, D. C., Backen, A. C., Lyon, M., Beddington, R. S., Wilson, V. A., and Gallagher, J. T. (2001) *J Biol Chem 276*, 35429-34.
28. Forsberg, E., Pejler, G., Ringvall, M., Lunderius, C., Tomasini-Johansson, B., Kusche-Gullberg, M., Eriksson, I., Ledin, J., Hellman, L., and Kjellen, L. (1999) *Nature 400*, 773-6.

29.   Tsuda, M., Kamimura, K., Nakato, H., Archer, M., Staatz, W., Fox, B., Humphrey, M., Olson, S., Futch, T., Kaluza, V., Siegfried, E., Stam, L., and Selleck, S. B. (1999) *Nature 400*, 276-80.

30.   Kjellen, L., and Lindahl, U. (1991) *Annu Rev Biochem 60*, 443-75.

31.   Ballinger, M. L., Nigro, J., Frontanilla, K. V., Dart, A. M., and Little, P. J. (2004) *Cell Mol Life Sci 61*, 1296-306.

32.   Delehedde, M., Lyon, M., Sergeant, N., Rahmoune, H., and Fernig, D. G. (2001) *J Mammary Gland Biol Neoplasia 6*, 253-73.

33.   Roden, L., Koerner, T., Olson, C., and Schwartz, N. B. (1985) *Fed Proc 44*, 373-80.

34.   Zhang, L., and Esko, J. D. (1994) *J Biol Chem 269*, 19295-9.

35.   McCormick, C., Duncan, G., Goutsos, K. T., and Tufaro, F. (2000) *Proc Natl Acad Sci U S A 97*, 668-73.

36.   Kobayashi, S., Morimoto, K., Shimizu, T., Takahashi, M., Kurosawa, H., and Shirasawa, T. (2000) *Biochem Biophys Res Commun 268*, 860-7.

37.   Lindahl, U., Kusche-Gullberg, M., and Kjellen, L. (1998) *J Biol Chem 273*, 24979-82.

38.   Habuchi, O. (2000) *Biochim Biophys Acta 1474*, 115-27.

39.   Habuchi, H., Tanaka, M., Habuchi, O., Yoshida, K., Suzuki, H., Ban, K., and Kimata, K. (2000) *J Biol Chem 275*, 2859-68.

40.   Shworak, N. W., Liu, J., Petros, L. M., Zhang, L., Kobayashi, M., Copeland, N. G., Jenkins, N. A., and Rosenberg, R. D. (1999) *J Biol Chem 274*, 5170-84.

41.   Freeman, C., and Hopwood, J. (1992) *Adv Exp Med Biol 313*, 121-34.

42.   Godder, K., Vlodavsky, I., Eldor, A., Weksler, B. B., Haimovitz-Freidman, A., and Fuks, Z. (1991) *J Cell Physiol 148*, 274-80.

43.   Vlodavsky, I., Bar-Shavit, R., Ishai-Michaeli, R., Bashkin, P., and Fuks, Z. (1991) *Trends Biochem Sci 16*, 268-71.

44.   Vlodavsky, I., Michaeli, R. I., Bar-Ner, M., Fridman, R., Horowitz, A. T., Fuks, Z., and Biran, S. (1988) *Isr J Med Sci 24*, 464-70.

45.   Vlodavsky, I., Eldor, A., Bar-Ner, M., Fridman, R., Cohen, I. R., and Klagsbrun, M. (1988) *Adv Exp Med Biol 233*, 201-10.

46.   Bai, X., Bame, K. J., Habuchi, H., Kimata, K., and Esko, J. D. (1997) *J Biol Chem 272*, 23172-9.

47.   Yanagishita, M. (1993) *Experientia 49*, 366-8.

48.   Taipale, J., and Keski-Oja, J. (1997) *Faseb J 11*, 51-9.

49.   Perrimon, N., and Bernfield, M. (2001) *Semin Cell Dev Biol 12*, 65-7.

50.   Perrimon, N., and Bernfield, M. (2000) *Nature 404*, 725-8.

51.   Rosenberg, R. D., Shworak, N. W., Liu, J., Schwartz, J. J., and Zhang, L. (1997) *J Clin Invest 100*, S67-75.

52.   Sasisekharan, R., Shriver, Z., Venkataraman, G., and Narayanasami, U. (2002) *Nat Rev Cancer 2*, 521-8.

53.   Liu, D., Shriver, Z., Venkataraman, G., El Shabrawi, Y., and Sasisekharan, R. (2002) *Proc Natl Acad Sci U S A 99*, 568-73.

54.   Sasisekharan, R., Moses, M. A., Nugent, M. A., Cooney, C. L., and Langer, R. (1994) *Proc Natl Acad Sci U S A 91*, 1524-8.

55.   Jin, L., Abrahams, J. P., Skinner, R., Petitou, M., Pike, R. N., and Carrell, R. W. (1997) *Proc Natl Acad Sci U S A 94*, 14683-8.

56.   Nugent, M. A., and Iozzo, R. V. (2000) *Int J Biochem Cell Biol 32*, 115-20.

57.   Derksen, P. W., Keehnen, R. M., Evers, L. M., van Oers, M. H., Spaargaren, M., and Pals, S. T. (2002) *Blood 99*, 1405-10.

58.   Dhoot, G. K., Gustafsson, M. K., Ai, X., Sun, W., Standiford, D. M., and Emerson, C. P., Jr. (2001) *Science 293*, 1663-6.

59.   Kolset, S. O., and Salmivirta, M. (1999) *Cell Mol Life Sci 56*, 857-70.

60.   Ma, Y. Q., and Geng, J. G. (2002) *J Immunol 168*, 1690-6.

61.   Ma, Y. Q., and Geng, J. G. (2000) *J Immunol 165*, 558-65.

62.   Utani, A., Nomizu, M., Matsuura, H., Kato, K., Kobayashi, T., Takeda, U., Aota, S., Nielsen, P. K., and Shinkai, H. (2001) *J Biol Chem 276*, 28779-88.

63.   Lundmark, K., Tran, P. K., Kinsella, M. G., Clowes, A. W., Wight, T. N., and Hedin, U. (2001) *J Cell Physiol 188*, 67-74.

64.   Iozzo, R. V., and San Antonio, J. D. (2001) *J Clin Invest 108*, 349-55.

65.   Lindahl, B., Westling, C., Gimenez-Gallego, G., Lindahl, U., and Salmivirta, M. (1999) *J Biol Chem 274*, 30631-5.

66.   Hulett, M. D., Freeman, C., Hamdorf, B. J., Baker, R. T., Harris, M. J., and Parish, C. R. (1999) *Nat Med 5*, 803-9.

67.   Shriver, Z., Hu, Y., and Sasisekharan, R. (1998) *J Biol Chem 273*, 10160-7.

68.   Myette, J. R., Shriver, Z., Claycamp, C., McLean, M. W., Venkataraman, G., and Sasisekharan, R. (2003) *J Biol Chem 278*, 12157-66.

69.   Shriver, Z., Liu, D., and Sasisekharan, R. (2002) *Trends Cardiovasc Med 12*, 71-7.

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

71.   Habuchi, H., Suzuki, S., Saito, T., Tamura, T., Harada, T., Yoshida, K., and Kimata, K. (1992) *Biochem J 285 (Pt 3)*, 805-13.

72.   Guimond, S. E., and Turnbull, J. E. (1999) *Curr Biol 9*, 1343-6.

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

74.   Kwan, C. P., Venkataraman, G., Shriver, Z., Raman, R., Liu, D., Qi, Y., Varticovski, L., and Sasisekharan, R. (2001) *J Biol Chem 276*, 23421-9.

75.   Chang, Z., Meyer, K., Rapraeger, A. C., and Friedl, A. (2000) *Faseb J 14*, 137-44.

76.   Casu, B., Choay, J., Ferro, D. R., Gatti, G., Jacquinet, J. C., Petitou, M., Provasoli, A., Ragazzi, M., Sinay, P., and Torri, G. (1986) *Nature 322*, 215-6.

77.   Mikhailov, D., Linhardt, R. J., and Mayo, K. H. (1997) *Biochem J 328 (Pt 1)*, 51-61.

78.   Mulloy, B., Forster, M. J., Jones, C., and Davies, D. B. (1993) *Biochem J 293 (Pt 3)*, 849-58.

79.   Steitz, T. A. (1990) *Q Rev Biophys 23*, 205-80.

80.   Lin, X., and Perrimon, N. (1999) *Nature 400*, 281-4.

81.   Binari, R. C., Staveley, B. E., Johnson, W. A., Godavarti, R., Sasisekharan, R., and Manoukian, A. S. (1997) *Development 124*, 2623-32.

82.   Cumberledge, S., and Reichsman, F. (1997) *Trends Genet 13*, 421-3.

83.   Folkman, J., Klagsbrun, M., Sasse, J., Wadzinski, M., Ingber, D., and Vlodavsky, I. (1988) *Am J Pathol 130*, 393-400.

84.   Dowd, C. J., Cooney, C. L., and Nugent, M. A. (1999) *J Biol Chem 274*, 5236-44.

85.   Rosenberg, R. D. (2001) *N Engl J Med 344*, 673-5.

86.   Desai, U. R., Petitou, M., Bjork, I., and Olson, S. T. (1998) *J Biol Chem 273*, 7478-87.

87. Atha, D. H., Stephens, A. W., Rimon, A., and Rosenberg, R. D. (1984) *Biochemistry 23*, 5801-12.
88. Cohen, M. M., Jr. (2003) *Am J Med Genet 123A*, 5-28.
89. Uren, A., Reichsman, F., Anest, V., Taylor, W. G., Muraiso, K., Bottaro, D. P., Cumberledge, S., and Rubin, J. S. (2000) *J Biol Chem 275*, 4374-82.
90. Wang, S., Ai, X., Freeman, S. D., Pownall, M. E., Lu, Q., Kessler, D. S., and Emerson, C. P., Jr. (2004) *Proc Natl Acad Sci U S A 101*, 4833-8.
91. Allen, B. L., and Rapraeger, A. C. (2003) *J Cell Biol 163*, 637-48.
92. Bellaiche, Y., The, I., and Perrimon, N. (1998) *Nature 394*, 85-8.
93. Hacker, U., Lin, X., and Perrimon, N. (1997) *Development 124*, 3565-73.
94. Haerry, T. E., Heslip, T. R., Marsh, J. L., and O'Connor, M. B. (1997) *Development 124*, 3055-64.
95. The, I., Bellaiche, Y., and Perrimon, N. (1999) *Mol Cell 4*, 633-9.
96. Sen, J., Goltz, J. S., Stevens, L., and Stein, D. (1998) *Cell 95*, 471-81.
97. Princivalle, M., and de Agostini, A. (2002) *Int J Dev Biol 46*, 267-78.
98. Nakajima, M., Irimura, T., Di Ferrante, D., Di Ferrante, N., and Nicolson, G. L. (1983) *Science 220*, 611-3.
99. Vlodavsky, I., Friedmann, Y., Elkin, M., Aingorn, H., Atzmon, R., Ishai-Michaeli, R., Bitan, M., Pappo, O., Peretz, T., Michal, I., Spector, L., and Pecker, I. (1999) *Nat Med 5*, 793-802.
100. Uno, F., Fujiwara, T., Takata, Y., Ohtani, S., Katsuda, K., Takaoka, M., Ohkawa, T., Naomoto, Y., Nakajima, M., and Tanaka, N. (2001) *Cancer Res 61*, 7855-60.
101. Borsig, L., Wong, R., Hynes, R. O., Varki, N. M., and Varki, A. (2002) *Proc Natl Acad Sci U S A 99*, 2193-8.
102. Borsig, L., Wong, R., Feramisco, J., Nadeau, D. R., Varki, N. M., and Varki, A. (2001) *Proc Natl Acad Sci U S A 98*, 3352-7.
103. Ernst, S., Venkataraman, G., Winkler, S., Godavarti, R., Langer, R., Cooney, C. L., and Sasisekharan, R. (1996) *Biochem J 315 (Pt 2)*, 589-97.
104. Pojasek, K., Shriver, Z., Hu, Y., and Sasisekharan, R. (2000) *Biochemistry 39*, 4012-9.
105. Jayson, G. C., Lyon, M., Paraskeva, C., Turnbull, J. E., Deakin, J. A., and Gallagher, J. T. (1998) *J Biol Chem 273*, 51-7.
106. Timar, J., Ladanyi, A., Lapis, K., and Moczar, M. (1992) *Am J Pathol 141*, 467-74.
107. Nakanishi, H., Oguri, K., Yoshida, K., Itano, N., Takenaga, K., Kazama, T., Yoshida, A., and Okayama, M. (1992) *Biochem J 288 (Pt 1)*, 215-24.
108. Safaiyan, F., Lindahl, U., and Salmivirta, M. (1998) *Eur J Biochem 252*, 576-82.
109. Presta, M., Leali, D., Stabile, H., Ronca, R., Camozzi, M., Coco, L., Moroni, E., Liekens, S., and Rusnati, M. (2003) *Curr Pharm Des 9*, 553-66.
110. Folkman, J., and Shing, Y. (1992) *Adv Exp Med Biol 313*, 355-64.
111. Karumanchi, S. A., Jha, V., Ramchandran, R., Karihaloo, A., Tsiokas, L., Chan, B., Dhanabal, M., Hanai, J. I., Venkataraman, G., Shriver, Z., Keiser, N., Kalluri, R., Zeng, H., Mukhopadhyay, D., Chen, R. L., Lander, A. D., Hagihara, K., Yamaguchi, Y., Sasisekharan, R., Cantley, L., and Sukhatme, V. P. (2001) *Mol Cell 7*, 811-22.
112. Payza, A. N., and Korn, E. D. (1956) *Nature 177*, 88-9.
113. Korn, E. D., and Payza, A. N. (1956) *Biochim Biophys Acta 20*, 596-7.
114. Dietrich, C. P., Silva, M. E., and Michelacci, Y. M. (1973) *J Biol Chem 248*, 6408-15.

115. Linhardt, R. J., Turnbull, J. E., Wang, H. M., Loganathan, D., and Gallagher, J. T. (1990) *Biochemistry 29*, 2611-7.
116. Godavarti, R., and Sasisekharan, R. (1996) *Biochem Biophys Res Commun 229*, 770-7.
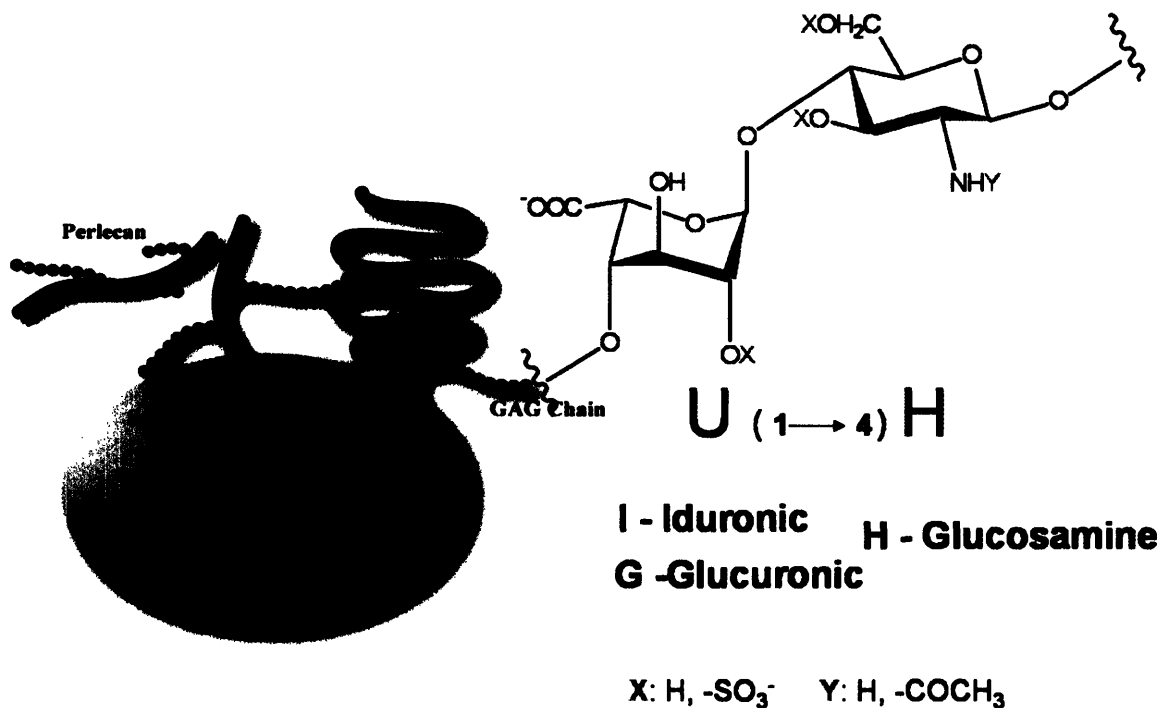
Figure 4.1. Heparan Sulfate Glycosaminoglycans (HSGAGs) repeat unit. HSGAGs are extracellular and cell surface polysaccharides. They are made up of a disaccharide repeat unit of a uronic acid (U) linked 1,4 to a glucosamine (G).There are two possibilities for the uronic component of the disaccharide unit: Iduronic acid or Glucuronic acid. The chemical modifications that may take place physiologically are O-sulfation at the second carbon of the uronic acid (2-O) and the O-sulfation at the third and sixth carbons of the glucosamine (3-O and 6-O), as well as N-sulfation (N-S) or N-acetylation (N-Ac) of the second carbon of the glucosamine.

115
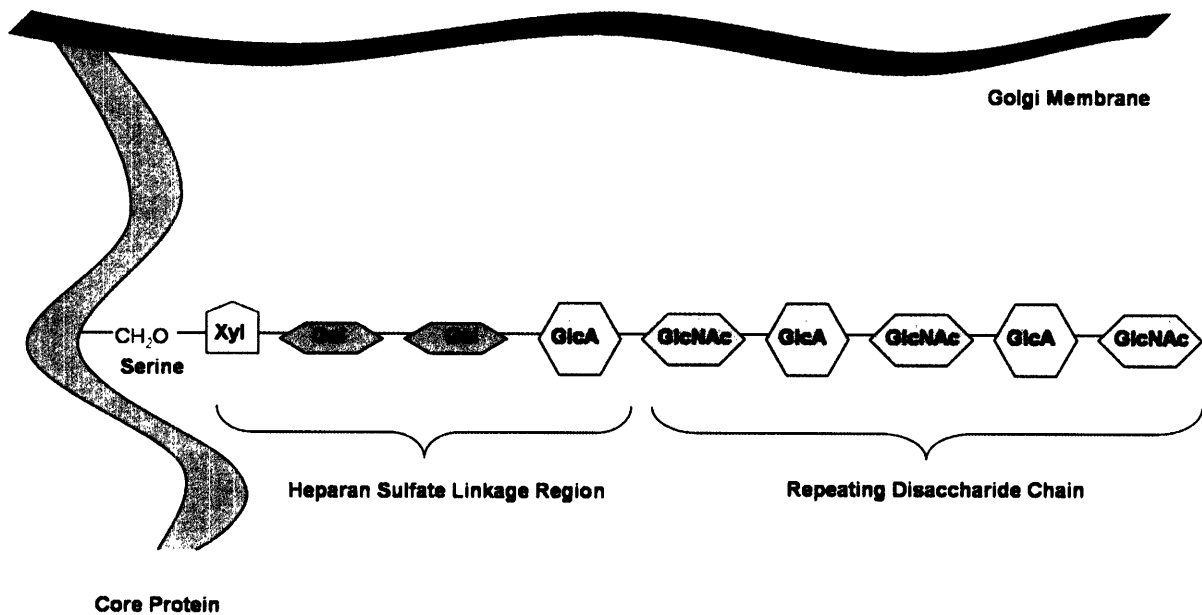
Figure 4.2. Schematics of initiation of HSGAG biosynthesis. HSGAG biosynthesis starts by the addition of the tetrasaccharide linkage sequence [xylose (Xyl)- galactose (Gal)- galactose (Gal)- glucuronic acid (GlcA)] to the serine residue of a core proteoglycan. This is followed by the addition of N-acetyl-glucosamine (GlcNAc) to the linkage tetrasaccharide. The HSGAG chain is then extended by subsequent alternating additions of glucuronic acid and N-acetyl-glucosamine.
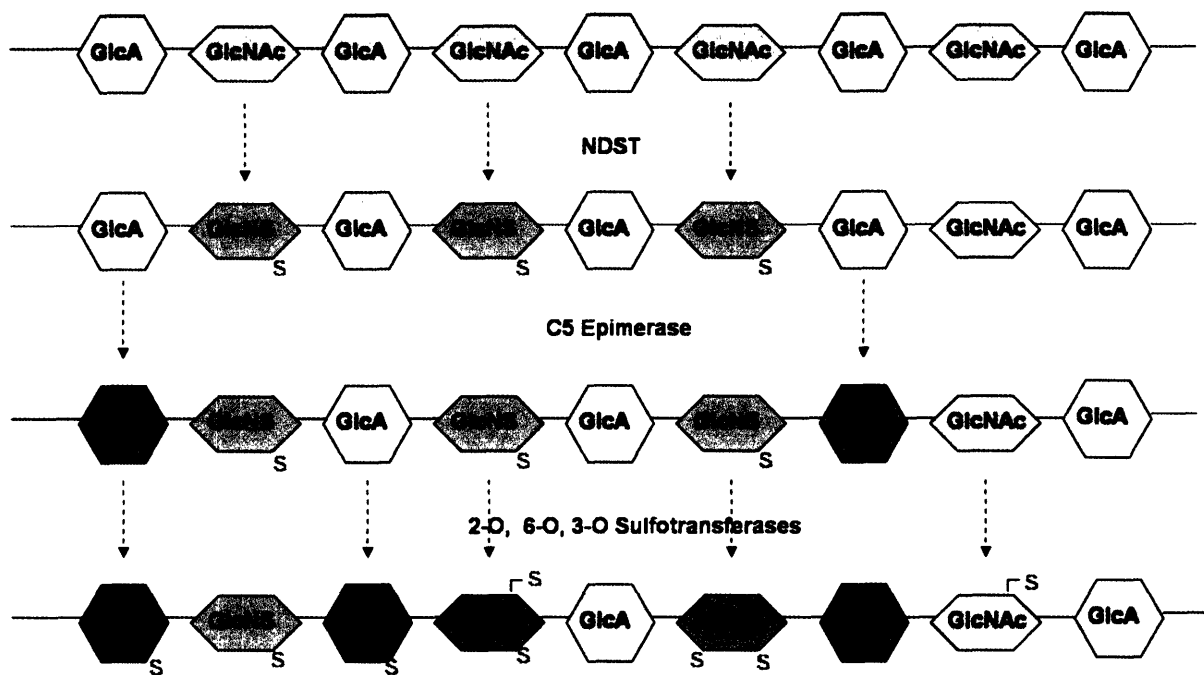
Figure 4.3. Schematics of modification of the nascent HLGAG chain. The nascent chain is modified by the sequential activity of a series of enzymes which leads to N-deacetylation - sulfation, epimerization, and O-sulfations. N-deacetylase-N-sulfotransferase (NDST) converts some N-acetyl-glucosamines (GlcNAc) to N-sulfated-glucosamines (GlcNS). This is followed by the action of epimerase which converts some glucuronic acids (GlcA) to iduronic acids (IdoA). Lastly the 2-O, 3-O, and 6-O heparan sulfate sulfotransferases act to further modify the HSGAG chain. As a result, a highly heterogeneous HSGAG chain is formed.
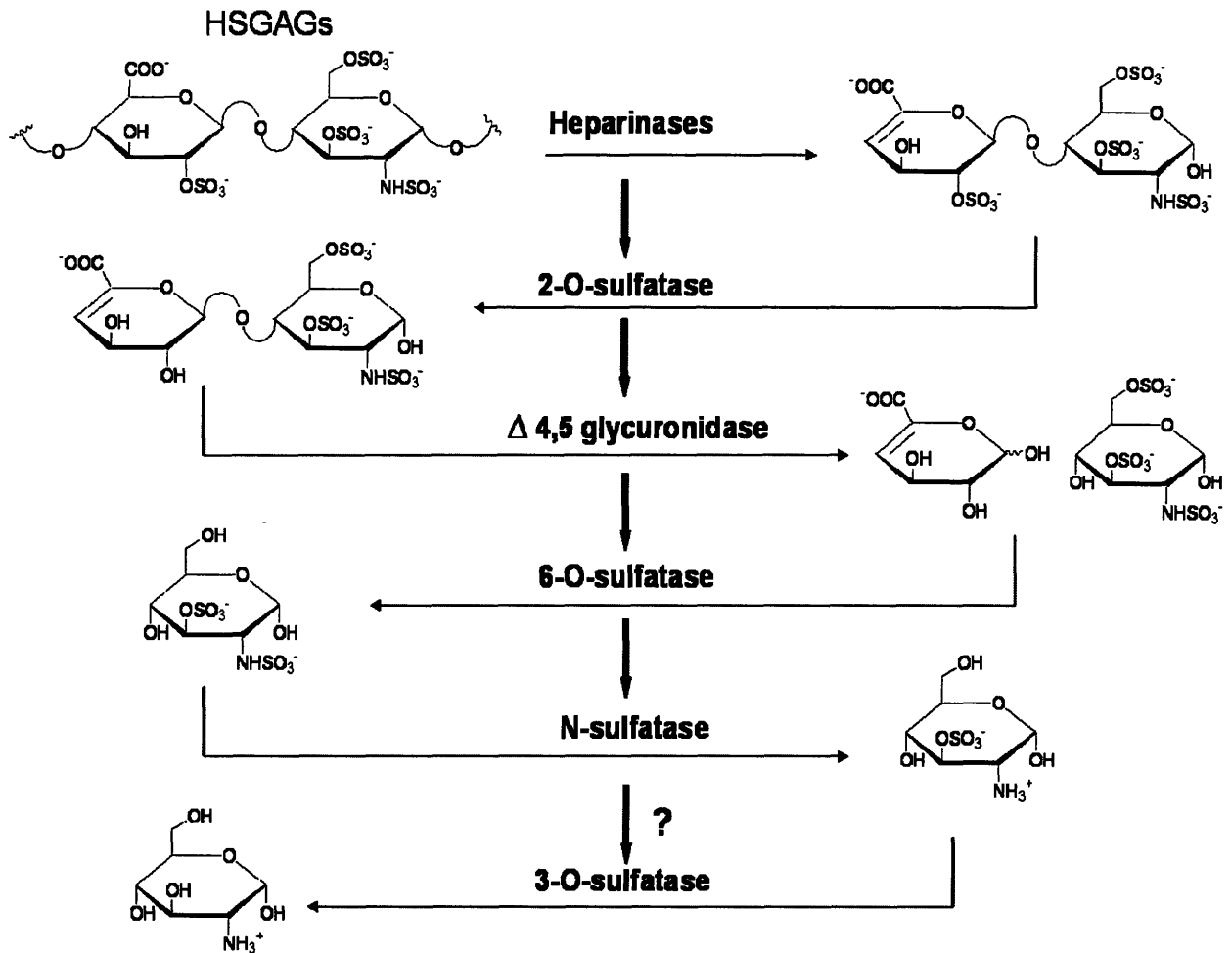
Figure 4.4. HSGAG degrading enzymes from *Flavobacterium heparinum*. Among these enzymes are heparinases, 2-O-sulfatase, Δ4,5-glycuronidase, 6-O-sulfatase, 3-O-sultafase and N-sulfatase. Each of these enzymes has unique substrate specificities and act on the HSGAG chain in a sequential order. The reactions catalysed by these enzymes and the sequential order are shown. It is not exactly known whether N-sulfatase or 3-O-sulfatase acts first.

CHAPTER 5

Molecular Cloning, Recombinant Expression and

Biochemical Characterization of the Heparin/Heparan Sulfate

Δ4,5 Glycuronidase from Flavo Heparinum

Synopsis: The soil bacterium Flavobacterium heparinum produces several enzymes that degrade heparan sulfate glycosaminoglycans (HSGAGs) in a sequence-specific manner. Among others, these enzymes include the heparinases and an unusual glycuronidase that hydrolyzes the unsaturated Delta4,5 uronic acid at the nonreducing end of oligosaccharides resulting from prior heparinase eliminative cleavage. We report here the molecular cloning of the Delta4,5 glycuronidase gene from the flavobacterial genome and its recombinant expression in Escherichia coli as a highly active enzyme. We also report the biochemical and kinetic characterization of this enzyme, including an analysis of its substrate specificity. We find that the Delta4,5 glycuronidase discriminates on the basis of both the glycosidic linkage and the sulfation pattern within its saccharide substrate. In particular, we find that the glycuronidase displays a strong preference for 1→ 4 linkages, making this enzyme specific to heparin/heparan sulfate rather than 1→ 3 linked glycosaminoglycans such as chondroitin/dermatan sulfate or hyaluronan. Finally, we demonstrate the utility of this enzyme in the sequencing of heparinase-derived HSGAG oligosaccharides.

CHAPTER 5

Molecular Cloning, Recombinant Expression and

Biochemical Characterization of the Heparin/Heparan Sulfate

Δ4,5 Glycuronidase from Flavo Heparinum

5.1 Introduction

Glycosaminoglycans (GAGs) are linear, acidic polysaccharides that exist ubiquitously in nature as residents of the extracellular matrix and at the cell surface of many different organisms (*1*). In addition to a structural role, GAGs act as critical modulators of a number of biochemical signaling events (*2*) requisite for cell growth and differentiation, cell adhesion and migration, and tissue morphogenesis. The heparan sulfate glycosaminoglycans (HSGAGs) are a particular class of GAGs that have emerged as key players in many different biological processes ranging from angiogenesis (*3*) and cancer biology (*4*) to microbial pathogenesis (*5*). HSGAGs have also recently been shown to play a fundamental role in multiple aspects of development (*6*).

HSGAGs are able to mediate these different biological events as a consequence of their considerable structural diversity (*7, 8*). This chemical heterogeneity is due largely to the differential sulfation pattern present for each disaccharide unit within the oligosaccharide chain. Adding to this structural variability is the C5 epimerization of the uronic acid that occurs at select chain positions. Taken together, these structural determinants impart a unique sequence to each HSGAG chain. In turn, this sequence dictates a critical HSGAG-protein interaction (*9*).

A determination of GAG sequence, however, has been technically challenging. To begin with, these polysaccharides are naturally present in very limited quantities, which, unlike other biopolymers such as proteins and nucleic acids, cannot be readily amplified. Second, due to their highly charged character and structural heterogeneity, HSGAGs are not easily isolated from biological sources in a highly purified state. Compounding this reality has historically been the lack of sequence-specific tools to cleave these linear polysaccharide chains in a manner analogous to DNA sequencing or restriction mapping. Recent chemical and enzymatic approaches have been used to cleave GAGs in a sequence-specific fashion. These HSGAG-degradation procedures have been used in conjunction with several analytical methods such as gel electrophoresis (10, 11) HPLC, and mass spectroscopy (12) to sequence the polysaccharides. We have refined this GAG sequencing approach by using tandem capillary electrophoresis/MALDI-MS analyses carried out within a bioinformatics framework (13). We have used this method to rapidly sequence biologically important HSGAGs, including saccharide sequences involved in modulating anticoagulation (14). The use of additional glycosaminoglycan-degrading enzymes will increase the analytical scope of this and other sequencing procedures.

Flavobacteria synthesize many glycosaminoglycan-degrading enzymes as an integral part of their catabolic life cycle (15). For this reason, this particular soil bacterium has been invaluable for identifying several enzymes potentially useful to the *in vitro* characterization of glycosaminoglycan structure. The biosynthesis of many of these enzymes can be induced several fold under growth conditions of high glycosaminoglycan content (16, 17). One of the first such heparin-degrading enzymes purified from adapted *F. heparinum*[21] cultures included

---

[1] While *Flavobacterium heparinum* is the historical (and commonly used) taxonomy, the current scientific name for this microorganism is *Cytophaga heparina*.

the heparin lyases (i.e., heparinases) (*16*), which cleave large chains of heparin and heparan sulfate into constituent oligosaccharides, each possessing a Δ4,5 unsaturated uronidate at the non-reducing end. In a likewise manner, an unusual glycuronidase that specifically hydrolyzes this unsaturated uronic acid was also identified and partially purified (*18*).

In this study, we report the first known molecular cloning of the Δ4,5 glycuronidase from the *F. heparinum* genome and its subsequent recombinant expression in *E. coli* as a soluble, highly active enzyme. We also present a biochemical characterization of the glycuronidase that includes a demonstration of its kinetic equivalence to the native enzyme and a description of its unusual substrate specificity.

5.2 Experimental Procedures

Chemicals and reagents

Unless otherwise stated, biochemicals were purchased from Sigma Aldrich Chemical

(St. Louis, MO). Unsaturated disaccharides were purchased from Calbiochem (San Diego,

CA). Reagents for λZAP II genomic library construction and screening were obtained from

Stratagene (La Jolla, CA). Additional molecular cloning reagents were obtained from the

manufacturers listed.

Bacterial strains and growth conditions

F. heparinum (*Cytophaga heparina*) was obtained as a lyophilized stock from

American Type Culture Collection (ATCC, Manassas, VA), stock no. 13125. *E. coli* strains

included TOP10 (Invitrogen, Carlsbad CA) for PCR cloning and subcloning and BL21(DE3)

(Novagen, Madison WI) for recombinant protein expression. Bacteriophage host strains

XL1-Blue MRF' and SOLR were obtained from Stratagene.

Large-scale Purification of Δ4,5 glycuronidase from Flavobacterium

The Δ4,5 glycuronidase was purified from a 10 liter fermentation culture essentially

as described for the *Flavobacterial* 2-O-sulfatase (*19*). Briefly, the cell pellet was harvested

by centrifugation, resuspended in 200 mL of sodium phosphate buffer, pH 7.0 and lysed by

repeated passage through an Aminco French pressure cell. The bacterial lysate was

chromatographed on a CM-Sepharose CL-6B column (17 x 3.2 cm), equilibrated with 10

mM sodium phosphate pH 6.5. Δ4,5 glycuronidase activity was eluted using a salt gradient of

0 to 0.3 M NaCl. Pooled fractions were diluted in sodium phosphate pH 6.7 buffer and applied to hydroxyapatite (Bio-Gel HTP, 16x2cm). Glycuronidase activity was eluted by application of a 0.01 to 0.5 M sodium phosphate gradient. Pooled enzyme fractions were desalted using Sephadex G-50M before applying to a Taurine-Sepharose CL-4B column (14x2cm). As a final purification step, the flow-through from Taurine Sepharose chromatography was reapplied to a CM Sepharose column and eluted by a NaCl gradient as described. Purification to homogeneity was verified by SDS-PAGE which indicated a single protein with an apparent molecular weight of approximately 45 kDa.

Generation of glycuronidase peptides and protein sequencing

Δ4,5 glycuronidase (approximately 1 nmol) was denatured in 50 μL 8M Urea, 0.4 M ammonium bicarbonate and reduced with 5 mM dithiothreitol at 65°C, cooled to room temperature, and alkylated with 10 mM iodoacetic acid for 20 minutes. The reaction was quenched with water. To this reaction was added 4% (w/w) of trypsin, and the digestion was carried out at 37°C overnight. The resulting peptides were separated using gradient reverse-phase HPLC (2-80% acetonitrile in 0.1% trifluoroacetic acid) over the course of 120 minutes. Tryptic peptides were monitored at 210 and 277nm. Peptide peaks were collected and sequenced using an on-line Model 120 phenylthiohydantoin-derivative analyzer (Biopolymers Laboratory, Massachusetts Institute of Technology).

Molecular cloning of the Δ4,5 glycuronidase gene from F. heparinum genomic DNA

Flavobacterial genomic DNA was isolated from a 10 mL *Flavobacterial* culture using the Qiagen DNeasy DNA purification kit according to the Manufacturer's instructions

for gram-negative bacteria (approximately 2 x 10⁹ cells per column). Following purification,

genomic DNA was ethanol precipitated and resuspended in TE buffer (10 mM Tris, pH 7.5, 1

mM EDTA) at 0.5 mg/mL. The quality of genomic DNA was confirmed

spectrophotometrically at 260/280 nm and by electrophoresis on a 0.5% agarose gel

The following degenerate primers were synthesized from peptides corresponding

initially to peaks 19 and 24:

5' GARACNCAYCARGGNYTNACNAAYGAR 3' (peak 19 forward),

5' YTCRTTNGTNARNCCYTGRTGNGTYTC 3' (peak 19 reverse);

5' AAYTAYGCNGAYTAYTAYTAY 3' (peak 24 forward);

5' RTARTARTARTCNGCRTARTT 3' (peak 24 reverse).

All four primers were screened in a PCR assay using all possible pairings (forward and

reverse). The PCR reaction conditions included 200 ng of genomic DNA, 200 picomoles for

each forward and reverse primer, 200 μM dNTPs, 1 unit of Vent DNA Polymerase (New

England Biolabs) in a 100 μl reaction volume. 35 cycles were completed using a 52°C

annealing temperature and 2 minute extensions at 72°C. A 450 bp product amplified using

primers 19 forward and 24 reverse was gel purified and subject to direct DNA sequencing

which confirmed the inclusion of translated sequence corresponding to peptide peaks 19 and

24 in addition to peak 12. The same DNA was also ³²P radiolabeled by random priming using

200 μCi α³²P[dCTP] at 6000 Ci/mmole (NEN, Boston, MA), 50-100 ng of DNA and the

Prime-it II random priming kit (Stratagene) (probe1). Unincorporated ³²P dNTPs were

removed by gel filtration using G-50 Quick-spin columns (Roche biochemicals, Piscataway

New Jersey).

DNA hybridization probe 2 was initially created by PCR as described above except

using degenerate primers 26 CARACNTAYACNCCNGGNATGAAY (peak 26 forward) and

20 picomoles of reverse, non-degenerate primer 54 (TTCATGGTCGTAACCGCATG).

Direct sequencing of this PCR fragment confirmed the presence of gene sequence

corresponding to peak 26 and peak 13 peptides. Δ4,5 5' specific DNA probe 3 used in DNA

southern hybridizations (below) was PCR amplified from genomic DNA using primers 68

(TATACACCAGGCATGAACCC)) and 74 (CCCAGTATAAATACTCCAGGT).

Plaque hybridization screening of F. heparinum genomic library

A λ ZAP II genomic library was constructed as described (20). The amplified library

$(1 \times 10^{10}$ pfu/mL) was plated out at approximately 1×10^6 pfu (~50,000 pfu/plate) on 100 x

150 mm LB plates. Plaque lifts on to nylon membranes (Nytran Supercharge, Schleicher and

Schuell) and subsequent hybridization screenings were completed using standard methods

and solutions (21). DNA was crosslinked to each filter by UV-irradiation (Stratagene

Stratalinker) for 30 seconds at 1200 joules/cm^3. Hybridizations were carried at 42°C using

10^7-10^8 cpm of radiolabeled probe (at approximately 0.25 ng/mL). Low stringency washes

were carried out at room temperature in 2X SSC, 0.1 % SDS; high stringency washes carried

out at 58-60°C in 0.2X SSC and 0.1% SDS. Hybridized plaques were visualized by phosphor

imaging (Molecular Dynamics) and/or ^{32}P autoradiography. Tertiary screens of positive

clones were completed and the recombinant phage was excised as a double-stranded

phagemid (pBluescript) using the Ex-Assist interference-resistant helper phage according to

the manufacturer's protocol (Stratagene). Recombinants were characterized by DNA

sequencing using both T7 and T3 primers.

Creation of a flavobacterium Bgl II-EcoR1 subgenomic library for isolation of the Δ4,5 5'

terminus

Southern DNA hybridizations were completed according to standard protocols (_21_)

using ^{32}P radiolabeled probe 3. 1 µg of genomic DNA was cut with 20 units of _Eco_ R1, _Bgl_

II, and _Hind_ III individually or as double digests and the restriction products were resolved

by gel electrophoresis on a 1 % agarose gel run in 1X TAE buffer. Based on this Southern

analysis, a 1.5 kb DNA fragment containing the 5' Δ4,5 terminus was identified. 10 µg of

Flavobacterial genomic DNA was therefore digested with _Bgl_-II-_Eco_ R1 and the DNA

resolved on a preparative gel run under identical conditions as those described for the

analytical gel. DNA ranging from approximately 1-2 kb was gel purified and ligated into

pLITMUS as a _Bgl_ II-_Eco_R1 cassette. Positive clones were identified by PCR colony

screening using primers 68 and 74 and confirmed by DNA sequencing.

PCR cloning of Δ4,5 gene and recombinant expression in E. coli

The full-length glycuronidase gene was directly PCR amplified from genomic DNA

using forward primer 85 TGTTCTAGA**CATATG**AAATCACTACTCAGTGC and reverse

primer 86 GT**CTCGAG**GATCCTTAAGACTGATTAATTGTT (with _Nde_ 1 and _Xho_1

restriction sites denoted in bold), 200 ng genomic DNA, and Vent DNA Polymerase for 35

cycles. dA overhangs were generated in a final 10 minute extension at 72°C using AmpliTaq

DNA polymerase (Applied Biosystems). PCR products were gel purified, ligated into the

TOPO/TA PCR cloning vector (Invitrogen), and transformed into One-shot TOP10

chemically competent cells. Positive clones were identified by blue/white colony selection

and confirmed by PCR colony screening. The 1.2 kb Δ4,5 gene was subcloned into

expression plasmid pET28a (Novagen) as an *Nde* 1-*Xho* 1 cassette. Final expression clones were confirmed by plasmid DNA sequencing.

For the expression of Δ4,5 glycuronidase beginning with M21 (Δ4,5$^{\Delta20}$), the forward primer 95 TGT TCT AGA **CAT ATG** ACA GTT ACG AAA GGC AA (also containing an *Nde* 1 restriction site near its 5' terminus) was used in place of primer 85 (above). 50 ng of the original expression plasmid pET28a/Δ4,5 was used as the DNA template in PCR reactions involving a total of 20 cycles. Otherwise, cloning was as described for the full-length gene. Both pET28a/Δ4,5 and pET28aΔ4,5$^{\Delta20}$ plasmids were transformed into BL21 (DE3) for expression as N-terminal 6X His-tagged proteins. 1 liter cultures were grown at room temperature (~20°C) in LB media supplemented with 40 μg/mL kanamycin. Protein expression was induced with 500 μM IPTG added at an A$_{600}$ of 1.0. Induced cultures were allowed to grow for 15 hours (also at room temperature).

Recombinant Δ4,5 glycuronidase purification

Bacterial cells were harvested by centrifugation at 6000 x g for 20 minutes and resuspended in 40 mL of binding buffer (50 mM Tris-HCL, pH 7.9, 0.5 M NaCl, and 10 mM imidazole). Lysis was initiated by the addition of 0.1 mg/mL lysozyme (20 minutes at room temperature) followed by intermittent sonication in an ice-water bath using a Misonex XL sonicator at 40-50% output. The crude lysate was fractionated by low-speed centrifugation (18,000 x g; 4°C; 15 minutes) and the supernatant was filtered through a 0.45 micron filter. The 6X-His tag Δ4,5 glycuronidase was purified by Ni^{+2} chelation chromatography on a 5 mL Hi-Trap column (Pharmacia Biotech, NJ) pre-charged with 200 mM NiSO$_4$ and subsequently equilibrated with binding buffer. Column washes included an intermediate step

with 50 mM imidazole. The Δ4,5 enzyme was eluted from the column in 5 mL fractions using high imidazole elution buffer (50 mM Tris-HCL, pH 7.9, 0.5 M NaCl, and 250 mM imidazole). Peak fractions were dialyzed overnight against 4 liters of phosphate buffer (0.1M sodium phosphate, pH 7.0, 0.5 M NaCl) to remove the imidazole.

The 6X His tag was effectively cleaved by adding biotinylated thrombin at 2 units/milligram of recombinant protein, overnight at 4°C with gentle inversion. Thrombin was captured by binding to streptavidin agarose at 4°C for two hours using the Thrombin Capture Kit (Novagen). The cleaved peptide MGSSHHHHHHSSGLVPR was removed by final dialysis against a 1000-fold volume of phosphate buffer.

Protein concentrations were determined by the Bio-Rad protein assay and confirmed by UV spectroscopy using a theoretical molar extinction coefficient ε=88,900 for Δ4,5[Δ20]. Protein purity was assessed by SDS-PAGE followed by Coomassie Brilliant Blue staining.

Computational methods

Signal sequence predictions were made by SignalP V1.1 using the von Heijne computational method (*22*) with maximum Y and S cutoffs set at 0.36 and 0.88, respectively. Glycuronidase multiple sequence alignments were made from select BLASTP database sequences (with scores exceeding 120 bits and less than 6% gaps) using the CLUSTAL W program (version 1.81) preset to an open gap penalty of 10.0, a gap extension penalty of 0.20, and both hydrophilic and residue-specific gap penalties turned on.

Assay for enzyme activity and determination of kinetic parameters

Standard reactions were carried out at 30°C and included 100 mM sodium phosphate buffer, pH 6.4, 50 mM NaCl, 500 μM disaccharide substrate and 200 nM enzyme in a 100 μl reaction volume. Hydrolysis of unsaturated heparin disaccharides was determined spectrophotometrically by measuring the loss of the $\Delta 4,5$ chromophore measured at 232 nm. Substrate hydrolysis was calculated using the following molar extinction coefficients empirically determined for each unsaturated disaccharide substrate: ΔUH_{NAc}, ε_{232}=4,524; $\Delta UH_{NAc,6S}$, ε_{232}=4,300; ΔUH_{NS}, ε_{232}=6,600; $\Delta UH_{NS,6S}$, ε_{232}=6,075; $\Delta UH_{NH2,6S}$, ε_{232}=4,826; $\Delta U_{2S}H_{NS,6S}$, ε_{232}=4,433. Initial rates (V_o) were extrapolated from linear activities representing <10% substrate turnover and fit to pseudo first-order kinetics. For kinetic experiments, disaccharide concentration for each respective substrate was varied from 48 to 400 μM concentrations. K_m and k_{cat} values were extrapolated from V_o vs. [S] curves fit to the Michaelis-Menten equation by a non-linear least squares regression.

For experiments measuring the relative effect of ionic strength on glycuronidase activity, the NaCl concentration was varied from 0.05 to 1 M in 0.1 M sodium phosphate buffer (pH 6.4), 200 μM $\Delta UH_{NS,6S}$ and 100 nM enzyme under otherwise standard reaction conditions. The effect of pH on catalytic activity was kinetically determined at varying $\Delta UH_{NS,6S}$ concentrations in 100 mM sodium phosphate buffer at pH 5.2, 5.6, 6.0, 6.4, 6.8, 7.2 and 7.8. Data were fit to Michaelis-Menten kinetics as described above and the relative k_{cat}/K_m ratios plotted as a function of pH.

130

Detection of Δ4,5 glycuronidase activity by capillary electrophoresis

200 μg of heparin (Celsus Laboratories) was subject to exhaustive heparinase

digestion as described (*13*) with certain modifications that included a 50 mM PIPES buffer,

pH 6.5 with 100 mM NaCl in a 100 μl reaction volume. Following heparinase treatment, 25

picomoles of Δ4,5 glycuronidase were added to one-half of the original reaction (pre-

equilibrated at 30°C). 20 μl aliquots were removed at 1 minute and 30 minutes and activity

quenched by heating at 95°C for 10 minutes. 20 μl of the minus Δ4,5 control (also heated for

10 minutes) was used as the 0 time point. Unsaturated disaccharide products were resolved

by capillary electrophoresis run for 25 minutes under positive polarity mode as previously

described (*12*).

Molecular mass determinations by MALDI-MS

The molecular weight of both the wild type and recombinantly expressed forms of the

enzyme were determined by matrix-assisted laser desorption ionization mass spectrometry

(MALDI-MS) using a Voyager Elite time-of-flight instrument (PerSeptive Biosystems,

Framingham, MA). Prior to mass analysis, the N-terminal histidine tag of the recombinant

protein was removed. For MALDI-MS analysis, 0.5 μl of an aqueous protein solution (0.1-1

mg/mL) was added to the target followed by the addition of 1μl of a saturated sinapinic acid

matrix solution in 50% acetonitrile/water. The sample was allowed to air dry and then placed

into the instrument. Mass spectra were collected in the linear mode using delayed extraction.

Instrument parameters were adjusted to maximize signal. All spectra were calibrated

externally using a mixture of ovalbumin and carbonic anhydrase (Sigma-Aldrich).

5.3 Results

Molecular cloning of the Δ4,5 glycuronidase gene from F. heparinum genome

To clone the Δ4,5 glycuronidase gene, we isolated a series of Δ4,5 glycuronidase–

derived peptides after a protease treatment of the purified enzyme. The native enzyme was

directly purified from fermentation cultures of *F. heparinum* using a 5-step chromatography

scheme as previously described (19). The extent of purity was ultimately characterized by

reverse phase chromatography, which indicated a single major peak (Figure 5.1A). We were

able to generate a number of peptides by a limit trypsin digestion of the purified enzyme. 26

peptide fragments were resolved by reverse phase chromatography (Figure 5.1B). From these

26, at least eight peptides (corresponding to major peaks 8, 12, 13, 19, 24, and 26) were of

sufficient yield and purity and were selected for protein sequence determination (Figure

5.1C).

Based on this information, we designed degenerate primers corresponding to peaks

19, 24, and 26. These primers were used to PCR amplify Δ4,5 specific sequences to be used

as suitable DNA hybridization probes for screening the *Flavobacterial* genomic library. A

combination of two primer pairs in particular (peak 19 forward and peak 24 reverse) gave a

discrete PCR product of approximately 450 base pairs. The translation of the corresponding

DNA sequence indicated that it contained the expected amino acid sequence corresponding

to peaks 19 and 24. The peak 12 peptide also mapped to this region. We used this PCR DNA

fragment in the initial plaque hybridizations. The most 5' terminal clone obtained in this

screening included approximately one-half of the predicted gene size corresponding to the

carboxy terminus of the putative ORF. Invariably, all of the isolated clones possessed an *Eco*

R1 site at their respective 5' termini. In an attempt to isolate a clone from the phage library possessing the other half of the gene, we rescreened additional plaques, this time using a second N-terminal specific DNA hybridization probe (probe 2) in tandem with the original probe 1. This second strategy also failed to yield any clones with the fully-intact Δ4,5 gene.

Alternate approaches were taken in an attempt to obtain the 5' terminus of the glycuronidase gene. The size of this missing region was estimated, based on the molecular weight of the native protein, to be approximately 45 amino acids (135 base pairs). We completed DNA southern analyses to identify potentially useful DNA restriction sites flanking the 5' end of the Δ4,5 gene. This restriction mapping ultimately involved the use of the *Eco* R1 site within the gene in conjunction with hybridization probe 3 (whose 3' end lies just 5' to this restriction site) to positively bias our search for the remaining amino terminus. Based on this refined map, we successfully isolated and subcloned an approximately1.5 kb *Bgl* II- *Eco* R1 Δ4,5 fragment into pLITMUS. The 5' terminus of the Δ4,5 gene was obtained from direct DNA sequencing of this subgenomic clone.

A summary of the full-length gene sequence (compiled from the two overlapping cloning methods) is depicted in Figure 5.2. The Δ4,5 coding sequence is comprised of 1209 base pairs corresponding to an ORF that encodes a 402 amino acid protein. The predicted molecular weight of 45,621 Daltons for the translated protein corresponds very well with an empirical molecular mass of 45,586 Daltons for the purified *Flavobacterial* enzyme determined by MALDI-MS. All of the peptides for which we obtained sequence information map to this Δ4,5 ORF. Based on further primary sequence analyses, we have also identified a likely bacterial signal sequence spanning amino acids 1-20 also possessing a putative cleavage site between residues G20 and M21 (Figure 5.3). The presence within this

hydrophobic amino terminus of an AXXA peptidase cleavage motif further supports this presumption (*24*).

A search for related sequences in the NCBI protein database led to several functionally related enzymes. In a multiple sequence alignment of our cloned enzyme with select unsaturated glucuronyl hydrolases, we found a homology that generally corresponded to upwards of 30% identity and nearly 50% similarity at the primary sequence level (Figure 5.4). This relatedness spanned most of the enzyme sequence, excluding the N-terminus. Based on this alignment, we found several highly conserved positions within the *F. heparinum* Δ4,5 glycuronidase that included several aromatic and acidic residues, e.g., W73, F77, Y86, F119, W162, W244, W250, Y253 W292, and Y379; D116, D174, D213, D287, D293, D305, D371 and D378. Other invariant amino acids of possible catalytic importance include H115, H201, H218 and basic residues R151, R211, and R246.

Recombinant expression and purification of the Δ4,5 glycuronidase

Using PCR, we cloned from the *F. heparinum* genome both the full-length enzyme and the "mature" enzyme lacking the N-terminal 20 amino acid signal sequence ($\Delta4,5^{\Delta20}$) into a T7-based expression plasmid. Cloning into pET28a permitted the expression of the glycuronidase as an N-terminal 6X His-tag fusion protein. Pilot expression studies focused on the full-length enzyme. In these initial experiments, we examined several different induction conditions such as temperature, time and length of induction, and even IPTG concentrations. In every case, the full-length enzyme was present nearly exclusively as an insoluble fraction. Attempts to purify the enzyme directly from inclusion bodies and then refold the protein were not successful. Solubility was partially achieved by a combined use of

detergents (e.g., CHAPS), increasing salt concentrations, and the addition of glycerol; the partially purified enzyme, however, was largely inactive. Conversely, recombinant expression of the enzyme lacking the presumed signal sequence yielded remarkably different results. In this case, soluble recombinant expression levels routinely reached several hundred milligrams of protein per liter of induced cells. The specific activity of this enzyme on the heparin disaccharide ΔUH_{NAc} was likewise robust (see below).

A summary of the expression and purification of $\Delta 4,5^{\Delta 20}$ is summarized in Figure 5.5 and Table 5.1. A two-step purification comprised of Ni^{+2} chelation chromatography followed by thrombin cleavage to remove the 6X His purification tag typically yielded over 200 mg of pure enzyme as assessed by SDS-PAGE followed by Coomassie Brilliant Blue staining. Most notably, the specific activity of the recombinant enzyme acting upon ΔUH_{NAc} far exceeded the values reported in the literature for the wild type enzyme (*18*). While removal of the 6X His tag from the N-terminus of the enzyme was unnecessary for optimal hydrolytic activity, the presence of the histidine tag did appear to instigate protein precipitation over an extended time period, especially at higher enzyme concentrations. This tag was, therefore removed for all subsequent biochemical experiments. In this manner, the recombinant protein was stable at 4°C for at least two weeks during which time it remained in solution at protein concentrations exceeding 10 mg/mL under the buffer conditions described.

A molecular mass of 44,209 Daltons was determined for the recombinant enzyme (i.e., $\Delta 4,5^{\Delta 20}$) by MALDI-MS. This empirically-established molecular weight is consistent with its theoretical value of 43,956 Daltons based on its amino acid composition. This value physically differs by 1377 Daltons in comparison to a molecular weight of 45,586 Daltons likewise measured for the native enzyme. This mass differential is largely accounted for by

the engineered removal in the recombinant protein of the 20 amino acid signal sequence. However, we cannot exclude the possibility of differential posttranslational modifications such as glycosylation largely accounting for the observed differences between the two enzyme populations. Unfortunately, chemical blocking precluded us from determining the N-terminal sequence of the native protein.

Biochemical conditions for optimal enzyme activity

To determine the optimal reaction conditions for $\Delta 4,5$ glycuronidase activity, we analyzed initial reaction rates as a function of buffer, pH, temperature, and ionic strength (Figure 5.6). For these experiments, we used the unsaturated, disulfated heparin disaccharide substrate $\Delta UH_{NS,6S}$. Based on what is known about the degradation of heparin/heparan sulfate-like glycosaminoglycans by *flavobacteria* (*25*) as well as initial biochemical characterization of this and related enzymes (*18, 26*), we hypothesized that an unsaturated heparin disaccharide would be a logical substrate for the $\Delta 4,5$ glycuronidase. Enzyme activity was routinely monitored by a loss of absorbance at 232 nm, corresponding indirectly to the hydrolysis of the uronic acid from the non-reducing end. The observed decrease in UV absorption is presumably due to the chain opening of the unsaturated uronidate monosaccharide concomitant with keto-enol isomerization to form the α–keto acid, resulting in the loss of the λ_{232}-absorbing carbon-carbon bond (*27*).

Under the examined conditions, we observed a NaCl concentration-activity dependence with optimal enzymatic activity occurring at salt concentrations between 50 and 100 mM. At concentrations exceeding 100 mM NaCl, the enzyme demonstrated a significant decrease in activity that was sharply dependant on the ionic strength (Figure 5.6A), i.e., with

approximately 60% inhibition occurring at 250 mM NaCl relative to 100% activity measured at 100 mM NaCl.

The observed pH profile for the glycuronidase was bell-shaped (Figure 5.6B) with a pH optimum of 6.4. This result is consistent with previously published results (18). In addition, temperature titration experiments indicated that maximal enzymatic activity for the $\Delta4,5$ glycuronidase was achieved at 30°C. Interestingly, initial reaction rates were significantly reduced at higher temperatures, especially at 42°C (Figure 5.6C). Preincubation experiments at 30, 37, and 42°C to assess enzyme stability indicated no significant change in enzymatic activity when the enzyme was subsequently measured under the standard 30°C reaction conditions. The results of these experiments suggest that the observed decrease in the enzymatic reaction rate at higher temperatures is not due to a thermal lability of the $\Delta4,5$ glycuronidase, rather it reflects some intrinsic property of its catalytic function.

As a final variable for optimizing $\Delta4,5$ glycuronidase *in vitro* reaction conditions, we also considered any requirement for divalent metal ions, including Ca^{++}, Mg^{++}, Mn^{++}, and others. We found no evidence that any metal cofactor is either required for catalysis or activates the enzyme to any appreciable extent.

Having established the reaction conditions for optimal $\Delta4,5$ glycuronidase activity, we next compared the specific activity of the recombinant enzyme ($\Delta4,5^{\Delta20}$) to the native enzyme purified directly from *F. heparinum*. The activities of both enzyme fractions were measured in parallel under identical reaction conditions. In this comparison, the recombinant $\Delta4,5$ possessed a 1.5-fold higher initial velocity relative to the native enzyme. These results demonstrate that the recombinant enzyme is comparable in activity to the native form of the

137

enzyme and possesses the added benefit that over 1000-fold more enzyme can be obtained

per liter of culture as compared to the wild type.

Δ4,5 glycuronidase substrate specificity

Given the fact that the recombinant form of the Δ4,5 glycuronidase possessed similar

attributes to the native form and is undoubtedly free of contaminating HSGAG-degrading

activities, this form of the enzyme was used to characterize the substrate specificity of the

Δ4,5 glycuronidase using various unsaturated glycosaminoglycan disaccharide substrates.

These substrates examined included heparin, chondroitin, and hyaluronan disaccharides. In

particular, we considered the possibility of any structural discriminations pertaining to the

glycosidic linkage position ($1 \rightarrow 4$ vs. $1 \rightarrow 3$) and relative sulfation state within the

disaccharide. For each substrate, kinetic parameters were determined based on substrate

saturation profiles that fit Michaelis-Menten assumptions (Figure 5.7). These kinetic values

are listed in Table 5.2. For the heparin disaccharides, k_{cat} values varied significantly from

approximately 2 to 15 sec^{-1}, while the apparent K_m values for each respective disaccharide

were comparable, ranging from approximately 100 μM to 300 μM. The heparin disaccharide

$\Delta U_{2S}H_{NS6S}$ was unequivocally not a substrate at any of the concentrations studied, even

following an extended incubation time spanning several hours. For those unsaturated heparin

disaccharides that were hydrolyzed under the conditions tested and for which kinetic

parameters could be determined, a highly interesting substrate preference was apparent. In

this hierarchy, the two disaccharides ΔUH_{NAc} and $\Delta UH_{NAc,6S}$ were clearly the best substrates.

In comparison, $\Delta UH_{NH2,6S}$ and ΔUH_{NS} were relatively poor substrates. The kinetic values for

$\Delta UH_{NS,6S}$ fell roughly in the middle between these two boundaries.

None of the non-heparin disaccharides were hydrolyzed under the conditions established to measure substrate kinetics. These results indicate an unequivocal discrimination of the Δ4,5 glycuronidase with regards to linkage position, with a strong preference for 1→4 versus 1→3 linkages. At the same time, 1→3-containing disaccharides were hydrolyzed slowly to varying degrees when the enzyme reactions were conducted over a much longer time course (> 12 hours) and at considerably higher enzyme concentrations (data not shown). After approximately 18 hours, greater than 80% of a monosulfated chondroitin disaccharide ($\Delta UGal_{NAc,6S}$) was hydrolyzed, whereas the non-sulfated chondroitin ($\Delta UGal_{NAc}$) and the hyaluronan disaccharide (ΔUH) were still present at approximately 40% and 65%, respectively. The apparently positive effect of chondroitin 6-O-sulfation within the galactosamine is consistent with our results for the heparin substrates and provides further evidence for the influence of this position in dictating substrate specificity.

Based on the clear, kinetically defined substrate specificity for the unsaturated disaccharides, we set out to validate these results while, at the same time, investigating the utility of the Δ4,5 glycuronidase as an enzymatic tool for HSGAG compositional analysis. In this manner, the Δ4,5 glycuronidase could potentially be very useful in assessing the composition of disaccharides resulting from prior heparinase treatment of heparin/heparan sulfate. For this particular experiment, we pre-treated 200 μg of heparin with a heparinase cocktail. This exhaustive digestion was then directly followed by a relatively short (1 minute) or long (30 minute) Δ4,5 glycuronidase treatment carried out under optimal reaction conditions. The unsaturated disaccharide products were then resolved by capillary electrophoresis. The electrophoretic mobility profile for the Δ4,5 glycuronidase-treated

saccharides were then directly compared to the untreated control (i.e., heparinase treatment only) run under identical conditions (Figure 5.8). Seven disaccharide peaks were present in the capillary electrophoretogram corresponding to the heparinase only control (Figure 5.8A). A structural assignment for each one of these peaks was made based on previously established compositional analyses. Consistent with the observed substrate specificity for the $\Delta 4,5$ glycuronidase, the relative area of the peaks corresponding to 2-O-containing disaccharides remained unchanged over the time course of the experiment (>18 hours). On the other hand, peaks corresponding to disaccharides lacking the 2-O sulfate were eliminated to various extents over the course of the experiment. The relative rates of their disappearances corresponded exactly to their substrate preference as determined in the previous kinetic experiment. $\Delta UH_{NAc,6S}$ (peak 8) was essentially hydrolyzed within one minute; $\Delta UH_{NS,6S}$ (peak 4) and ΔUH_{NS} (peak 6) were hydrolyzed approximately 75 % and 50% , respectively. These two latter substrates were completely depleted by 30 minutes.

In addition to the assigned disaccharides, the $\Delta 4,5$ glycuronidase also acted on a heparinase-generated tetrasaccharide (identified as peak 2 in Figure 5.8). The assignment of Peak 2 as a tetrasaccharide was confirmed by MALDI-MS, which indicated a mass of 1036.9 Daltons that corresponds to a singly acetylated tetrasaccharide with four sulfates. Disaccharide analysis of this tetrasaccharide further indicated that it does not contain a 2-O sulfate at the non-reducing end. The $\Delta 4,5$ enzyme hydrolyzed approximately one-half of the unsaturated starting material after one minute. The relative rate of hydrolysis for this tetrasaccharide roughly corresponded to the rate observed for the disaccharide ΔUH_{NS} (peak 6). This exciting result clearly indicates that a longer chain saccharide such as a tetrasaccharide is in fact a substrate for the $\Delta 4,5$ catalyzed hydrolysis.

5.4 Discussion

In this paper, we report the first-known molecular cloning of the Δ4,5 glycuronidase gene from *F. heparinum* as well as its recombinant expression in *E. coli* as a highly active enzyme. We also report a biochemical characterization of its substrate specificity. Such a description provides important insight into possible structure-function relationships underlying its unusual catalytic function. It also demonstrates that the Δ4,5 glycuronidase, like the other GAG-degrading enzymes we have cloned (*20, 28*), cleaves oligosaccharides in a non-random fashion. For this reason, we feel that the Δ4,5 glycuronidase is useful as a complementary enzymatic tool in our PEN-MALDI method for sequencing complex polysaccharides (*13*).

Our attempts to recombinantly express the intact glycuronidase in *E. coli* as a soluble, active protein were unsuccessful. Several attempts to refold the protein from bacterial inclusion bodies also failed. The most likely explanation for this insolubility points to the presence of a very hydrophobic region within the wild-type protein sequence spanning the first 20 amino acids. This N-terminal sequence is also predicted to comprise a cleavable prokaryotic signal sequence with the most likely cleavage site occurring between position G20 and M21 (Figure 5.3). Within this sequence, we also find the alanine repeat AXXAXXAXXXXA that may serve as part of the actual cleavage recognition sequence (*24*). This signal peptide would indicate a periplasmic location for the glycuronidase with the N-terminus of the secreted (mature) protein beginning with M21. We recombinantly expressed this mature variant (Δ4,5$^{\Delta 20}$) in which the signal sequence was replaced entirely by an N-terminal 6X His purification tag.

Even in our initial purifications, it was evident that our cloned Δ4,5 glycuronidase possessed a specific activity that far exceeded the activity rates reported for the *Flavobacterial* enzyme in the early literature (*18*). These early studies however used only partially purified enzyme under much different reaction conditions that included milligram quantities of both substrate and enzyme. In a direct and more rigorous comparison between the recombinant and native enzymes, we find that the recombinant enzyme ($\Delta4,5^{\Delta20}$) possessed a roughly three-fold higher specific activity relative to the native *Flavobacterial* enzyme when measured under identical reaction conditions. From these results, it is clear that the activity of cloned enzyme is not compromised whatsoever by its recombinant expression in *E. coli*. Rather, the observed rate discrepancy is likely due to subtle differences in protein purification and storage and not to any intrinsic enzymatic properties.

The recombinant Δ4,5 glycuronidase exhibits a sharp ionic strength dependence as indicated by the shape of the NaCl titration curve (Figure 5.6A). These results are interesting given both the ionic character of the disulfated heparin disaccharide used in this experiment as well as the many ionic residues present within the enzyme that may function in substrate binding and/or catalysis; many of these charged residues are conserved in structurally and functionally related enzymes (Figure 5.4). From a substrate perspective, all of the unsaturated disaccharides examined possess a negative charge (at pH 6.4) due to the C6 carboxylate of the uronic acid. It is possible that this acid acts as a critical structural determinant, especially given its proximity to the Δ4,5 bond. Charge neutralization of 6-O sulfate (e.g., in ΔUH_{NS6S}) could possibly be another contributing factor. From the enzyme perspective, the recombinant glycuronidase ($\Delta4,5^{\Delta20}$) does possess 47 basic residues (theoretical pI of 8.5), including R151, R211, and R246 whose positions are invariantly conserved among the different

glycuronidases examined. These basic amino acids may possibly interact with the uronic acid carboxylate. At the same time, Δ4,5 also possesses 44 acidic residues. At least ten of these positions are highly conserved. Charge masking of some of these ionic residues (either acidic or basic) by increasing salt concentration might interfere with enzymatic activity. A similar observation of this ionic strength dependency has been made for the heparinases (29).

We also observed a bell-shaped pH profile with a 6.4 optimum. The 6.4 pH optimum generally agrees with results originally reported for the *F. heparinum* Δ4,5 as well as for more recent results published for an unsaturated glucuronyl hydrolase purified from *Bacillus sp.* GL1 (26). This result logically implicates one or more histidine residues functioning in catalysis. While there are 11 histidines present within the primary sequence, only three histidines (H115, H201, and H218) appear to be highly conserved. The precise catalytic function of these histidines remains to be determined. Interestingly, catalytically critical histidines also exist in all three heparin lyases (30) as well as chondroitin AC lyase (31) from *Flavobacterium heparinum*. While these latter two classes of enzymes cleave glycosaminoglycans by a somewhat different mechanism compared with the Δ4,5 glycuronidase (i.e., β-elimination vs. hydrolysis), all three enzymes would presumably involve acid-base catalysis *viz* the imidazole ring of the histidine.

We approached the question of substrate specificity from three structural perspectives: (1) the nature of the glycosidic linkage; (2) the relative sulfation pattern of the unsaturated disaccharide; and (3) the role of saccharide chain length (e.g., di- vs. tetrasaccharide). As expected, our results indicate that for the recombinant Δ4,5 glycuronidase, there is an unambiguous preference for the 1→4 linkage over the 1→3 linkage making heparin/heparan sulfate rather than chondroitin/dermatan and/or hyaluronan

the best substrate. This observation agrees with earlier, published reports (*23*) likewise indicating the critical importance of the 1→4 linkage position. It also clearly distinguishes this particular *Flavobacterial* unsaturated glycuronidase from the chondroitin glycuronidase isolated from the same bacterium; the latter enzyme has been shown to cleave only 1→3 linked disaccharides (*32*). It should be noted, however, that while this linkage position is clearly important for heparin/heparan sulfate Δ4,5 glycuronidase substrate discrimination, it is not absolute. Both chondroitin and hyaluronan Δ4,5 disaccharides were hydrolyzed, albeit at much slower rates and using higher enzyme concentrations than were required to hydrolyze heparin disaccharides.

We also present a kinetic pattern of the Δ4,5 glycuronidase with regards to the specific sulfation within a heparin disaccharide. First and foremost, we find that unsaturated saccharides containing a 2-O-sulfated uronidate (ΔU_{2S}) at the non-reducing end are uncleavable by the Δ4,5 glycuronidase. Furthermore, the inability of a 2-O-sulfated disaccharide to competitively inhibit the hydrolysis of non 2-O-containing disaccharide substrates (such as ΔUH_{NAc}) further suggests that the presence of a 2-O sulfate precludes binding of this saccharide to the enzyme. Therefore, 2-O sulfation, along with linkage position, is another unequivocal structural constraint. Based on these observations, it is also clear that both *in vivo* and *in vitro*, the 2-O position of the unsaturated uronidate must first be desulfated if the Δ4,5 glycuronidase is to subsequently hydrolyze the glycosidic bond.

In considering the effect of specific sulfate groups present on the glucosamine, the enzyme may be loosely summarized as having a graded preference for 6-O-sulfation but a clear selection against unsubstituted or sulfated amines. We must emphasize that this hierarchy is not an absolute distinction given the fact that the enzyme cleaved all the non 2-

O-containing heparin disaccharides that we examined. Instead, it is based on relative kinetic

parameters (Table 5.2). This apparent substrate discrimination at the N and 6 positions of the

glucosamine appears to be somewhat contextual, especially in the case of 6-O-sulfation. That

is, while 6-O sulfation may bestow a favorable selectivity to a saccharide substrate, this

positive effect is offset by the presence of a deacetylated amine (e.g., $\Delta UH_{NAc,6S}$ vs.

$\Delta UH_{NH2,6S}$ or $\Delta UH_{NS,6S}$). While the exact reason for this structural discrimination needs to be

investigated, it is tempting to speculate a positive role of the acetate itself, e.g., in satisfying

an important van der Waals interaction.

We feel that the unequivocal bias the $\Delta 4,5$ demonstrates against 2-O-sulfated

uronidates along with a so-called "N-position" discrimination for the glucosamine can be

exploited for use of the glycuronidase as an analytical tool for the compositional analyses of

heparinase-treated oligosaccharides. In fact, we were able to predict the extent and relative

rates by which specific disaccharide species would disappear (i.e., due to the glycuronidase-

dependent loss of absorbance at 232 nm.), based entirely on our kinetically defined substrate

specificity determinations (Figure 5.8). As expected, all 2-O-sulfate containing disaccharides

were refractory to hydrolysis by the $\Delta 4,5$ glycuronidase. On the other hand, the remaining

disaccharides were hydrolyzed in a time-dependent fashion that corresponded to their relative

substrate specificities (i.e., $\Delta UH_{NAc,6S} > \Delta UH_{NS,6S} > \Delta UH_{NS}$).

From this experiment, another important observation was made, namely that the $\Delta 4,5$

glycuronidase also hydrolyzes $\Delta 4,5$ unsaturated tetrasaccharides. It is also very interesting to

note that this particular tetrasaccharide is as good of a substrate as the disaccharide ΔUH_{NS}.

This observation may argue against a substrate discrimination used by the enzyme that is

negatively based on increasing molecular weight (at least up to a tetrasaccharide) as was first

reported (*23*). It also leaves open the possibility that the Δ4,5 glycuronidase interacts predominantly with the unsaturated disaccharide unit residing at the non-reducing end. The actual effect of oligosaccharide chain length on relative glycuronidase activity has to be more thoroughly examined. In particular, it will be useful to address structure-activity relationships as they pertain to internal saccharide positions.

The Δ4,5 unsaturated glycuronidase is an interesting and unusual enzyme whose catalytic mechanism merits further investigation. The work presented here is foundational to future studies relating molecular structure to this function. At the same time, its unique substrate specificity makes this enzyme an excellent tool for characterizing HSGAG composition and fine structure.

5.5 References

1. Habuchi, O. (2000) *Biochim Biophys Acta 1474*, 115-27.
2. Tumova, S., Woods, A., and Couchman, J. R. (2000) *Int J Biochem Cell Biol 32*, 269-88.
3. Folkman, J., Taylor, S., and Spillberg, C. (1983) *Ciba Found Symp 100*, 132-49.
4. Blackhall, F. H., Merry, C. L., Davies, E. J., and Jayson, G. C. (2001) *Br J Cancer 85*, 1094-8.
5. Shukla, D., Liu, J., Blaiklock, P., Shworak, N. W., Bai, X., Esko, J. D., Cohen, G. H., Eisenberg, R. J., Rosenberg, R. D., and Spear, P. G. (1999) *Cell 99*, 13-22.
6. Perrimon, N., and Bernfield, M. (2000) *Nature 404*, 725-8.
7. Lyon, M., Rushton, G., Askari, J. A., Humphries, M. J., and Gallagher, J. T. (2000) *J Biol Chem 275*, 4599-606.
8. Esko, J. D., and Lindahl, U. (2001) *J Clin Invest 108*, 169-73.
9. Conrad, H. (1998) *Heparin-binding proteins*, Academic press.
10. Drummond, K. J., Yates, E. A., and Turnbull, J. E. (2001) *Electrophoresis 22*, 304-10.
11. Calabro, A., Midura, R., Wang, A., West, L., Plaas, A., and Hascall, V. C. (2001) *Osteoarthritis Cartilage 9*, S16-22.
12. Rhomberg, A. J., Ernst, S., Sasisekharan, R., and Biemann, K. (1998) *Proc Natl Acad Sci U S A 95*, 4176-81.
13. Venkataraman, G., Shriver, Z., Raman, R., and Sasisekharan, R. (1999) *Science 286*, 537-42.
14. Shriver, Z., Raman, R., Venkataraman, G., Drummond, K., Turnbull, J., Toida, T., Linhardt, R., Biemann, K., and Sasisekharan, R. (2000) *Proc Natl Acad Sci U S A 97*, 10359-64.
15. Payza, A. N. a. K., E.D. (1956) *J. Biol. Chem. 223*, 853-858.
16. Linker, A., and Hovingh, P. (1965) *J Biol Chem 240*, 3724-8.
17. Dietrich, C. P. (1969) *Biochem J 111*, 91-5.
18. Warnick, C. T., and Linker, A. (1972) *Biochemistry 11*, 568-72.
19. McLean, M. W., Bruce, J. S., Long, W. F., and Williamson, F. B. (1984) *Eur J Biochem 145*, 607-15.
20. Sasisekharan, R., Bulmer, M., Moremen, K. W., Cooney, C. L., and Langer, R. (1993) *Proc Natl Acad Sci U S A 90*, 3660-4.
21. *Current Protocols in Molecular Biology* (1987), John Wiley and Sons, New York.
22. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) *Protein Eng 10*, 1-6.
23. Hovingh, P., and Linker, A. (1977) *Biochem J 165*, 287-93.
24. von Heijne, G. (1988) *Biochim Biophys Acta 947*, 307-33.
25. Linker, A. (1979) *Biochem J 183*, 711-20.
26. Hashimoto, W., Kobayashi, E., Nankai, H., Sato, N., Miya, T., Kawai, S., and Murata, K. (1999) *Arch Biochem Biophys 368*, 367-74.
27. Linker, A., Hoffman, P., Meyer, K.,Sampson, P., and Korn, E. D. (1960) *J. Biol. Chem 235*, 3061-3065.
28. Pojasek, K., Shriver, Z., Kiley, P., Venkataraman, G., and Sasisekharan, R. (2001) *Biochem Biophys Res Commun 286*, 343-51.

29.   Lohse, D. L., and Linhardt, R. J. (1992) *J Biol Chem 267*, 24347-55.
30.   Pojasek, K., Shriver, Z., Hu, Y., and Sasisekharan, R. (2000) *Biochemistry 39*, 4012-9.
31.   Huang, W., Boju, L., Tkalec, L., Su, H., Yang, H. O., Gunay, N. S., Linhardt, R. J., Kim, Y. S., Matte, A., and Cygler, M. (2001) *Biochemistry 40*, 2359-72.
32.   Gu, K., Linhardt, R. J., Laliberte, M., and Zimmermann, J. (1995) *Biochem J 312*, 569-77.

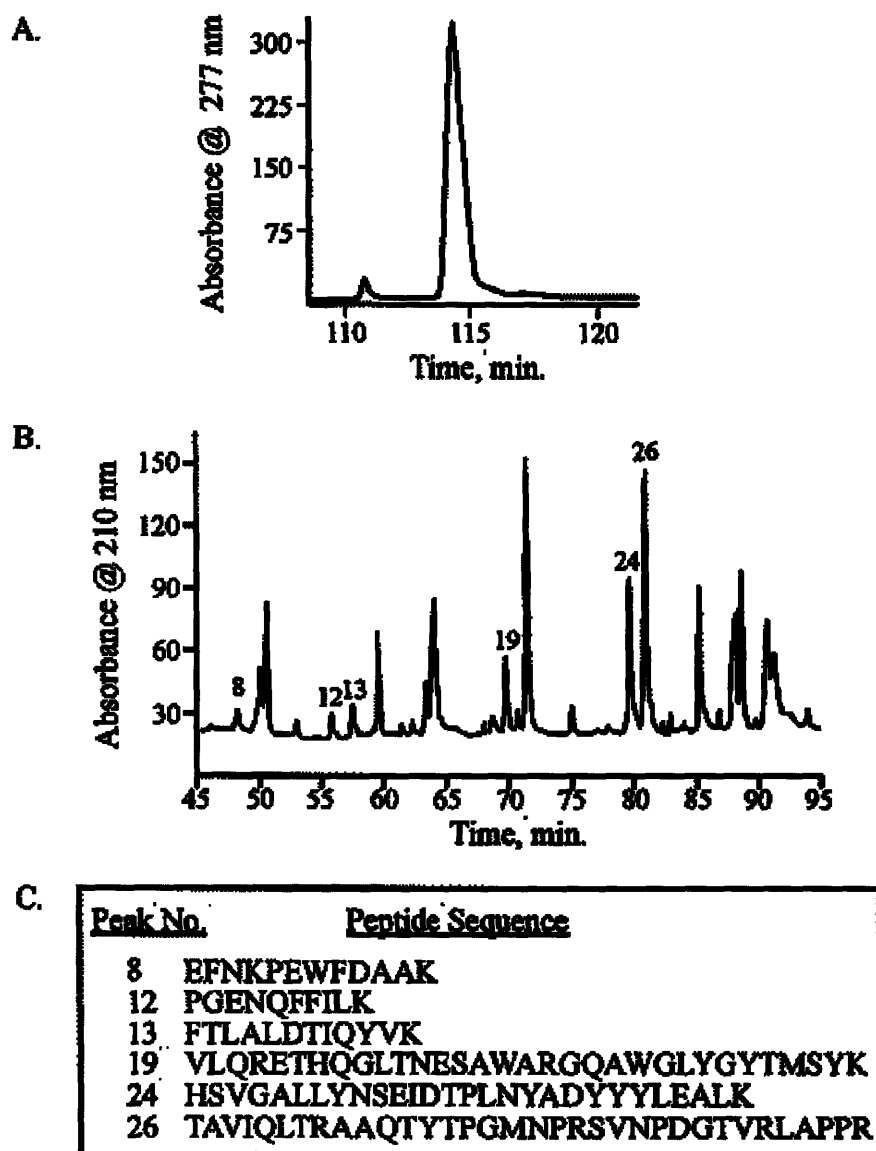**Figure 5.1.** Purification of Δ4,5 glycuronidase from Flavobacterium and resultant proteolytic products. A. C18 chromatogram of the purified enzyme. B. Purification of Δ4,5 peptides by reverse phase HPLC following trypsinization of the native protein. C. Amino acid sequence of select peptides isolated in B.
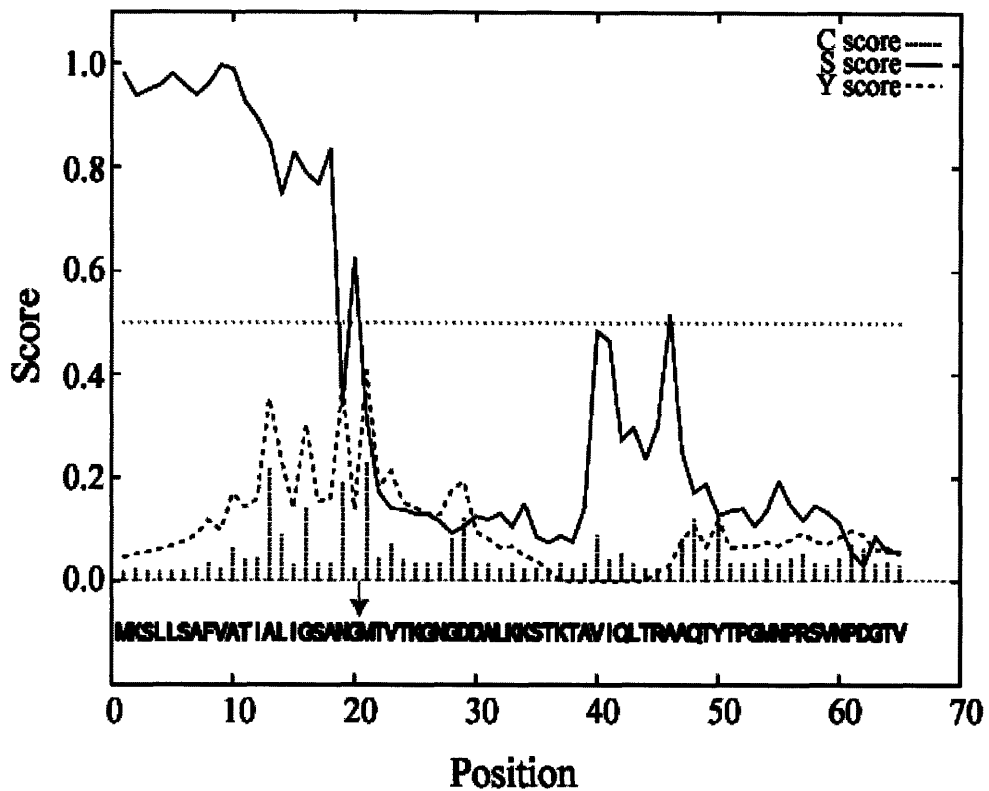
```
aagcggttttcctccgggcatcagacacaactccttttgtagttgtattttgggcatagagcgtaaggttaccc
gcaccaataattaatagcagtaaaaatctattcataacattttttttagttttacataatttggtttattgacaaa
aacgaatgtatagcactccaaccagaatgaatttaggtattgatcaaaatgctatcacttttgattaatcactca
ttatttcacaaaaatagatgctgtaatccaaaaaaaatcaggcattcaccaattcctaaatgaatatatccgata
cttaaagccttacatcttttacactccctaccttcatcattattaaaaaacaaaacttagatcggaaaataatt
atgaaatcactactcagtgcgtttgttgcgactattgcattaataggatctgcaaacgggatgacagttacgaaa
```

```
1    M  K  S  L  L  S  A  F  V  A  T  I  A  L  I  G  S  A  N  G↓ M  T  V  T  K
     ggcaacggcgatgactggttaaagaaatcaactaaaaccgcagtaatacagttaacacgggctgcacaaacctat
26   G  N  G  D  D  W  L  K  K  S  T  K  ████████████████████████████████████████
     acaccaggcatgaacccaaggtctgtcaatccggacgggacggtaaggctggcccctccccgcgactggaccaca
51   ████████████████████████████████████████████████████████████  P  R  D  W  T  T
     ggttttttcccgggaacgttgtggtatggttatgaactatcgggcgataaaaacctggcggccgaagccaaaaga
76   G  F  F  P  G  T  L  W  Y  G  Y  E  L  S  G  D  K  N  L  A  A  E  A  K  R
     tttacccttgccttagatacgatacaatatgttaaagatacgcacgacctgggctttatgttgtattgttcttat
101  ████████████████████████████████████████████  D  T  H  D  L  G  F  M  L  Y  C  S  Y
     ggcaatgcctaccgtgtaaccggagacaagatttacctgaagccattagaaacggtgcggccaatttatatgcc
126  G  N  A  Y  R  V  T  G  D  K  I  Y  L  K  P  L  E  N  G  A  A  N  L  Y  A
     cgtttcaataaaaaagtaggggccatccgctcatgggatttcggacactggcaatttccggtaattatagacaac
151  R  F  N  K  K  V  G  A  I  R  S  W  D  F  G  H  W  Q  F  P  V  I  I  D  N
     ctgatgaacctggagtatttatactgggcaggaaaggaattcaataagccagaatggttcgatgctgctaaaaca
176  L  M  N  L  E  Y  L  Y  W  A  G  K  ██████████████████████████████████████████  T
     catgcggttacgaccatgaaaaaccatttcagaaagattatagttcttaccatgtgatcagttacgataccctg
201  H  A  V  T  T  M  K  N  H  F  R  K  D  Y  S  S  Y  H  V  I  S  Y  D  T  L
     tctggaaaagtactgcaacgtgaaacccatcagggacttaccaacgaatcggcctgggcacgggggcaagcctgg
226  S  G  K  ████████████████████████████████████████████████████████████████████
     ggactttacggctataccatgagctataaggatacgaaagacaaaaaattcatcgaacatgcagagcatatcgct
251  ██████████████████████████████████  Y  K  D  T  K  D  K  K  F  I  E  H  A  E  H  I  A
     gctttcatcatgaaccaccctgcaatgccggcagataaaattccactttgggactttgatgtccacaaccgcgac
276  A  F  I  M  N  H  P  A  M  P  A  D  K  I  P  L  W  D  F  D  V  H  N  R  D
     aggtcgccaagggatgcttctgctgctgcagtaattgcttcagccctgctagacctgagcacgcaggtaaaagat
301  R  S  P  R  D  A  S  A  A  A  V  I  A  S  A  L  L  D  L  S  T  Q  V  K  D
     ggtcagaaatattttaaatttgccgaggatatcctgaaaacattgtcatcagatgaataccttggcgaaacccggc
326  G  Q  K  Y  F  K  F  A  E  D  I  L  K  T  L  S  S  D  E  Y  L  A  K  ██████
     gagaaccagttttttatattgaaacatagtgtaggtgcattgctgtacaattcggaaatcgatacacctttgaat
351  ██████████████████████████████████████████████████████████
     tatgccgactattactatctggaggctttaaaacgctatgcagagattaaaaaaattgacctgaaaacaattaat
376  ████████████████████████████████████████████████  R  Y  A  E  I  K  K  I  D  L  K  T  I  N
     cagtcttaatttttaaaccaacagcattataccaaaacagcaaaggctaccagggttggtagcctttgctgttttg
401  Q  S  STOP
     atagtttttaaaagaatttaaatcagtattgaaacattttggcagggattattttttttaatttgttgaacaagaatt
     aagaataaat
```

Figure 5.2. Δ4,5 glycuronidase gene sequence. Shown here are both the coding and flanking DNA sequences of the full length sequence. The coding sequence of 1209 base pairs contains an ORF encoding a putative protein of 402 amino acids. Initiation and termination codons are highlighted in bold. Possible Shine-Dalgarno (SD) sequence is boxed. Presumed signal sequence is underlined and its cleavage site delimited by a vertical arrow. Eco R1 restriction site is double-overscored. Also shown are the degenerate primer pairs (shown as arrows) used to PCR amplify DNA hybridization probes 1 and 2 as well as the relative positions of purified Δ4,5 peptides (shaded in gray) for which direct sequence information was obtained.

150

Figure 5.3. Assignment of a putative Δ4,5 glycuronidase signal sequence. Theoretical signal sequence determination using amino acids 1-65. Indices were calculated using SignalP V.1.1 using networks trained on gram-negative bacteria. A vertical arrow indicates putative cleavage site located between G20 and M21.
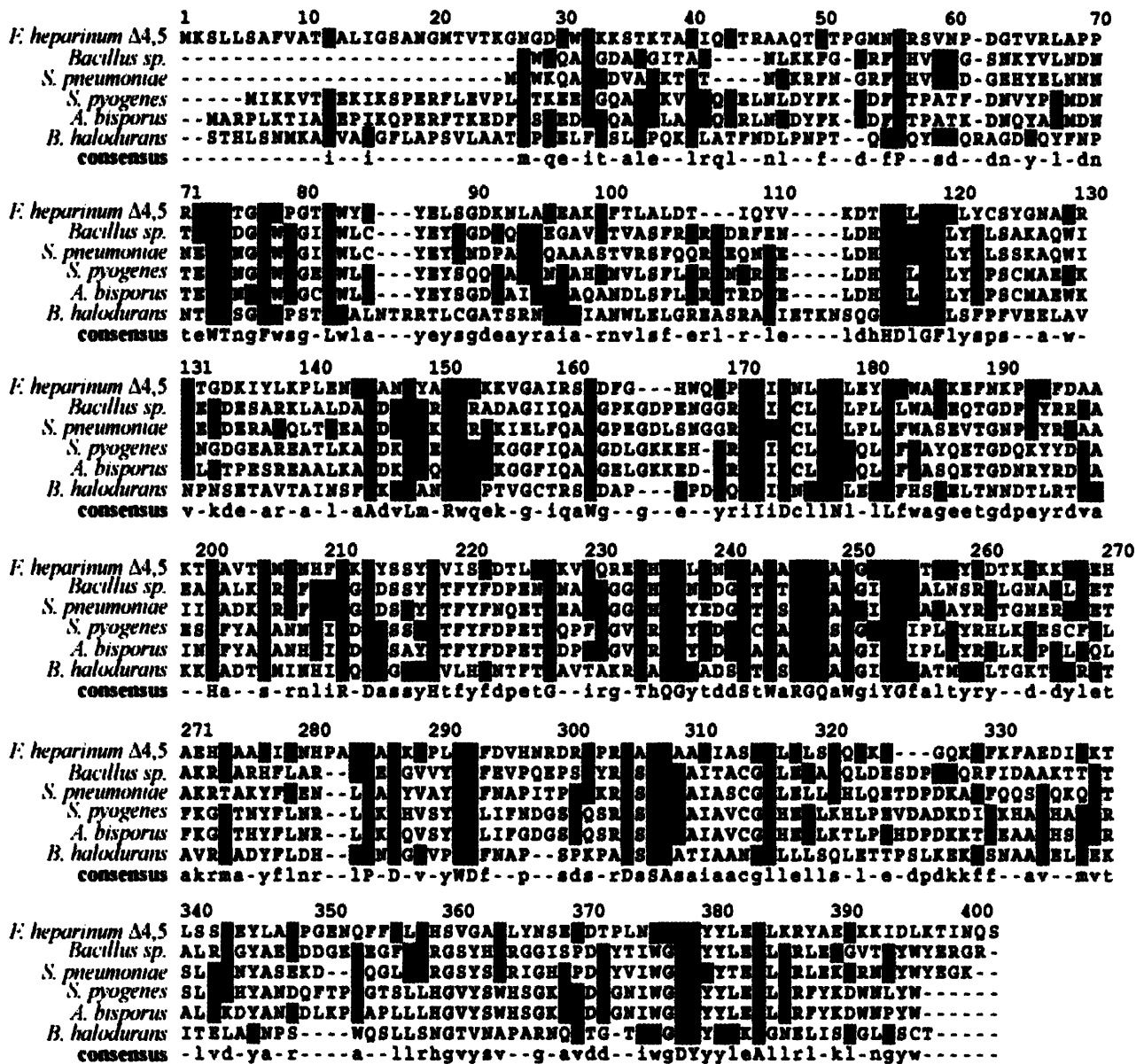
Figure 5.4. CLUSTAL W multiple sequence alignment of Δ4,5 glycuronidase and select unsaturated glucuronyl hydrolases. Protein sequences were selected from an initial BLASTP search of the protein database. Identical amino acids are shaded in dark gray, near invariant positions in charcoal, and conservative substitutions in light gray. Gen bank accession numbers are as follows: *Bacillus sp.* (AB019619); *Streptococcus pneumoniae* (AE008410); *Streptococcus pyogenes* (AE006517); *Agaricus bisporsus* (AJ271692); *Bactobacillus halodurans* (AP001514).
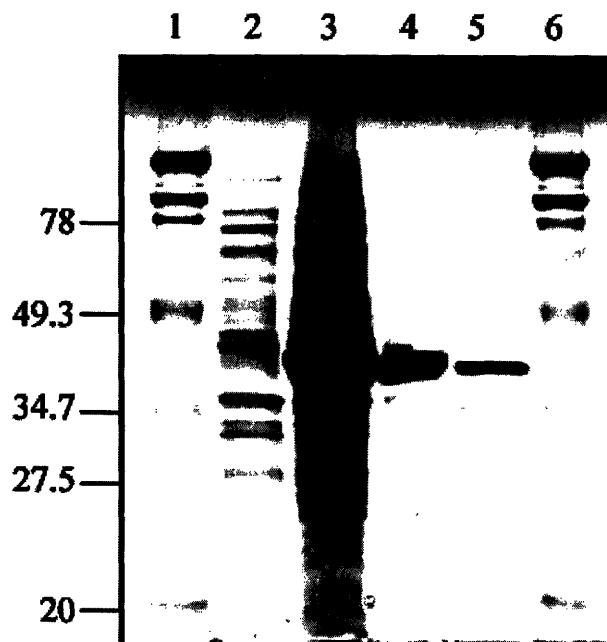
Figure 5.5. Recombinant $\Delta 4,5^{\Delta 20}$ protein expression and purification. SDS-PAGE of $\Delta 4,5$ protein fractions at various purification stages following expression in BL21 (DE3) as a 6XHIS N-terminal fusion protein. Shown here is a 12% gel stained with Coomassie-Brilliant blue. Lane 2, lysate from uninduced bacterial cells; Lane 3, crude cell lysate from induced cultures; Lane 4, Ni^{+2} chelation chromatography purification; Lane 5, thrombin cleavage to remove N-terminal 6X His purification tag. Molecular weight markers (Lanes 1 and 6) are also noted.
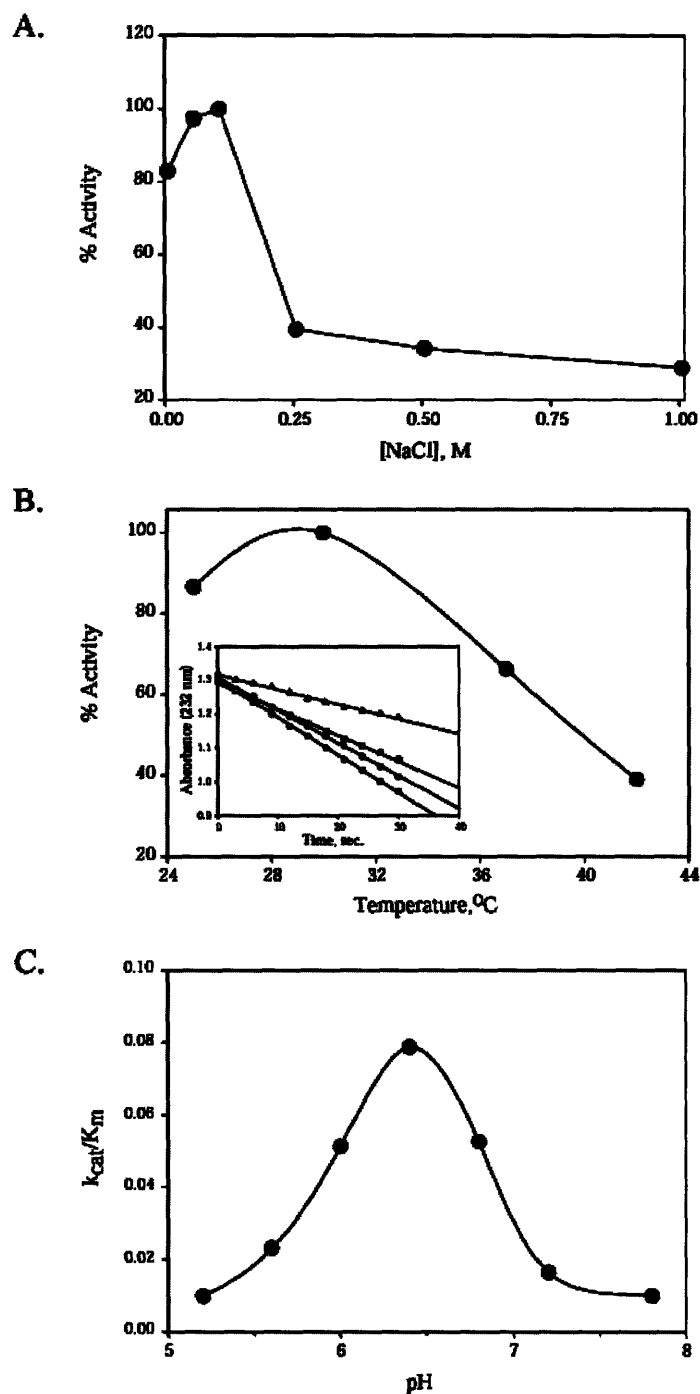
Figure 5.6. Δ4,5 glycuronidase biochemical reaction conditions. A. NaCl titration. B. Effect of reaction temperature. Inset (B): Δ4,5 glycuronidase reaction rates measured at varying temperatures by the disappearance of UV absorbance (232nm). 25°C, (O); 30°C, (●); 37°C, (□); 42°C, (▲). Relative enzyme activities in A and B were derived from initial rates and were normalized to 100% activity represented by 100 mM NaCl and 30°C, respectively. C. pH profile. K_m and k_{cat} values for the pH profile were extrapolated from Michaelis-Menten kinetics as described in the Methods. The disulfated heparin disaccharide $\Delta UH_{NS,6S}$ was used in all three experiments.
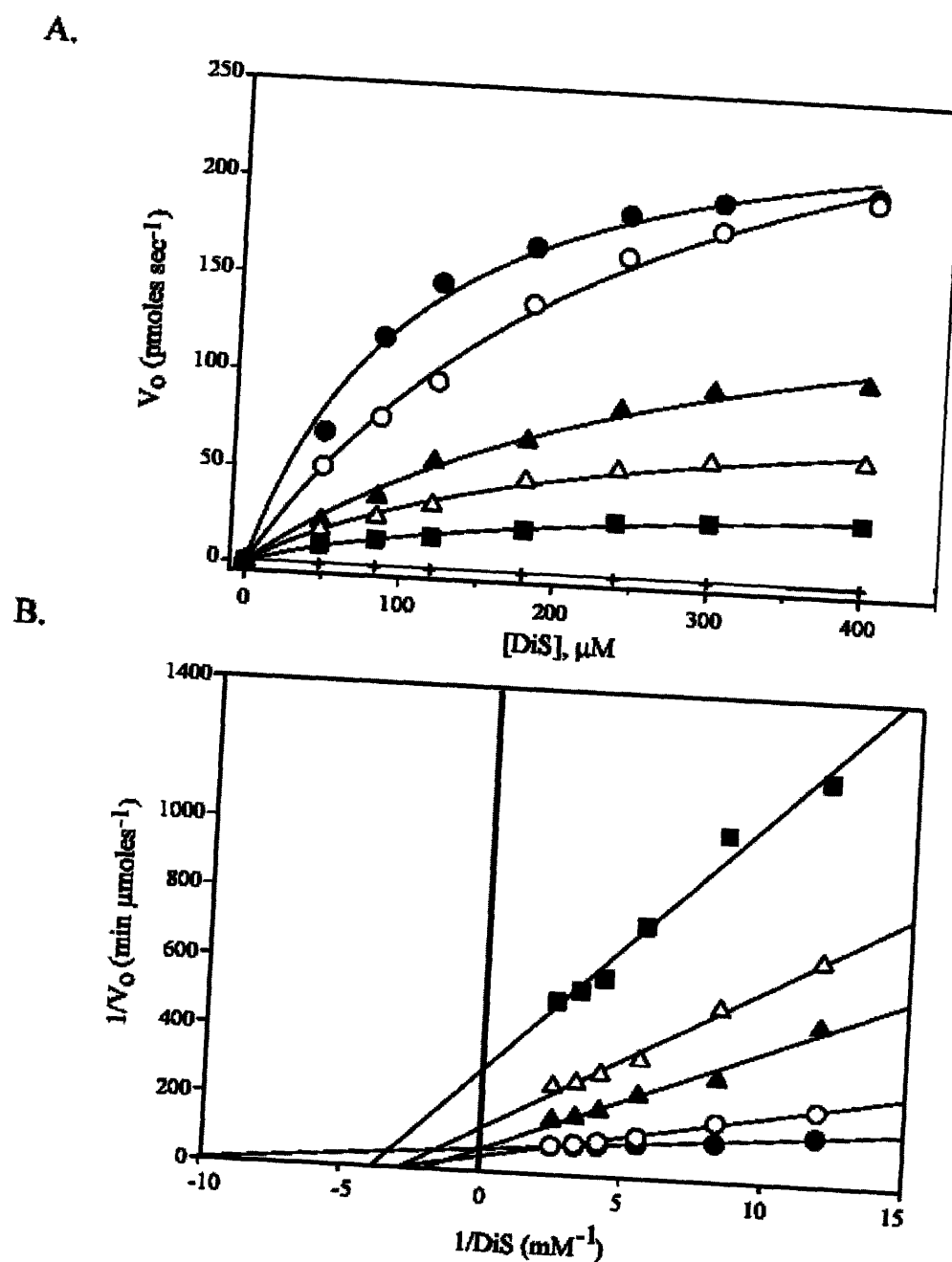
154

A.

B.

Figure 5.7. Disaccharide substrate specificity. **A.** Kinetic profiles for heparin disaccharides of varying sulfation. Initial rates were determined using 200 nM enzyme under standard conditions. V_o vs. [S] curves were fit to Michaelis-Menten steady state kinetics using a non-linear least squares analysis. **B.** Lineweaver-Burke representation of the data shown in **A.** $\Delta UH_{NAc,6S}$ (●); ΔUH_{Nac} (○); $\Delta UH_{NS,6S}$ (▲); ΔH_{NS} (△); ΔUH_{NH2} (■); $\Delta U_{2S}H_{NS,6S}$ (+).
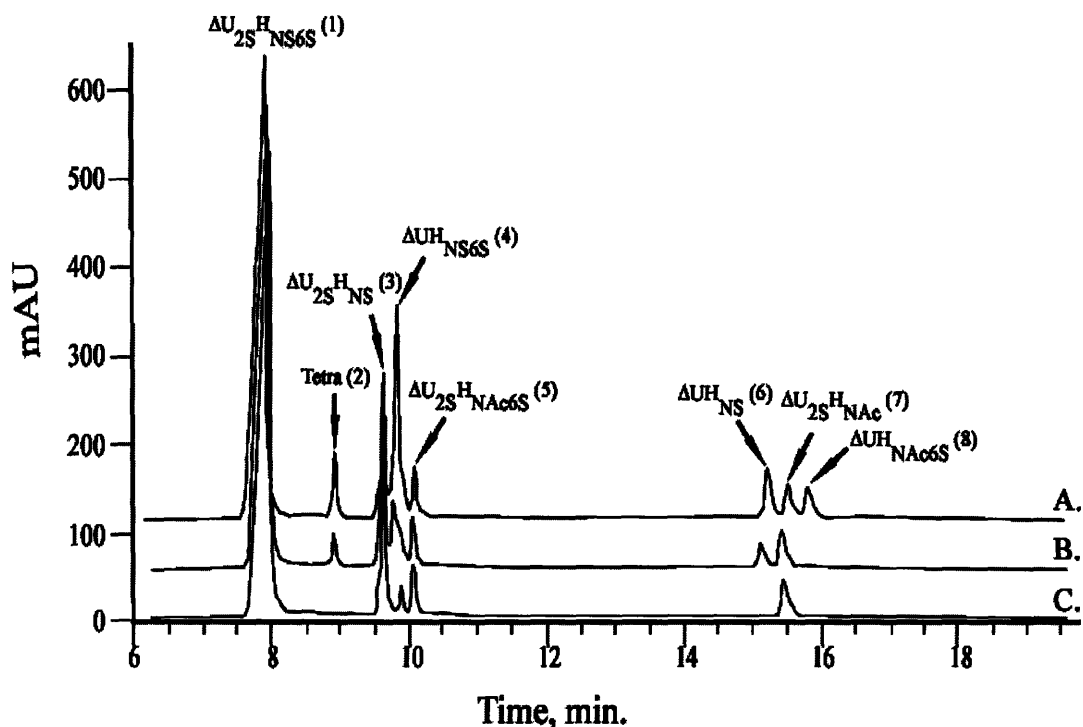
155

Figure 5.8. Tandem use of heparinases and Δ4,5 glycuronidase in HSGAG compositional analyses. 200 μg heparin was exhaustively digested with heparinases I, II, and III, after which Δ4,5 was added for a varying length of time. Unsaturated disaccharide products were resolved by capillary electrophoresis. Assignment of saccharide composition shown for each peak was confirmed by MALDI-MS. Each electrophoretogram represents varying time of Δ4,5 glycuronidase treatment: A. minus Δ4,5 enzyme control; B. 1 minute (partial) incubation; C. 30 minute (exhaustive) incubation.

Purification Step	Protein Yield (mg)	Specific Activity (μmoles DiS/ min/ mg protein)	Fold purification
Crude lysate	400	4.7	--
Ni^{+2} chromatography	205	12.9	2.7
Thrombin cleavage	205	13.6	2.9

Table 5.1. Purification summary for the recombinant $\Delta 4,5$ glycuronidase. Specific activities for each fraction were measured using 800 ng of protein and 120 μM of the unsulfated heparin disaccharide (DiS) ΔUH_{NAc} in a 100 μl reaction volume. The fold purification was calculated relative to the specific activity measured for the crude lysate.

Disaccharide substrates	k_{cat} (sec^{-1})	K_m (μM)	k_{cat}/K_m	Relative k_{cat}/K_m
ΔUH_{NAc}	15.3 $\pm$ 0.9	283 $\pm$ 31	0.054	0.49
$\Delta UH_{NAc, 6S}$	11.7 $\pm$ 0.6	107 $\pm$ 15	0.110	1.0
ΔUH_{NS}	4.9 $\pm$ 0.4	251 $\pm$ 40	0.020	0.18
$\Delta UH_{NS, 6S}$	8.8 $\pm$ 0.9	334 $\pm$ 57	0.026	0.24
$\Delta UH_{NH2, 6S}$	2.4 $\pm$ 0.2	235 $\pm$ 40	0.010	0.09
$\Delta U_{2S}H_{NS}$	N.A.	N.A	N.A.	N.A.

Table 5.2. Kinetic parameters for heparin disaccharides. k_{cat} and K_m values were determined from non-linear regression analyses of the Michaelis-Menten curves presented in Figure 7. *N.A., no activity was detected for$\Delta U_{2S}H_{NS}$.

CHAPTER 6

Heparan Sulfate Glycosaminoglycan Modulation of

Embryonic Stem Cell to Endothelial Cell Differentiation

Synopsis: The directed differentiation of embryonic stem (ES) cells holds an immense potential for regenerative medicine, necessitating an understanding of the mechanisms regulating self-renewal and cell fate decisions. Although attempts have been made to elucidate the key regulatory components of differentiation using transcriptiomic and proteomic approaches, these fail to capture the complete complexity of this process, which requires integration of the role of glycomics on differentiation. Heparan sulfate glycosaminoglycans (HSGAGs) are components of the extracellular matrix and constitute one of the major components of a cell's glycome. In the present study, we merged an enzymatic decomposition-based analysis of HSGAGs with a transcriptomal dissection of HSGAG biosynthesis to elucidate the role of the glycome in the differentiation of ES cells into endothelial cells. Murine ES cells, directed to differentiate under LIF-free conditions, progressively lost the stem cell marker, Oct-4, and acquired endothelial cells markers, such as von Willebrand Factor, VE-Cadherin, VEGF-R2 and eNOS, as detected by flow cytometry, confocal microscopy and real time-PCR. Compositional analysis of HSGAG structure by capillary electrophoresis revealed an increase in the quantity of HSGAGs with progressive differentiation, which was paralleled by an increase in the transcript levels of key HSGAG-biosynthetic enzymes. Ablation of the HSGAG-biosynthetic machinery through sodium chlorate-treatment, or the enzymatic decomposition of HSGAGs via treatment with heparinases, inhibited the formation of endothelial cells, although differentiation to other cell types did proceed as evidenced from the progressive loss of the Oct-4 signal. Reconstitution of the HSGAG moiety in sodium chlorate-treated cells by the exogenous addition of heparin partially recovered the formation of endothelial cells, suggesting that HSGAGs play a key role in the differentiation of ES cells into endothelial cells. Western blot analysis of the phospho-ERK levels suggest that HSGAGs may impinge on differentiation of ES cells into endothelial cells possibly through the MAPK pathway. The current study, for the first time, implicates the role of the glycome in the directed differentiation of embryonic stem cells, which should now be integrated into the systems biology approach for a molecular understanding of stem cell differentiation. Furthermore, the potential to regulate stem cell differentiation by modifying HSGAGs opens up new treatment modalities for regenerative medicine.

CHAPTER 6

Heparan Sulfate Glycosaminoglycan Modulation of

Embryonic Stem Cell to Endothelial Cell Differentiation

6.1 Introduction

Embryonic stem (ES) cells are pluripotent cells derived from the inner cell mass of developing blastocysts and have the unique ability to differentiate into any adult cell type (*1*). The ease of maintenance of ES cells *in vitro* in an undifferentiated pluripotent state, combined with their potential to differentiate into any cell type offers a unique opportunity to understand the biological processes that govern differentiation processes, capturing a perspective of the earliest stages of embryo development. Harnessing this potential of ES cells holds the key to regenerative medicine, contingent upon the development of methods to control their differentiation and expansion *in vitro*.

One of the attractive targets for regenerative medicine is the establishment of a viable vasculature. The vascular system is laid down during early developmental stages, and is fairly quiescent in the adult (*2, 3*). Neovascularization (the formation of new blood vessels) can also occur in the adult, albeit in a strictly regulated manner in certain physiological conditions, such as the reproductive cycle, tissue regeneration and wound healing (*4*). However, excessive and unregulated blood vessel formation has been implicated in multiple pathological conditions such as tumor angiogenesis and arthritis, making inhibition of neovascularization an attractive and tractable therapeutic goal. In contrast, some pathological conditions, such as diabetes, hypercholesterolemia and advanced age, are associated with impaired neovascularization which results in ischemic tissue (*5-7*). This impaired

neovascularization is in large part due to endothelial cell dysfunction and the promotion of neovascularization in such conditions has significant therapeutic implications in treating patients. One way to achieve this is via engineering of the impaired vasculature component *in vivo*, or the *ex-vivo* regeneration and transplantation of reconstituted vascular tissue.

The essential component of the vasculature that plays a key role in neovascularization is the endothelial cells lining the vessels. Numerous studies have attempted to regenerate the vascular tissue by using primary endothelial cells and endothelial progenitor cells (*8, 9*). However, these approaches are limited by the finite life span of primary cells, and by the low abundance of progenitor cells in circulating blood. These limitations necessitate the development of newer approaches for engineering of the vasculature.

Recent studies have revealed the potential for embryonic stem cells to differentiate into endothelial cells, making embryonic stem cells an attractive resource for regenerative therapies (*10*). However, the key limitation in using ES cells in regenerative therapies comes from the fact that upon differentiation, ES cells form a heterogeneous cell mixture as a result of their pluripotency. For example, murine ES cells were shown to give rise to multiple cell types including endothelial, hematopoietic, smooth muscle and neuronal cells *in vitro* (*11*). Furthermore, when injected in mice, ES cells have been shown to form teratocarcinomas and malignant transformations (*12, 13*). Thus, dissection of factors that regulate the differentiation process of ES cells is of extreme importance for developing regenerative therapies, so that ES cells can be directed to differentiate towards a homogeneous cell population. Among other factors, the extracellular matrix has been shown to be a critical modulator of ES differentiation (*11*).

One of the key components of the extracellular matrix (ECM) is a group of complex sugars called heparan sulfate glycosaminoglycans (HSGAGs). HSGAGs are resident components of the ECM and form a major part of a cell's glycome. HSGAGs interact with numerous proteins and play a dynamic role in various cellular events such as proliferation, morphogenesis, adhesion, migration and cell death, tumor metastasis and neovascularization (for an extensive review, see reference (*14*)). Interestingly, although there are many reports on the proteomal and transcriptomal analysis of stem cell differentiation (*15*), no previous studies has been performed on elucidating the role of the glycome component in stem cell differentiation.

In the current study, we report for the first time that the differentiation of stem cells into endothelial cells is impinged on by the glycome, specifically the HSGAG component of the cells. We demonstrate that modification of the HSGAG component of the ES cells, using enzymatic or pharmacological approaches, inhibits the differentiation into endothelial cells, albeit differentiation still proceeds as evident from the progressive loss of stem cell specific marker. Furthermore, we have shown that HSGAG modulation impinges on the mitogen activated protein kinase (MAPK) pathway, indicating that HSGAG modulation of ES cell differentiation could possibly be mediated through the MAPK pathway.

6.2 Experimental Procedures

Cell Culture

Mouse embryonic stem cells (J1) were grown in DMEM with 2mM L-glutamine

(GIBCO) adjusted to contain 100mM sodium pyruvate, 10mM nonessential amino acids and

1.5g/L sodium bicarbonate, with 15% fetal bovine serum (Hyclone) and supplemented with

100 units/mL penicillin G (Sigma), 100 µg/mL streptomycin sulfate (Sigma), 30 mM beta-

mercaptoethanol (Sigma) and 1000 units/ml murine Leukemia Inhibitory factor, LIF

(Chemicon). The cells were plated in 10cm cell culture dishes, coated with 0.1 % gelatin B

(Sigma) and grown at 37°C in a 5% CO_2 humidifier incubator. The culture medium was

changed every 2 days. The cells were subcultured at a split ratio of 1:10 when they reached

80% to 90% confluency, using a 0.25% trypsin-EDTA solution (Sigma). To induce

formation of embryoid bodies and differentiation into endothelial cells, J1 cells were plated

at a density of 1.25×10^5 cells/ 100 mm dish, or 3×10^4 cells/ 6 wells, and cultured in the

absence of LIF for defined time periods. Fresh media was replenished every 2 days.

Enzymatic and Chemical Treatment of Cells

Heparinase-III (Hep-III) was prepared as described previously (*16*). Heparinase-I

(Hep-I) was a generous gift of Momenta Pharmaceuticals. Sodium chlorate and heparin were

purchased from Sigma and Cambrex BioWhittaker respectively. To enzymatically modify

the HSGAG -glycome signature of the cells, cells were washed with phosphate buffer saline

and incubated with Hep-I or Hep-III for 30 minutes everyday at 37 °C in serum free DMEM.

Digested HSGAG residues were removed by washing the cells, and cells were then fed with

fresh media. In an additional experiment, cells were cultured with medium containing sodium chlorate which blocks the sulfation of HSGAGs. In a reversal experiment, heparin was added to the culture to overcome the chlorate-induced synthetic block. Treated cells were analyzed for differentiation into endothelial cells between day 3- day 15. Hep-III was added to the cells at a concentration of 2.5µg/mL, Hep-I was added at a concentration of 1.5µg/mL, sodium chlorate at 10mM, and heparin (20µg). The concentrations of enzymes were optimized to exhaustively cleave all cell surface HSGAGs in the given period of time. All enzymes and chemicals were diluted in serum free DMEM.

Flow Cytometry

Cells were harvested at defined time points during differentiation, incubated with rat monoclonal antibody against CD16/CD32 (1:50 dilution, Pharmingen) to block Fcϒ receptors, and incubated further with the primary antibody against Von Willebrand Factor (vWF) (Dako, rabbit polyclonal, added at a 1:100 dilution). An isotype-matched polyclonal antibody IgG (Pharmingen) was used as the control. Cells were then washed twice and were incubated with FITC-conjugated secondary antibody (Jackson ImmunoResearch) in a 1:100 dilution, washed twice and taken up in OptiMEM (GIBCO/BRL) and analysed on a Becton Dickinson FACScan flow cytometer (excitation 488 nm, argon laser; emission 580/30).

Confocal Microscopy

For microscopic analysis of endothelial cells, the differentiating J1 cells were fixed in cold methanol at designated time points, blocked in goat serum, and probed overnight with a rabbit primary antibody against vWF, an endothelial cell marker, or Oct-4, a stem cell

marker. The sections were washed and re-probed with a goat secondary antibody coupled to FITC. The nuclei were counterstained with propidium iodide. Images were captured using a Leica LSM510 confocal microscope at a 512×512 pixels resolution. Fluorochromes were excited with 488 nm and 543 nm laser lines, and the images were captured using 505-530 BP and 565-615 BP filters at a 512×512 pixel resolution.

Real time PCR

RNA was isolated from J1 cells using TRIzol (Invitrogen) and RNAlater (Qiagen) according to the manufacturer's protocol. Single stranded cDNA was generated via oligo-dT primed reverse transcription, and genes serving as markers for differentiation were quantified using an Abi Prism 7700 real-time PCR thermocycler (Applied Biosystems). Expression levels were obtained for the marker genes Oct-4, VE-cadherin, VEGF-R2, and eNOS, with β -actin used as a control. PCR conditions involved denaturation for 10 min at 95°C, followed by 40 cycles of denaturation for 20 sec at 94°C, and annealing and extension for 1 min at 60°C. cDNA isolated from J1 cells and primers for the described genes were mixed with SYBR Green PCR Mastermix (Applied Biosystems) for real-time quantification according to the manufacturers instructions. Primers used were 5'-ccaatcagcttgggctagag-3' and 5'-ctgggaaaggtgtccctgta-3' for Oct-4; 5'-accgagagaaacaggctgaa-3'and 5'-agacggggaagttgtcattg-3'for VE-Cadherin; 5'-ggacagtgctccaaccaaat-3' and 5'-gttcacactgcagacccaga-3' for TIE-2; 5'-gctttcggtagtgggatgaa-3' and 5'-ggccttccatttctgtacca-3' for VEGF-R2; 5'-tcttcgttcagccatcacag-3' and 5'-cctatagcccgcatagcgta-3' for ENOS; 5'-agccatgtacgtagccatcc-'3 and 5'-ctctcagctgtggtggtgaa-'3 for β-Actin. The sulfotransferase primers and conditions were chosen as described elsewhere (Kiziltepe *et al.,* manuscript in preparation). Relative and

normalized levels of gene expression were obtained according to the equation: $2^{-(Ct\,(gene)\,-\,Ct\,(\beta\text{-}Actin))}$ where Ct is the cycle number at which amplification of each gene crossed an arbitrary threshold within the exponential phase of amplification. Gene expression levels were further normalized to expression levels prior to differentiation.

Isolation and Compositional Analysis of HSGAGs

Cell surface HSGAG fragments were isolated from J1 cells at various stages of differentiation. Briefly, cells were washed with PBS, and treated with trypsin/EDTA (GibcoBRL) at 37 °C for 25 min to harvest cell surface proteoglycans. The resulting cell/trypsin solution was boiled for 10 min to deactivate the trypsin and other proteins. The solution was centrifuged at 4500g, the supernatant was collected and concentrated by centrifuging in a Centriprep-3 (Amicon). The concentrated supernatant was run through ultrafree-DEAE (Pharmacia) that had been equilibrated with 0.1 M sodium phosphate buffer, pH 6.0, that contained 0.15 M NaCl. The bound HSGAG fragments were washed and eluted with 0.1 M sodium phosphate buffer, pH 6.0, that contained 1.0 M NaCl. The fragments were then concentrated and buffer-exchanged into ultra-pure water by application to a Microcon filter (molecular weight cutoff = 3,000 Da). The samples were exhaustively digested overnight with a mixture of Hep-I-III (1 milliunit each) in 25 mM sodium acetate and 1 mM calcium acetate, pH 7.0. The samples were analyzed by capillary electrophoresis using a high-sensitivity flow cell under reverse polarity with a running buffer of 50 mM Tris/phosphate, pH 2.5. Identities of the resultant saccharides were determined based on co-migration with known standards.

Western Blotting

For elucidating the biochemical pathways involved downstream of the modulation of the cell surface sugars, protein contents of the embryoid bodies were solubilized by rapid mixing with 3× SDS sample buffer under reducing conditions. Equivalent amounts of protein per sample were electrophoretically resolved on 4-12% gradient polyacrylamide gels, and transferred onto a nitrocellulose (0.22 μm) membrane. The membrane was subsequently probed with a phosphor-ERK antibody (1:800 dilution, Cell Signaling Technologies), which specifically detects the phosphorylated forms of ERK1 and 2. The signal was amplified using a 1:2,000 dilution of the appropriate horseradish peroxidase-conjugated secondary antibody (BioRad), and the immunocomplexes were visualized using enhanced chemiluminescence detection (Amersham Life Science). The signal was normalized to the expression of total ERK1/2, which was detected on the same blot using ERK1/2 specific antibodies (Santa Cruz, used at a 1:200 dilution). Quantification of the luminescence signal was carried out by a Kodak 2000R imager.

Statistical analysis

Statistical significance was tested using the Students t-test or one-way ANOVA followed by Dunnets or Friedman's Post-Hoc test. (Graphpad Prism 3 software). $P<0.05$ was considered to be significant.

6.3 Results

ES cells differentiate significantly towards an endothelial cell population.

To optimize the conditions for efficient differentiation of ES cells into endothelial

cells *in vitro*, we analyzed the expression levels of ES cell specific and endothelial cell

specific markers at different stages of differentiation under different cell culture conditions.

Specifically, we tested the effects of cell density, the ECM content (*i.e.*, gelatin, collagen,

laminin, matrigel), exogenous addition of growth factors (*i.e.*, VEGF, FGF, HGF and insulin)

and different types of serum (*i.e.*, Sigma FBS and Hyclone FBS). The optimal conditions that

gave a maximum of 30% differentiation into endothelial cells were achieved by plating the

undifferentiated cells onto gelatin B coated tissue culture dishes, at a concentration of

1.25×10^5 cells/ 100 mm dish, in the presence of 15% Hyclone FBS, 30 mM beta-

mercaptoethanol, 1 mM sodium pyruvate, and in the absence of LIF without further

passaging for 7-15 days. No effects of exogenous addition of growth factors were detected,

possibly due to the presence of sufficient levels of growth factors either coming from the

FBS or produced by the autocrine signaling of ES cells. The ES cell specific marker analyzed

was Octamer-4 (Oct-4) (*17*) and endothelial cell specific markers analyzed were von

Willebrand Factor (vWF) (*18*), vascular endothelial growth factor receptor-2 (VEGF-R2)

(*19*), vascular endothelial cadherin (VE-cad) (*20*), and endothelial cell specific nitric oxide

synthase (eNOS) (*21*). The flow-cytometry analysis of the expression of vWF in

differentiating ES cells showed an increase in the proportion of cells expressing detectable

levels of vWF over time (Figure 6.1A). Similarly, confocal microscopy analysis of

differentiating ES cells showed that expression levels of vWF protein increased during

differentiation reaching a maximum between days 7-15 (Figure 6.1B). Consistently, the expression levels of the ES cell marker, Oct-4, decreased progressively with time (Figure 6.1B). This temporally converse expression of vWF and Oct-4 indicated an efficient differentiation of ES cells towards an endothelial cell population. Since vWF is also expressed in megakaryocytes, although to a lesser extent (18), we confirmed our results by analyzing the temporal transcriptomal expression of other endothelial cell-specific markers, VEGF-R2, VE-cadherin and eNOS during differentiation. The real-time PCR results showed a decrease in the transcript levels of Oct-4, and an increase in the transcript signals of VEGF-R2, VE-cadherin, and eNOS, 7 days after induction of differentiation (Figure 6.1C). These results are consistent with the results from the confocal microscopy and flow-cytometry experiments and suggest significant differentiation towards an endothelial cell population.

HSGAG synthesis is upregulated during differentiation of ES cells.

To investigate the role of HSGAGs in the differentiation of ES cells, first we analyzed the compositional changes in HSGAGs during differentiation. For this purpose, cell surface HSGAGs were harvested at different stages of differentiation, normalized to cell number and subjected to compositional analysis of the comprising disaccharide units by capillary electrophoresis. Although there was very low levels of detectable HSGAGs on undifferentiated cells (at day 3), there was a progressive and dramatic increase in the total levels of sulfated HSGAGs with differentiation (Figure 6.2A).

To further confirm these results, we also investigated changes in the genetic expression of some critical HSGAG biosynthetic enzymes during differentiation. Biosynthesis of HSGAGs in mammals is initiated by the formation of a tetrasaccharide

linkage (Glucuronic Acid-Galactose-Galactose-Xylose) to a proteoglycan core (22). After the initial formation of this linkage tetrasaccharide, the alternating addition of glucuronic acid and N-acetyl-glucosamine from their UDP-sugar nucleotide precursors forms a repeating 1,4-linked disaccharide chain. The disaccharide chain is further modified by a series of sulfotransferases, of which N-deacetylase-N-sulfotransferase and the 2-O, 3-O, and 6-O heparan sulfate sulfotransferases play a key role. Tissue and substrate specific isoforms of many of these sulfotransferases have been discovered, indicating a further level of complexity in the biosynthesis of HSGAGs (for reviews see references (23, 24)). It is the sulfotransferases that give HSGAGs their "signature" structure and thus these enzymes are critical in modulating specific structure-function relationships of HSGAGs (14). We performed real time PCR analysis for 2-O, 3-O, and 6-O sulfotransferases, N-deacetylase-N-sulfotransferase, and their isoforms. This analysis demonstrated that there were significantly higher levels of all sulfotransferase transcipts as differentiation progressed with time, supporting the results from the compositional analysis of HSGAGs (Figure 6.2B). These results suggest a possible role for HSGAGs in differentiation of ES cells.

HSGAGs modulate differentiation of ES cells into endothelial cells.

To dissect the role of HSGAG structure in ES cell differentiation, we modified the cell surface and extracellular HSGAG moiety of differentiating ES cells using an enzymatic and pharmacological approach. Specifically, differentiating ES cells were incubated with Hep -I or Hep-III for 30 minutes every day to cleave the cell surface and ECM HSGAGs at structurally distinct sites, or cultured in media supplemented with sodium chlorate, which is a pharmacological inhibitor of the HSGAG biosynthesis (25). Hep-I cleaves the HSGAGs at

the highly sulfated regions, and Hep-III cleaves the HSGAGs at the undersulfated regions

(*26, 27*). At the end of enzymatic degradation, the medium containing the enzymes and

HSGAG fragments was replaced with fresh medium. At different time points, the

differentiation of ES cells into endothelial cells was monitored using flow cytometry,

confocal microscopy and real-time PCR. Flow cytometry analysis revealed that all the

treatments inhibited the expression of vWF factor, although to different extents (Figure

6.3A). Hep-I, which cleaves the HSGAGs at the highly sulfated regions, and sodium

chlorate, which inhibits HSGAG biosynthesis, both decreased the proportion of cells

expressing detectable levels of vWF by approximately 5 fold. Hep-III, which cleaves the

HSGAGs at the undersulfated regions, although milder, also showed an inhibitory effect by

~3 fold (Figure 6.3B).

The inhibitory effects of Hep-I, Hep-III and sodium chlorate treatments on vWF

expression were also confirmed by confocal microscopy studies (Figure 6.4A). Interestingly,

none of the treatments had significant effects on Oct-4 levels of differentiating cells (Figure

6.4B). These findings suggest that treatment with Hep-I, Hep-III or sodium chlorate does not

inhibit differentiation of ES cells, but specifically inhibit their differentiation into endothelial

cells. Most importantly, addition of heparin to sodium chlorate treated ES cells during

differentiation reconstituted conditions that favor differentiation towards endothelial cells as

detected by increased vWF staining (Figure 6.4C).

Real time PCR experiments further confirmed these findings. While there was no

effect on Oct-4 transcript levels with any of the treatments, Hep-I and sodium chlorate

treatments significantly decreased the transcript levels of VEGF-R2, VE-Cadherin and eNOS

(Figure 6.5). Hep-III treatment also decreased the transcript levels of VE-Cadherin and

171

eNOS, however, surprisingly increased the transcript levels of VEGF-R2 (Figure 6.5).

Orthogonal effects of Hep-I and Hep-III in cell phenotype have also been reported elsewhere

(28) and suggests that HSGAGs modulate differentiation of ES cells into endothelial cells in

a structurally specific manner. Consistent with the confocal microscopy results, real time

PCR results also showed that addition of heparin to sodium chlorate treated ES cells

reconstituted conditions that favor differentiation towards endothelial cells as detected by

increased VEGF-R2, VE-cadherin and eNOS transcript levels.

HSGAGs may impinge on ES cell differentiation via MAPK pathway.

The MAPK pathway is the downstream convergence point of the signaling of several

angiogenic factors such as VEGF, FGF, HGF, EGF, PDGF and angiopoietins (29). Given the

significant role of these factors in endothelial cell proliferation (for review see reference

(30)), we set out to reveal if HSGAG modulation was impinging on the MAPK pathway of

differentiating ES cells. As shown in Figure 6.6A, 6.6B, treatment of differentiating ES cells

with the HSGAG-modifying enzymes- Hep-I or Hep-III, or the pharmacological inhibitor-

sodium chlorate, inhibits the phosphorylation of ERK1/2. Interestingly this inhibition is

reversed by the addition of exogenous heparin. These results suggest that MAPK pathway is

involved in differentiation of ES cells, and HSGAGs are critical modulators of this pathway.

6.4 Discussion

Neovascularization is involved in many physiological processes such as development, reproduction, tissue regeneration and wound healing, and is highly regulated through an 'angiogenic switch' (3, 4). However, aberrant neovascularization underlies many pathophysiological conditions. For example, in diabetes, hypercholesterolemia and advanced age, dysfunctional endothelial cells and impaired neovascularization can result in ischemic tissue (5-7). This can lead to chronic wounds and cause the loss of extremities in over 3% of all diabetics, or result in impaired cardiovascular function in coronary artery disease[1]. For such cases, endothelial cell regeneration and transplantation has potential therapeutic implications in treating patients.

In a recent study, Levenberg et al demonstrated that embryonic stem cells exhibited the potential to differentiate into endothelial cells, and form vessel-like structures *in vitro* and *in vivo* (31). In the current study, we used endothelial-specific markers such as vWF, VE-cadherin, eNOS, and VEGF-R2 to follow the differentiation of ES cells into endothelial cells over a period of 7-15 days. Both confocal microscopy and flow cytometry analysis suggested that the stems cells progressively differentiate into endothelial cells, which was supported by the real-time PCR data. Consistent with earlier reports, we could observe the formation of cord-like primordial vascular structures by confocal microscopy by day 10 (data not shown), indicating that the embryoid bodies could serve as an interesting model to study the molecular mechanisms of vasculogenesis and early angiogenesis.

Although differentiation of embryonic stem cells into endothelial cells is an established concept, limited knowledge exists on the possible factors that impinge on this differentiation process. Current studies to elucidate these factors and underlying mechanisms

---

[1] American Diabetes and Heart association

have primarily been based on transcriptomic and proteomic approaches (*15*). Numerous studies, including from our laboratory, have established that a third dimension, the glycome, which includes the cell surface and extracellular matrix HSGAGs, also impinges on the cellular phenotype. However, no previous studies were done on the role of HSGAGs in stem cell differentiation, primarily owing to the complexities involved in dissecting HSGAG structure-function relationships.

In the current study, we dissected the role of HSGAGs in ES cell differentiation into endothelial cells. Our experiments suggest that the modulation of the HSGAG moiety of differentiating ES cells, by either enzymatic degradation or by inhibition of their biosynthesis, inhibits their differentiation into endothelial cells. However, despite these modulations of the HSGAGs in the cellular microenvironment, there was still significant decrease in the levels of the ES cell-specific marker, Oct-4. This suggests that although the differentiation of ES cells specifically into endothelial cells is inhibited following HSGAG modulation, the overall differentiation process is not inhibited. Intriguingly, in the current study we detected low levels HSGAGs on the surface of undifferentiated ES cell, and there was a progressive increase in the highly-sulfated HSGAG signal with differentiation. Altogether, these findings suggest that although the HSGAG component of differentiating ES cells are not critical in determining their differentiation state, they play a key role in determining which cell lineage the differentiation process will yield. Further investigation is necessary to determine the cell lineage that is obtained.

To dissect the possible signaling pathways impinged by HSGAGs of differentiating ES cells, we studied the effect of the HSGAG modulation on MAPK pathway. MAPK is a key convergence point in the signal transduction pathways of multiple angiogenic factors,

including tyrosine kinase receptor ligands such as FGF, VEGF, HGF, EGF, PDGF and angiopoietins (*29, 32*). The potential for some of these factors to promote the differentiation of stem cells into endothelial cells is well established (*33-35*), and all these factors can use HSGAGs as secondary ligand binding sites (*36*). Our studies showed inhibition of ERK phosphorylation following treatment with HSGAG-degrading enzymes or sodium chlorate. This indicates that HSGAGs are important modulators of the upstream MAPK pathway in differentiation of ES cells into endothelial cells. It is possible that paracrine or autocrine signaling through the growth factors, which act through the MAPK pathway, could lead to endothelial cell enrichment as a result of differentiation. Intriguingly, in a recent study, the inhibition of the MAPK pathway was shown to promote the self renewal of embryonic stem cells (*37*), indicating a key role for the MAPK pathway in stem cell differentiation. Further studies are warranted to dissect out the exact players in the differentiation process, such as by using pharmacological inhibitors such as PTK787 against VEGFR or antibodies against FGFR to selectively knock out signaling pathways.

The current study is the first to clearly dissect the role of the glycome, specifically the role of HSGAGs, in modulating the differentiation of embryonic stem cells into endothelial cells potentially creating the opportunity to harness the optimal HSGAG compositions for directed differentiation, the goal for regenerative medicine.

6.5 References

1. Cowan, C. A., Klimanskaya, I., McMahon, J., Atienza, J., Witmyer, J., Zucker, J. P., Wang, S., Morton, C. C., McMahon, A. P., Powers, D., and Melton, D. A. (2004) *N Engl J Med 350*, 1353-6.
2. Folkman, J. (2003) *Curr Mol Med 3*, 643-51.
3. Folkman, J. (1997) *Exs 79*, 1-8.
4. Zetter, B. R. (1988) *Chest 93*, 159S-166S.
5. Rivard, A., Fabre, J. E., Silver, M., Chen, D., Murohara, T., Kearney, M., Magner, M., Asahara, T., and Isner, J. M. (1999) *Circulation 99*, 111-20.
6. Rivard, A., Silver, M., Chen, D., Kearney, M., Magner, M., Annex, B., Peters, K., and Isner, J. M. (1999) *Am J Pathol 154*, 355-63.
7. Van Belle, E., Rivard, A., Chen, D., Silver, M., Bunting, S., Ferrara, N., Symes, J. F., Bauters, C., and Isner, J. M. (1997) *Circulation 96*, 2667-74.
8. Joyce, N. C., and Zhu, C. C. (2004) *Cornea 23*, S8-S19.
9. Murasawa, S. (2004) *Nippon Ronen Igakkai Zasshi 41*, 48-50.
10. Gerecht-Nir, S., Ziskind, A., Cohen, S., and Itskovitz-Eldor, J. (2003) *Lab Invest 83*, 1811-20.
11. Yamashita, J., Itoh, H., Hirashima, M., Ogawa, M., Nishikawa, S., Yurugi, T., Naito, M., and Nakao, K. (2000) *Nature 408*, 92-6.
12. Erdo, F., Trapp, T., Buhrle, C., Fleischmann, B., and Hossmann, K. A. (2004) *Orv Hetil 145*, 1307-13.
13. Erdo, F., Buhrle, C., Blunk, J., Hoehn, M., Xia, Y., Fleischmann, B., Focking, M., Kustermann, E., Kolossov, E., Hescheler, J., Hossmann, K. A., and Trapp, T. (2003) *J Cereb Blood Flow Metab 23*, 780-5.
14. Sasisekharan, R., and Venkataraman, G. (2000) *Curr Opin Chem Biol 4*, 626-31.
15. Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R., Lebkowski, J., and Stanton, L. W. (2004) *Nat Biotechnol 22*, 707-16.
16. Godavarti, R., Cooney, C. L., Langer, R., and Sasisekharan, R. (1996) *Biochemistry 35*, 6846-52.
17. Pesce, M., Gross, M. K., and Scholer, H. R. (1998) *Bioessays 20*, 722-32.
18. Sadler, J. E. (1991) *J Biol Chem 266*, 22777-80.
19. Yamaguchi, T. P., Dumont, D. J., Conlon, R. A., Breitman, M. L., and Rossant, J. (1993) *Development 118*, 489-98.
20. Lampugnani, M. G., Resnati, M., Raiteri, M., Pigott, R., Pisacane, A., Houen, G., Ruco, L. P., and Dejana, E. (1992) *J Cell Biol 118*, 1511-22.
21. Alderton, W. K., Cooper, C. E., and Knowles, R. G. (2001) *Biochem J 357*, 593-615.
22. Varki A, C. R., Esko J, Freeze H, Hart G, Marth J (1999) *Essentials of Glycobiology*, Cold Spring Harbor, New York.
23. Lindahl, U., Kusche-Gullberg, M., and Kjellen, L. (1998) *J Biol Chem 273*, 24979-82.
24. Habuchi, O. (2000) *Biochim Biophys Acta 1474*, 115-27.
25. Safaiyan, F., Kolset, S. O., Prydz, K., Gottfridsson, E., Lindahl, U., and Salmivirta, M. (1999) *J Biol Chem 274*, 36267-73.

26.   Linhardt, R. J., Turnbull, J. E., Wang, H. M., Loganathan, D., and Gallagher, J. T. (1990) *Biochemistry 29*, 2611-7.

27.   Godavarti, R., and Sasisekharan, R. (1996) *Biochem Biophys Res Commun 229*, 770-7.

28.   Liu, D., Shriver, Z., Venkataraman, G., El Shabrawi, Y., and Sasisekharan, R. (2002) *Proc Natl Acad Sci U S A 99*, 568-73.

29.   Sengupta, S., Gherardi, E., Sellers, L. A., Wood, J. M., Sasisekharan, R., and Fan, T. P. (2003) *Arterioscler Thromb Vasc Biol 23*, 69-75.

30.   Le Querrec, A., Duval, D., and Tobelem, G. (1993) *Baillieres Clin Haematol 6*, 711-30.

31.   Levenberg, S., Golub, J. S., Amit, M., Itskovitz-Eldor, J., and Langer, R. (2002) *Proc Natl Acad Sci U S A 99*, 4391-6.

32.   Griffioen, A. W., and Molema, G. (2000) *Pharmacol Rev 52*, 237-68.

33.   Keller, G. M. (1995) *Curr Opin Cell Biol 7*, 862-9.

34.   Darland, D. C., and D'Amore, P. A. (2001) *Curr Top Dev Biol 52*, 107-49.

35.   Hirashima, M., Kataoka, H., Nishikawa, S., and Matsuyoshi, N. (1999) *Blood 93*, 1253-63.

36.   Keiser, N., Venkataraman, G., Shriver, Z., and Sasisekharan, R. (2001) *Nat Med 7*, 123-8.

37.   Qi, X., Li, T. G., Hao, J., Hu, J., Wang, J., Simmons, H., Miura, S., Mishina, Y., and Zhao, G. Q. (2004) *Proc Natl Acad Sci U S A 101*, 6027-32.
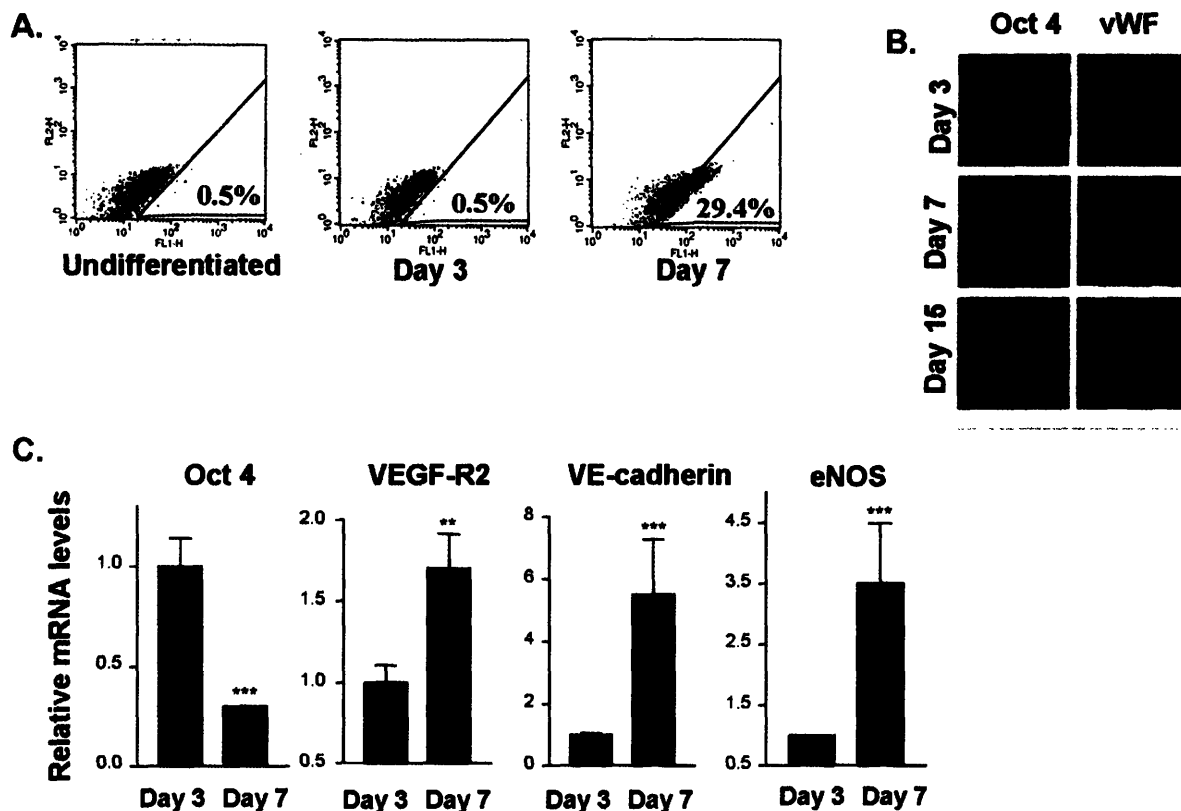
Figure 6.1. Embryonic stem cell differentiation into endothelial cells. The differentiation of ES cells into endothelial cells was detected by using cell specific markers. vWF, VEGF-R2, VE-cadherin, and eNOS were used as endothelial cell specific markers and Oct-4 was used as ES cell specific marker. (A) Flow cytometry analysis of vWF at different stages of differentiation. (B) Confocal images of vWF and Oct-4 staining in differentiating ES cells. (C) Real-time PCR data of VEGF-R2, VE-cadherin, eNOS, and Oct-4 at different stages of differentiation. The relative mRNA levels are normalized to day 3, in which no significant differentiation was observed. Altogether, these results show that Oct-4 transcription and expression progressively diminishes with differentiation, while that of vWF, VEGF-R2, VE-cadherin, and eNOS increases, suggesting efficient differentiation towards an endothelial cell population. Representative images are shown.
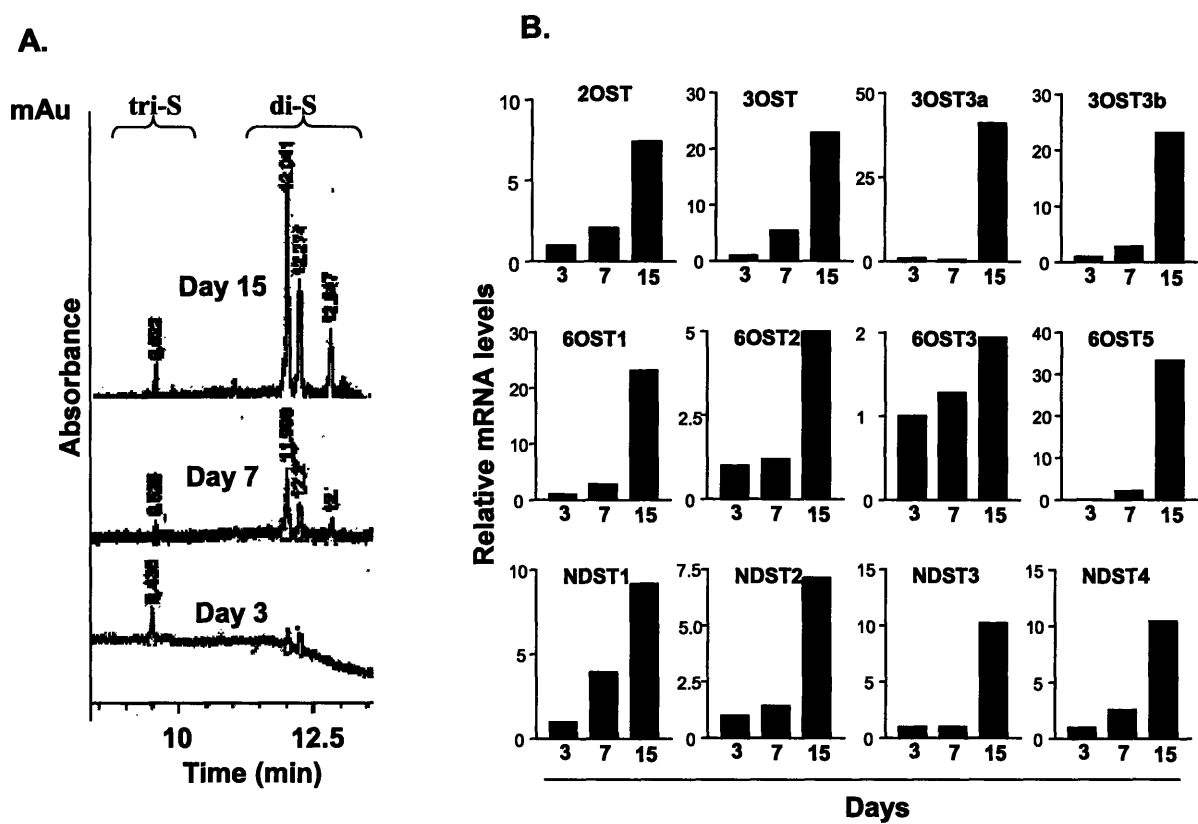
Figure 6.2. Analysis of HSGAG composition and HSGAG biosynthesis during differentiation of ES cells. (A) Capillary electrophoresis chromatograms of HSGAG components of ES cells at different stages of differentiation. The isolated HSGAGs were subjected to heparinase I-III digestion and analyzed via capillary electrophoresis. The signal for di- and tri- sulfated disaccharides increased significantly upon differentiation. (B) Real time PCR data of HSGAG biosynthetic enzymes at different stages of differentiation. The transcription of HSGAG biosynthetic enzymes was monitored using real-time PCR. Graphs show the relative transcriptional levels of 2O-sulfotransferases (2OST), 3O-sulfotransferases (3OST), 6O-sulfotransferases (6OST), N-Deacetylase-Sulfotransferases (NDST) and their isoforms at different stages of differentiation. The transcripts of these enzymes progressively increases as ES cells differentiate. Data from representative experiments are shown.
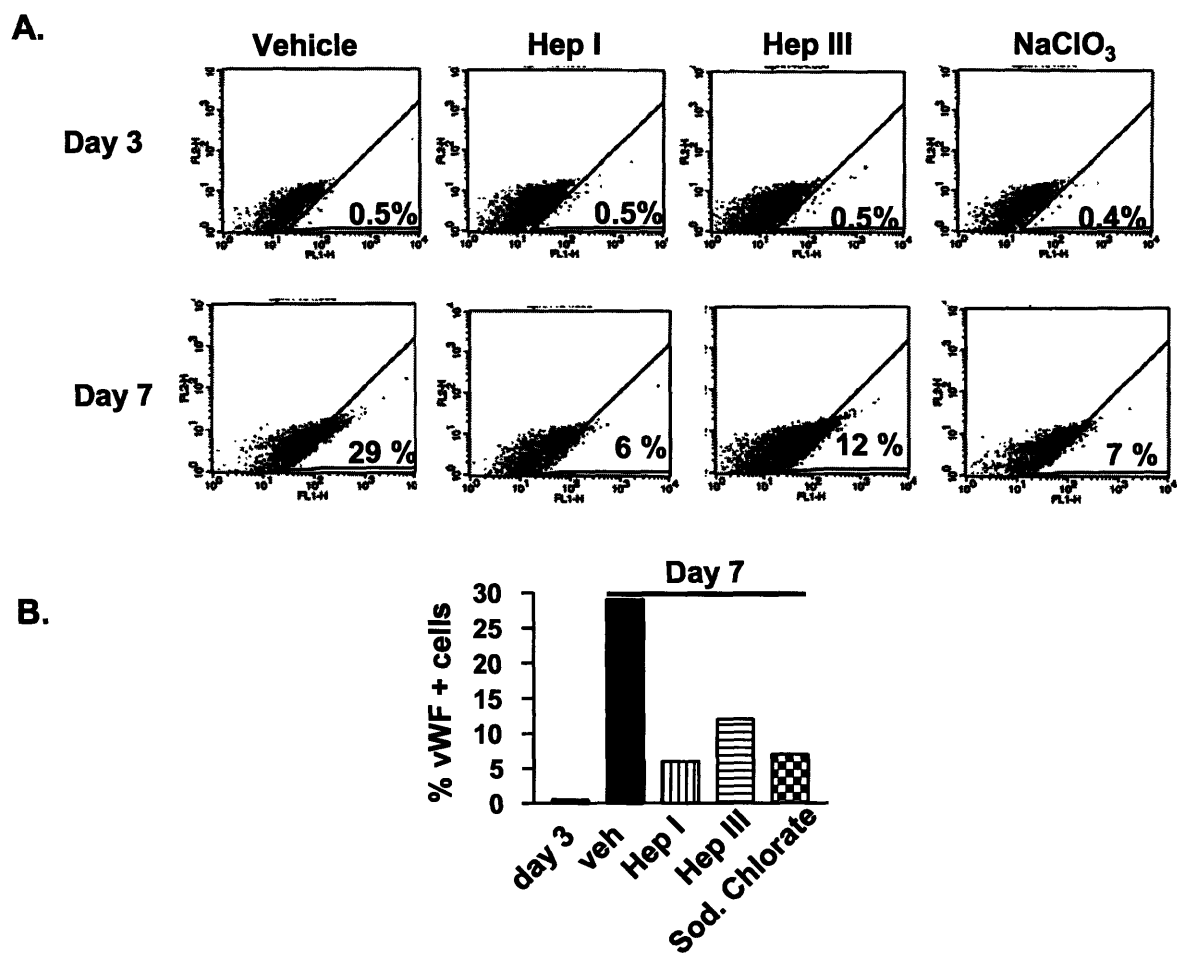
A.

B.

Figure 6.3. Flow cytometry analysis of the effects of enzymatic or pharmacological modification of the HSGAGs on differentiation of ES cells into endothelial cells. (A) Effects of treatments on vWF staining at different stages of differentiation. (B) The bar plot of % cells staining positively for vWF in the flow cytometry experiment. Extracellular degradation of HSGAGs by heparinase I and heparinase III treatment, as well as inhibition of HSGAG biosynthesis by sodium chlorate, inhibits differentiation of ES cells into endothelial cells.
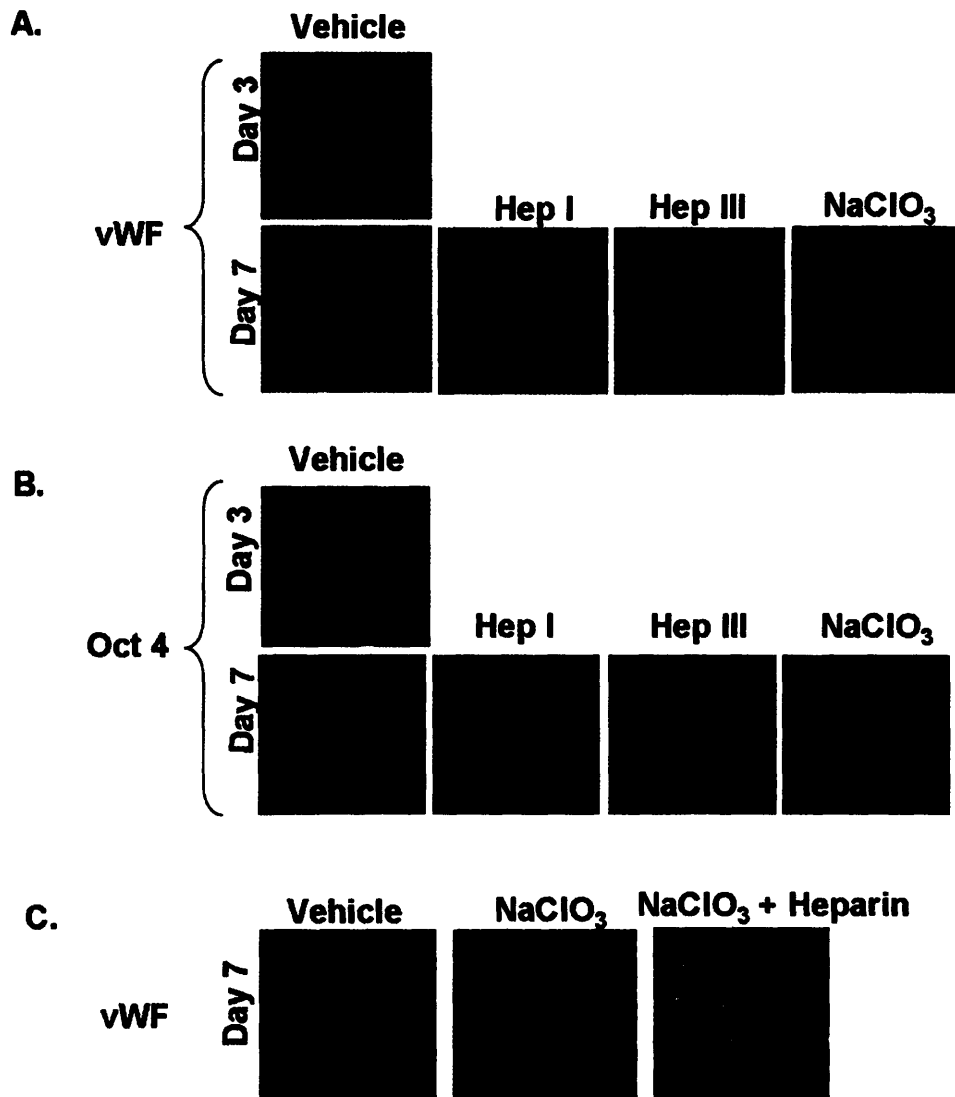
Figure 6.4. Confocal microscopy analysis of the effects of enzymatic or pharmacological modification of the HSGAGs on differentiation of ES cells into endothelial cells. (A) vWF staining. Extracellular degradation of HSGAGs by heparinase I and heparinase III treatment, as well as inhibition of HSGAG biosynthesis by sodium chlorate, inhibited differentiation of ES cells into endothelial cells as detected by vWF staining. (B) Oct-4 staining. Although differentiation towards endothelial cells is inhibited, overall differentiation still did proceed as evidenced by Oct-4 staining. (C) Reconstitution experiment by addition of heparin. Addition of exogeneous heparin to sodium chlorate treated ES cells reconstituted conditions that favor differentiation towards endothelial cells as detected by increased vWF staining.
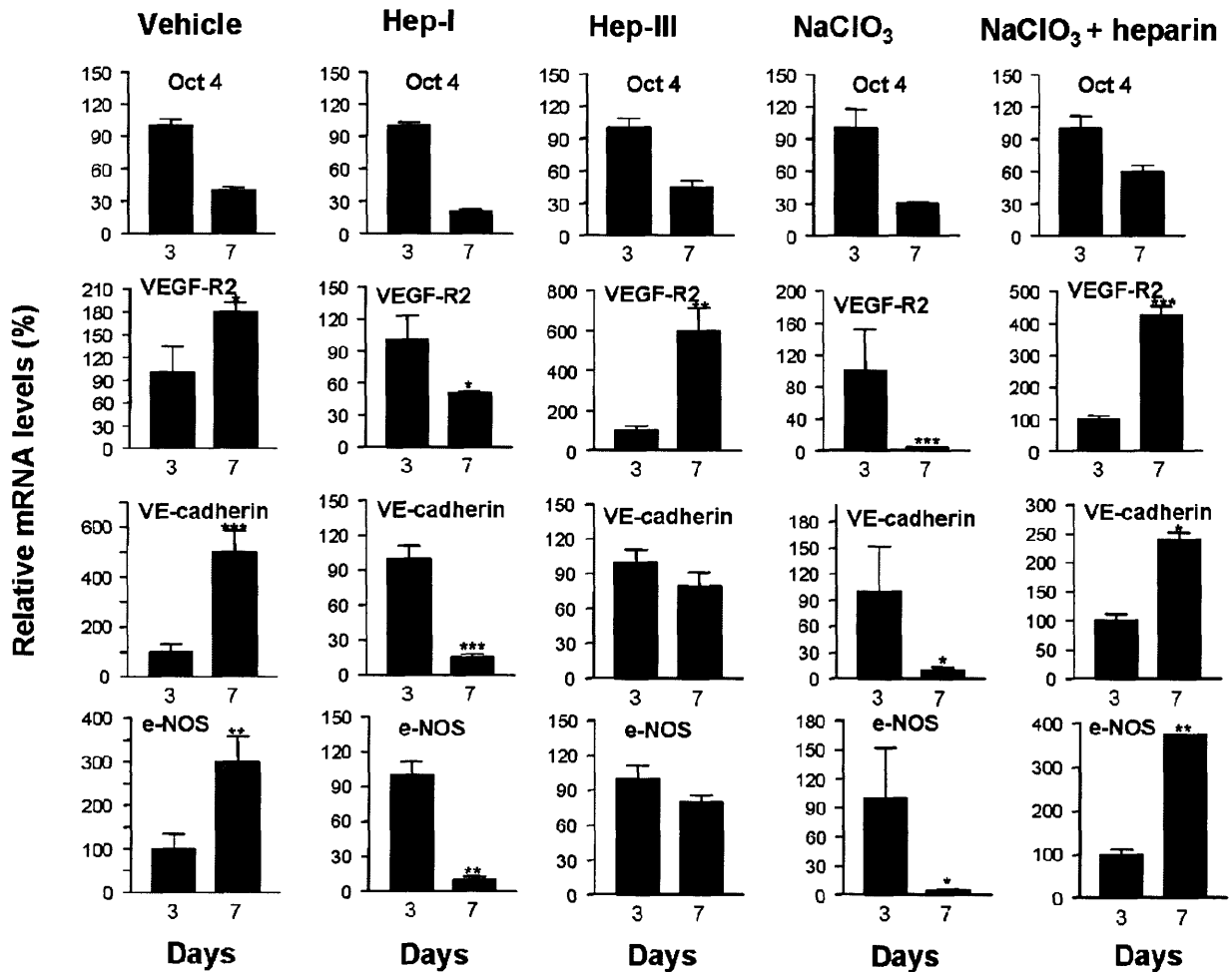
Figure 6.5. Real time PCR analysis of the effects of enzymatic or pharmacological modification of the HSGAGs on differentiation of ES cells into endothelial cells. Extracellular degradation of HSGAGs by heparinase I treatment, as well as inhibition of HSGAG biosynthesis by sodium chlorate inhibited differentiation of ES cells into endothelial cells (as detected by a decrease in VEGF-R2, VE-cadherin, and eNOS transcript levels), albeit differentiation still took place (as detected by a concurrent decrease in Oct-4 transcript levels). Although heparinase III treatment had effects consistent with Heparinase I in general, it showed an orthogonal effect for VEGF-R2 transcript levels. Addition of exogeneous heparin to sodium chlorate treated ES cells reconstituted conditions that favor differentiation towards endothelial cells as detected by increased VEGF-R2, VE-cadherin and e-NOS transcript levels.
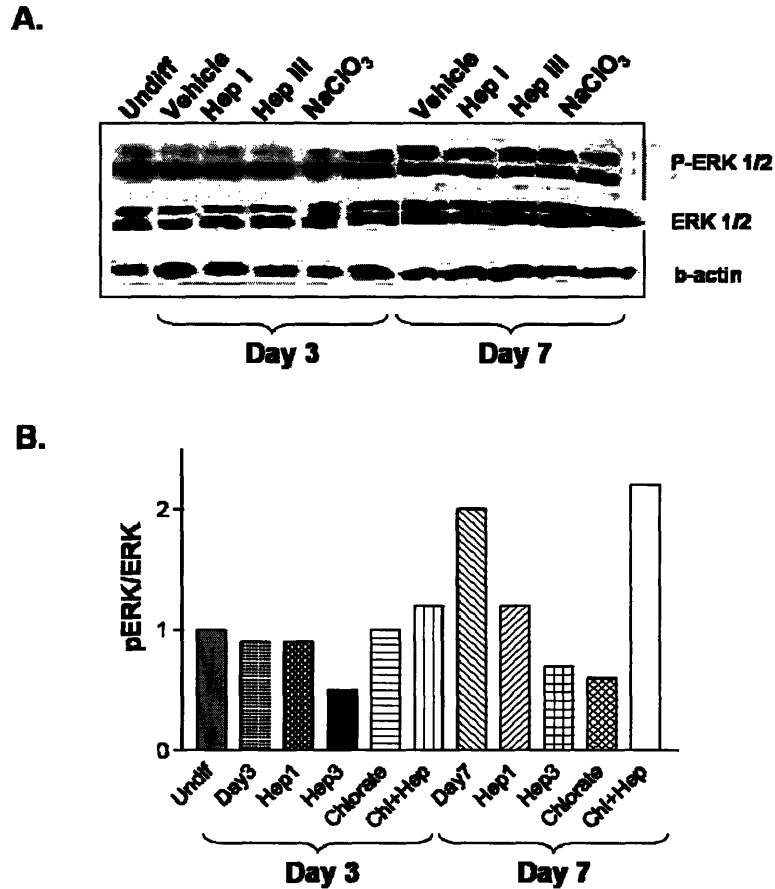
Figure 6.6. Effects of HSGAG modulation on the MAPK pathway in differentiating ES cells. (A) Western blots performed with ERK and phospho-ERK antibodies. Western blots showed that treatment with heparinases or sodium chlorate inhibited the phosphorylation of ERK. This inhibition is reversed by the addition of exogenous heparin. (B) The bar plot of (A) that shows the ratio of pERK/ERK with different treatments. These results suggest that MAPK pathway is involved in differentiation of ES cells, and HSGAGs are critical modulators of this pathway.

Chapter 7

Conclusion and Future Directions

Chapter 7

Conclusion and Future Directions

A significant part of a cell's glycome is constituted by Heparan Sulfate Glycosaminoglycans (HSGAGs) - the resident components of the extracellular matrix. With their 48 possible disaccharide repeat units, HSGAGs are the most information dense biopolymers. It is becoming increasingly clear that HSGAGs impinge on critical signaling pathways and modulate fundamental biological processes by binding and modulating the functions of many proteins. Thus, an understanding of the role that HSGAGs play in biological processes adds a very important dimension in our understanding of biological processes in the post-genomic era.

Understanding specific structure-function relationship of HSGAGs has been a very challenging task. This is primarily due to the high heterogeneity and structural diversity, combined with the low availability of HSGAGs on the cell surface and extracellular matrix. In addition, as a result of their highly charged character, it is difficult to isolate HSGAGs from biological sources in a highly purified state. As a result of these technical challenges, the development of tools to study specific structure-function relationship of HSGAGs lagged behind that of proteins and DNA. Recently, enzymatic approaches to cleave HSGAGs in a structurally specific fashion have been developed to reveal sequence information of HSGAGs, in a manner analogous to the use of restriction enzymes in DNA restriction mapping and the use of proteases in obtaining sequence information of proteins.

One of the main focuses of our laboratory has been to clone, recombinantly express and characterize HSGAG degrading enzymes from *Flavobacterium heparinum*. These

enzymes include heparinase-I, heparinase-II, heparinase-III, Δ4,5-glycuronidase, 2-O-sulfatase, 3-O-sultafase and N-sulfatase (*1*). With their unique substrate specificities, each of these enzymes serves as invaluable tools that can be used to obtain HSGAG sequence information. Prior to the studies explained in Chapter 5 of this thesis, the repertoire of HSGAG recombinant enzymes of our laboratory was limited to heparinase-I, heparinase-II and heparinase-III. Chapter 5 describes the cloning, recombinant expression and characterization of Δ4,5-glycuronidase, which is an invaluable addition to the repertoire of recombinant enzymes that can be used for characterizing HSGAG composition and fine structure.

The studies presented in Chapter 5 is foundational to future structure-function studies of Δ4,5-glycuronidase. This enzyme is particularly suitable for crystallographic structural studies due to the high yields of pure protein that can be obtained from bacterial cell cultures. Further biochemical characterization of Δ4,5-glycuronidase is warranted to reveal the amino acids critical for substrate binding and catalytic mechanism. Dissection of the catalytic mechanism of Δ4,5-glycuronidase would make it possible to engineer variants of this enzyme with unique substrate specificities, which can be added to the repertoire of HSGAG degrading enzymes to expand the tools available for HSGAG sequencing. In short, Δ4,5 glycuronidase is an interesting enzyme whose catalytic mechanism merits further investigation.

The ultimate goal of our laboratory is to have a broad repertoire of recombinant HSGAG degrading enzymes with unique substrate specificities, and apply these enzymes to the micro-sequencing of HSGAGs. With the addition of Δ4,5-glycuronidase, the current repertoire includes heparinase-I, heparinase-II, heparinase-III, Δ4,5-glycuronidase (*2-4*).

Very recently, another HSGAG degrading enzyme from *Flavobacterium heparinum*, 2-O sulfotase, was added to this repertoire (*5*). One of the future aims of the Sasisekharan lab is to continue expanding this *Flavobacterium*-derived repertoire of HSGAG degrading enzymes. A complete repertoire would be possible by the addition of several other enzymes such as, 6-O sulfatase, the 3-O sulfatase and the N-sulfatase, to the existing repertoire. A powerful HSGAG sequencing methodology based on heparinases, MALDI-MS, and capillary electrophoresis was previously developed (*6*), and each new addition of a HSGAG degrading enzyme to the repertoire enables the further improvement of this sequencing methodology.

In addition to their role in sequencing, *Flavobacterium*-derived HSGAG degrading enzymes also provide us with invaluable tools that can be used to modify HSGAGs of the cell surface and ECM to reveal the role of HSGAGs in biological processes. For instance, heparinases have been used to reveal the role of HSGAGs in tumor progression, metastasis and angiogenesis (*7, 8*). Chapter 6 of this thesis describes studies done to elucidate the role of HSGAGs in differentiation of ES cells into endothelial cells, which has tremendous implications in the field of regenerative medicine. In these studies, we demonstrated that modification of the HSGAG moiety of the ES cells, specifically by exogenous addition of heparinases, or by inhibiting their biosynthesis, inhibited ES cell differentiation into endothelial cells. Interestingly, differentiation still did proceed as evident from the progressive loss of the stem cell specific marker, indicating that cells other than of the endothelial lineage were yielded. Further investigation is necessary to determine which cell population was obtained as a result of differentiation when the HSGAG moiety was modified. Additionally, we showed that HSGAG modulation impinges on the mitogen activated protein kinase (MAPK) pathway, indicating that HSGAG modulation of ES cell

differentiation could possibly be mediated through the MAPK pathway. The dissection of the exact pathways impinged on by HSGAGs that modulate stem cell differentiation requires further investigation.

Our results show that HSGAGs are critical for ES cell differentiation into endothelial cells. Intriguingly, earlier studies from our laboratory had demonstrated that the HSGAGs are key for the survival of endothelial cells, and there is a switch from highly sulfated to undersulfated with an overall loss of HSGAG when the cells shift from a healthy to apoptotic state (*9*). This, in context of the observation that HSGAGs play a key role in the differentiation of stem cells to endothelial cells, raises a very interesting question: How does HSGAG fine structure impinge on the dynamic state of endothelial cells? For example, would inhibition of HSGAG sulfotranserases (*i.e.*, 3-O, 2-O or 6-O) inhibit differentiation of ES cells into endothelial cells? If so, which sulfotransferases are critical for differentiation? Similarly, would the inhibition of sulfotransferases induce apoptosis in an endothelial cell? If so, which sulfotransferases are critical? These questions are very important in understanding the effects of HSGAGs on the biology of endothelial cells and needs to be further explored. Most importantly, questions similar to these can be asked about different types of cells as well as distinct types of glycosaminoglycans to obtain a more global understanding of the role of glycosaminoglycans on the biology of a cell.

Additionally, the current finding has significant implications in blood vessel engineering. For example, by providing ES cells with the right type of HSGAG microenvironment, it might be possible to direct their differentiation towards endothelial cells - the essential components of blood vessels. Our results also show that enzymatic degradation of the HSGAGs of differentiating cells or ablation of the HSGAG-biosynthetic

machinery inhibited the formation of endothelial cells. This finding has important implications in regeneration of tissues other than blood vessels. For example, by inhibiting differentiation towards endothelial cells, it is possible to enrich the differentiating cell population in other cell types (*i.e.*, neurons, bone marrow and muscle tissue). In short, by modulating the HSGAG moiety of ES cells, it is possible to direct differentiation towards a more homogeneous cell population. Thus, the HSGAG paradigm we discovered in ES cell differentiation opens new approaches and strategies for regenerative medicine and enables the development of novel molecular therapeutics.

It is becoming increasingly clear that HSGAGs impinge on proteome, and as a result can also affect the transcriptome (*10*). The studies reported in Part B of this thesis are the first to show that differentiation of stem cells into endothelial cells is impinged on by the glycome, specifically the HSGAG component of the glycome. Thus, this study emphasizes the significance of the glycome as the third component after the proteome and the transcriptome, which should now be integrated into the systems biology approach for a molecular understanding of physiological and pathophysiological processes.

References

1. Ernst, S., Langer, R., Cooney, C. L., and Sasisekharan, R. (1995) *Crit Rev Biochem Mol Biol 30*, 387-444.
2. Sasisekharan, R., Bulmer, M., Moremen, K. W., Cooney, C. L., and Langer, R. (1993) *Proc Natl Acad Sci U S A 90*, 3660-4.
3. Godavarti, R., Davis, M., Venkataraman, G., Cooney, C., Langer, R., and Sasisekharan, R. (1996) *Biochem Biophys Res Commun 225*, 751-8.
4. Shriver, Z., Hu, Y., Pojasek, K., and Sasisekharan, R. (1998) *J Biol Chem 273*, 22904-12.
5. Myette, J. R., Shriver, Z., Claycamp, C., McLean, M. W., Venkataraman, G., and Sasisekharan, R. (2003) *J Biol Chem 278*, 12157-66.
6. Venkataraman, G., Shriver, Z., Raman, R., and Sasisekharan, R. (1999) *Science 286*, 537-42.
7. Liu, D., Shriver, Z., Venkataraman, G., El Shabrawi, Y., and Sasisekharan, R. (2002) *Proc Natl Acad Sci U S A 99*, 568-73.
8. Sasisekharan, R., Moses, M. A., Nugent, M. A., Cooney, C. L., and Langer, R. (1994) *Proc Natl Acad Sci U S A 91*, 1524-8.
9. Johnson, N. A., Sengupta, S., Saidi, S. A., Lessan, K., Charnock-Jones, S. D., Scott, L., Stephens, R., Freeman, T. C., Tom, B. D., Harris, M., Denyer, G., Sundaram, M., Sasisekharan, R., Smith, S. K., and Print, C. G. (2004) *Faseb J 18*, 188-90.
10. Sasisekharan, R., and Venkataraman, G. (2000) *Curr Opin Chem Biol 4*, 626-31.

TANYEL KIZILTEPE
391 Broadway, Apt 505, Somerville, MA 02145
(617) 7928917
kiziltep@mit.edu

EDUCATION

Department of Chemistry, Massachusetts Institute of Technology (Cambridge, MA)
Ph.D. in Biochemistry (1999 – 2005).
Co-supervisors: Professor Bevin Engelward and Professor Ram Sasisekharan
Thesis: Nitric Oxide-Induced DNA Recombination & Glycosaminoglycan Mediated Differentiation in Stem Cells

Department of Chemistry, Bilkent University (Ankara, Turkey)
B.S. in Chemistry (1993-1998).

AWARDS

- Fulbright Fellowship for Doctoral Studies. Twelve fellowships awarded nationally (1999).
- Bilkent University Full-Tuition Merit Scholarship. Ten scholarships awarded nationally (1993-1998).
- Turkish Science & Technology Institute Award. Thirty scholarships awarded nationally (1993-1998).

RESEARCH and TEACHING EXPERIENCE

Department of Biological Engineering, Massachusetts Institute of Technology (Cambridge, MA)
Teaching Assistant of the Graduate Level Course: Mechanisms of Drug Actions (Fall 2004).

Department of Biological Engineering, Massachusetts Institute of Technology (Cambridge, MA)
Research Assistant of Prof. Ram Sasisekharan and Prof. Bevin Engelward (2001-2005).

- Elucidated the chemical mechanism of inflammation-induced DNA recombination in mammalian cells to understand the relationship between inflammation and cancer. Specifically, I developed molecular biological and genetic tools to study homologous recombination, and the effects of replication on homologous recombination in mammalian cells. By using these tools, I dissected the types of reactive nitrogen species and the classes of DNA lesions that induce recombination in mammalian cells. This work is currently in press *(Chemistry & Biology).*

- Dissected the role of glycosaminoglycans in the differentiation of embryonic stem cells into endothelial cells and elucidated the pathways that are impinged upon by the modulation of the glycosaminoglycans during the differentiation process. This study has potential applications in both tissue regeneration and the inhibition of tumor neovascularization. We have filed a patent and are currently developing the data into a manuscript.

- Developed a dual-color fluorescence imaging-based system for the dissection of anti-angiogenic and chemotherapeutic activity of molecules. This system enables the study of tumor-angiogenesis relationship on the same platform, expanding the rapid screening of novel molecules in a simulated clinicopathological setting. This study was published in FASEB J. (13):1565-7 (2004). Furthermore, this technology was used to develop a novel nanocell based cancer therapy through temporal targetting of both tumor cells and neovasculature. This work is currently being reviewed by Nature.

- Cloned, expressed and kinetically characterized a glycosaminoglycan degrading enzyme - $\Delta4,5$ glycuronidase. This study was published in Biochemistry 41(23):7424 (2002). This enzyme, with its use in dissecting structure-function relationships of glycosaminoglycans, served as one of the building blocks of Momenta Pharmaceuticals.

Department of Chemistry, Massachusetts Institute of Technology (Cambridge, MA)
Research Assistant of Prof. Lawrence Stern (1999-2001)

- Expressed and purified recombinant MHC proteins. Performed thermodynamic and kinetic measurements for receptor - ligand binding interactions for cell surface MHC molecules.

Max Planck Institute for Biophysics (Frankfurt, Germany)
Summer Intern (1999): Purified and crystallized proteins for electron microscopy studies

Asta Medica, Cancer Research Department (Frankfurt, Germany)
Summer Intern (1998): Synthesized peptides to be screened as potential anti-cancer therapeutics.

LEADERSHIP and TEAM WORK

MIT $50K Entrepreneurship Competition-2004 (Cambridge, MA)
Idea Creator and Team Coordinator (2004)
- Managed the team Dynamic BioSystems to develop a business plan on protein stabilization platform technologies.
- Selected as a semi-finalist; idea publicized in Boston Business Journal.

Division of Biological Engineering, Massachusetts Institute of Technology (Cambridge, MA)
Supervisor of Undergraduate Students (2002-present)
- Supervised undergraduate students' research projects.
- Coordinated undergraduate students to accomplish routine laboratory tasks.

TECHNICAL EXPERIENCE

Cell Biology / Toxicology / Immunology: Tissue Culture, Toxicity Assays (Proliferation, Survival and Apoptosis), Flow Cytometry (Cell Cycle Analysis, Immunofluorescence Staining), Immunoprecipitation and Western Blots, T Cell / B Cell Activation Assays, Glycosaminoglycan Analysis.
Protein Biochemistry / Enzymology: Recombinant Protein Expression, Refolding and Purification of Proteins (Nickel column, Ion Exchange, HPLC, Gel Chromatography, SDS page), Steady State Kinetic Measurements (UV/Vis Spectrometry), ELISA, FRET, Surface Plasmon Resonance, Mass Spectrometry.
Molecular Biology / Genetics: Gene Cloning, Recombinant DNA Technology, Genetic Engineering of Mammalian Cells, Stable and Transient Gene Expression, Southern Blots, Real Time PCR.
Animal Studies: Injection and dissection of tumors in mice, Histology.

SKILLS, INTERESTS and ACTIVITIES

Technical: Experienced with Fortran 77 and Matlab. Proficient in MS Word, Excel, PowerPoint and Photoshop.
Languages: Turkish, native; English, fluent; German, intermediate.
Athletics: National Champion of Turkey in Ladies Figure Skating (1993), and the winner of several other national medals. Internationally certified Figure Skating Trainer.

PUBLICATIONS

1. "Delineation of the Chemical Pathways Underlying Inflammation-Induced DNA Recombination in Mammals," **Kiziltepe T**, Yan A, Dong M, Jonnalagadda VS, Dedon PC, Engelward BP, Chemistry & Biology (in press, 2004).
2. "Effects of Heparan Sulfate Glycosaminoglycans on Stem Cell Differentiation into Endothelial Cells", **Kiziltepe T**, Sengupta S, Eavarone D, Keiser N, Sasisekharan R (manuscript in preparation).
3. "A dual-color fluorescence imaging-based system for the dissection of anti-angiogenic and chemotherapeutic activity of molecules," Sengupta S, **Kiziltepe T**, Sasisekharan R, FASEB J. (13):1565-7 (2004).
4. "Novel Cancer Therapy Through Temporal Targetting Of Both Tumor Cells And Neovasculature Using Nanocells," Sengupta S, Eavarone D, Zhao G, Capila I, Watson N, **Kiziltepe T**, Sasisekharan R, Nature (in review, 2004).
5. "Molecular cloning of the heparin/heparan sulfate Δ4,5 unsaturated glycuronidase from Flavobacterium-heparinum, its recombinant expression in Escherichia coli, and biochemical determination of its unique substrate specificity," Myette JR, Shriver Z, **Kiziltepe T**, McLean MW, Venkataraman G, Sasisekharan R, Biochemistry 41(23):7424 (2002).
6. "Theoretical Analysis of Substituent Effects on Building Blocks of Conducting Polymers: 3,4'-Substituted Bithiophenes," Salzner U, **Kiziltepe T**, J Org Chem. 64(3):764 (1999).

PATENTS

1. "Modulation of Embryonic Stem Cell Differentiation," **Kiziltepe T**, Sengupta S, Eavarone D, Sasisekharan R, USPTO (filed, 2004).