RICE UNIVERSITY

Evolution of the Perlecan/HSPG2 gene and Regulation of its Expression by Inflammatory Cytokines in Normal Tissue Models and Cancer

by

Curtis Robert Warren, M.S.

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, DISSERTATION COMMITTEE

Mary C. Farach-Carson
Ralph and Dorothy Looney Professor of Biochemistry and Cell Biology
Vice Provost for Translational Bioscience

Daniel D. Carson
Dean, Natural Sciences
Schlumberger Chair of Advanced Studies and Research
Professor of Biochemistry and Cell Biology

K. Jane Grande-Allen
Professor of Bioengineering & Associate Chair, External Partnerships

Michael Gustin
Professor of Biochemistry and Cell Biology

Daniel Wagner
Associate Professor of Biochemistry and Cell Biology

HOUSTON, TEXAS
2014
ABSTRACT

Evolution of the Perlecan/HSPG2 gene and Regulation of its Expression by Inflammatory Cytokines in Normal Tissue Models and Cancer

by

Curtis Robert Warren

Perlecan is the large heparan sulfate proteoglycan common to all basement membranes. It has numerous functions in maintenance of BM integrity, cell signaling and scaffolding protein interactions. Perlecan accumulation is elevated in wound healing and is essential to organismal development. In this work the evolution of perlecan and its role in the simplest and most ancient animals are explored. Transcriptional regulation of the HSPG2 gene also is examined in human prostate cancer and associated stromal cells. The protein was elevated in the reactive stroma of primary prostate cancer and TNF-α was identified as the primary driver of HSPG2 expression induction in various prostate cancer, prostate stromal and bone marrow stromal cell lines. Various aspects of this response echo the fibroblastic response to wounding and tumor progression. HSPG2 homologues were found in the genomes of the cnidarian, Nematostella vectensis, and the placozoan, Trichoplax adhaerens. Thus the last common ancestor to encode a perlecan homologue is the placozoan Trichoplax adhaerens. N. vectensis perl elevation was identified as part of the gene expression profile of complex regenerating structures in the oral region of the animal following wounding. This is a conserved expression pattern of the gene which is still found in wound healing of modern mammals. These studies both demonstrate a role for perlecan in wound healing and pathological states, corroborating the hypothesis that the perlecan gene’s primary evolutionary role is to support tissues in times of remodeling.
Dedication

I dedicate this work to my parents; my father Dr. Kevin Warren who taught me to marvel at nature, and my mother Cyndy Warren who taught me the importance of kindness.
Acknowledgements

First and foremost I must acknowledge the boundless generosity and patience of my advisor Cindy Farach-Carson. I was fortunate to work in her lab as an undergraduate with no experience and enthusiasm as my only qualification. Seven years later through her constant guidance and mentorship I have achieved a level of competency that I hope will live up to her standard.

It is important to thank all members of the Farach-Carson and Carson labs. Especially Dan Carson, who has been an unofficial secondary advisor for all of us in Cindy’s lab. His input has been instrumental in my development as a scientist.

I would like to thank my committee members, Drs. Michael Gustin, Daniel Wagner and Matthew Bennett who have steered my research during yearly progress reviews, as well as Dr. Jane Grande-Allen and Dr. Carson who participated in my defense. Critical collaborators were Lewis Francis, James Spurlin, Nik Putnam, Carlos Nossa and most of all the undergraduate student who worked with me Elias Kassir.

Last I should thank my wife Maggie for understanding the move to Texas to pursue this project, and for moving with me although living with the heat is not really her cup of tea. She made all of the hard times easy.
# Table of Contents

List of tables ..................................................................................................................... XI

List of illustrations .............................................................................................................. XII

List of abbreviations .......................................................................................................... XV

Chapter 1: General Introduction (partly as published in Farach-Carson et al. 2013 in Matrix Biology) ......................................................................................................................... 1

1. Perlecan/HSPG2 gene and protein; perlecan expression in normal biology ………. 1
   a. Structure of the HSPG2 gene and its conservation ............................................. 1
   b. Structure of the perlecan protein and its conservation ....................................... 4
      i. The structure of perlecan protein domains ....................................................... 5
         1. Domain I ............................................................................................................. 7
         2. Domain II ........................................................................................................... 7
         3. Domain III ....................................................................................................... 8
         4. Domain IV ....................................................................................................... 9
         5. Domain V ........................................................................................................... 13

2. Perlecan expression in normal tissues and the effect of HSPG2 loss ………………… 14
   a. Perlecan in the basement membrane ................................................................. 14
   b. Perlecan in pericellular matrices ...................................................................... 20
   c. Disorders associated with the HSPG2 mutation ............................................... 21

3. Perlecan function in invertebrates gives clues to its evolutionary origin ………….. 22
   a. unc-52 .................................................................................................................... 22
4. trol……………………………………………………………………………………………………23
5. Evolutionary clues about perlecan and BM assembly………………………………25
6. Prostate cancer biology……………………………………………………………………26
   a. Epidemiology…………………………………………………………………………………26
   b. Primary disease……………………………………………………………………………28
      i. Prostatic intraepithelial neoplasm…………………………………………………..29
      ii. The Gleason scoring system………………………………………………………….30
      iii. Stromal reaction……………………………………………………………………..30
      iv. Tumor-associated macrophages…………………………………………………..34
   c. Prostate cancer metastasis……………………………………………………………..34
      i. Epithelial to mesenchymal transition………………………………………………34
      ii. Bone metastasis……………………………………………………………………..35

Chapter 2: HSPG2/perlecan gene and protein expression changes during prostate
cancer disease progression……………………………………………………………………38

1. Introduction…………………………………………………………………………………..38
2. Materials & Methods ………………………………………………………………………39
   a. Cell culture………………………………………………………………………………….39
      i. Treatments………………………………………………………………………………40
      ii. Co-cultures……………………………………………………………………………..40
   b. Plasmid construction…………………………………………………………………..41
   c. Transient transfection………………………………………………………………….41
   d. Real-time polymerase chain reaction (see chapter 3)……………………………..41
   e. Immunoblotting for secreted perlecan (see chapter 3)………………………..42
f. WST-1 assay for cell number (see chapter 3).................................42

g. Genomatix analysis of the HSPG2 5' upstream region.................42

h. Luciferase reporter assay (see chapter 3)..................................42

3. Results..........................................................................................43

a. HSPG2 mRNA expression in prostate cancer, bone marrow and prostate stromal cell lines and primary bone marrow stromal cells.........43

b. Perlecan protein secretion in prostate cancer, bone marrow and prostate stromal cell lines and primary bone marrow stromal cells.........45

c. Changes in cell number during cytokine treatments.....................45

d. Analysis of perlecan promoter in silico......................................49

e. HSPG2 promoter-reporter constructs used in this study..............51

f. HSPG2 promoter-reporter activity in prostate cancer, bone marrow and prostate stromal cell lines in response to treatments and co-culture.....53

g. Morphological changes in stromal cells induced by inflammatory cytokines.................................................................61

4. Discussion.......................................................................................63

Chapter 3: Regulation of the HSPG2 gene by inflammatory cytokines (as published in Warren et al. 2014 in the Journal of Cellular Biochemistry)..............68

1. Abstract.........................................................................................69

2. Introduction..................................................................................70

3. Materials and Methods...............................................................73

a. Immunohistochemistry.........................................................73

b. Cell culture................................................................................73
c. Plasmid construction ................................................................. 74

d. Transient transfection ............................................................ 75

e. Luciferase reporter assay ......................................................... 75

f. Real-time polymerase chain reaction ........................................ 75

g. Immunoblotting for secreted perlecan ....................................... 76

h. WST-1 assay for viable cell number .......................................... 76

i. TNF-α enzyme-linked immunosorbent assay ............................. 77

j. Immunostaining for NFκB ......................................................... 77

k. Quantitative chromatin immunoprecipitation ............................. 78

l. Statistics .................................................................................. 78

4. Results ..................................................................................... 79

a. Perlecan is upregulated in desmoplastic stroma of PCa. ............... 79

b. Perlecan secretion and HSPG2 mRNA expression increase in response to cytokine treatment. ......................................................... 81

c. The HSPG2 promoter has conserved NFκB binding elements more than 2.4 kb upstream of the start site of transcription. ......................... 84

d. Bone marrow stromal, prostate stromal and prostate cancer cell lines secrete very low levels of TNF-α ......................................................... 86

e. The HSPG2 promoter is activated by TNF-α in a dose-dependent manner. .86

f. The HSPG2 promoter is active in stromal and prostate cancer cells ....... 87

g. TNF-α treatment of HS-27a and WPMY-1, but not HS-5, stromal cell lines increases NFκB nuclear localization ............................................. 88
h. *HSPG2* promoter responsiveness to TNF-α resides more than 2.4 kb upstream of the start site of transcription in the more distal NFκB binding site.

i. NFκB binds the distal region of the *HSPG2* promoter.

5. Discussion

Chapter 4: Evolution of the *HSPG2* gene and *HSPG2* in tissue regeneration (as prepared for submission)
e. Phylogenetic analysis of *HSPG2* homologues……………………………..116

f. *Perl* expression profile in *T. adhaerens* …………………………………119

g. *Perl* expression profile in *N. vectensis* ………………………………..119

h. *Perl* expression is activated during *N. vectensis* tissue regeneration………………………………………………………………………………121

5. Discussion…………………………………………………………………………123

Chapter 5: Summary, General Discussion & Future Directions…………………….128

1. Summary: Perlecan expression in prostate cancer and stromal cells…………128

2. Summary: Perlecan in basal metazoans………………………………………133

3. Future Directions for the study of perlecan in prostate cancer………………136

4. Future directions for the study of perlecan and ECM in basal metazoan……138

References…………………………………………………………………………142

Appendix A. Oligonucleotide primers and riboprobes used in this study………164

Appendix B. *HSPG2* gene ChIP-seq data from ENCODE…………………….166

Appendix C. *T. adhaerens* gene orthologues for heparan sulfate pathway components………………………………………………………………………………167

Appendix D. Analysis of *O. carmela* and *M. leidyi* perlecan BLAST hits for perlecan………………………………………………………………………………168

Appendix E. A circular isoform of the *HSPG2* transcript……………………169
List of Tables

Table 1.1 Perlecan binding partners involved in tissue structure and function ..........18

Table 1.2 Bound active molecules that can be released by perlecan turnover ..........19

Table A.1 Oligonucleotide primers used in this study ......................................164

Table C.1 T.adhaerens gene orthologues for heparan sulfate pathway components .....170
List of Illustrations

Figure 1.1 Conservation of the perlecan gene and protein ..................................................3

Figure 12 Schematic of perlecan ....................................................................................6

Figure 1.3 Colorized analysis of Ig-modules in domain IV showing two clear patterns...12

Figure 1.4 The typical progression of prostate cancer .....................................................33

Figure 2.1 Basal transcription of HSPG2 mRNA in various cell lines .........................44

Figure 2.2 Perlecan secretion varies during cytokine treatment in different primary bone
marrow stromal cell populations .................................................................................47

Figure 2.3 Cell number is affected by treatment with TNF-α and/or TGFβ1 ...............48

Figure 2.4 Conserved CREs in the region 5′ of the HSPG2 gene .............................50

Figure 2.5 HSPG2 promoter-reporter constructs used in this study ..........................52

Figure 2.6 Baseline promoter-reporter activity levels are highest in stromal cell lines...54

Figure 2.7 The 50 base pair region housing consensus NFκB elements is insufficient to
drive the TNF-α response .......................................................................................56

Figure 2.8 HSPG2 promoter-reporter activity levels are unresponsive to many
treatments .............................................................................................................58

Figure 2.9 HSPG2 promoter-reporter activity levels are unresponsive to physical co-
culture and conditioned medium .......................................................................60
Figure 2.10 Stromal cell morphology changes under TNF-α and TGFβ1 combination treatment…………………………………………………………………………………62

Figure 3.1 Perlecan accumulates in the reactive stroma of prostate cancer…………………80

Figure 3.2 Expression of HSPG2 mRNA and perlecan increase in HS-27a and LNCaP cells during TGFβ1 and TNF-α treatment…………………………………………………………83

Figure 3.3 Conserved NFκB consensus elements are present in the HSPG2 promoter….85

Figure 3.4 HSPG2 promoter activity increases in response to TNF-α in a dose-dependent manner…………………………………………………………………………………89

Figure 3.5 TNF-α treatment triggers nuclear translocation of NFκB in HS-27a and WPMY-1, but not HS-5, cells………………………………………………………………………90

Figure 3.6 Deletion of the distal 200 base pairs of the HSPG2 promoter or mutation of NFκB consensus elements eliminates responsiveness to TNF-α and TGFβ1……………92

Figure 3.7 p65 binds the HSPG2 5’ regulatory region in LNCaP cells during TNF-α treatment…………………………………………………………………………………………94

Figure 3.8 Model for HSPG2 induction by tumor and immune cell interaction………96

Figure 4.1 Perlecan orthologues or constituents in model organisms………………..115

Figure 4.2 Molecular phylogenetic analysis of 76 perlecan homologues……………..117

Figure 4.3 T. adhaerens perl 1 expression in the medial cell layer………………..118

Figure 4.4 Perl is expressed in the ectoderm of the scapus, irregularly in the mesentery, and irregularly in the inner cell layer of the tentacles of N. vectensis…………………120
Figure 4.5 *Perl* is expressed in the regenerating oral region and tentacles at stage 2.5 through stage 4-5………………………………………………………………………………122

Figure A.1 Riboprobes synthesized for *T. adhaerens* and *N. vectensis in situ* hybridization…………………………………………………………………………………………165

Figure A.2 Negative controls for *in situ* hybridization of *N. vectensis*……………………………………………………………………………………………………166

Figure B.1 Histone modifications of the 5’ region and intron 1 of the *HSPG2* gene…..167

Figure B.2 TNF-α secretion by cell lines…………………………………………………………168

Figure B.3 Negative controls for immunofluorescent imaging of p65 nuclear localization…………………………………………………………………………………………169

Figure D.1 Search of *O. carmela* transcriptome for *HSPG2* homologues………………171

Figure E.1 The *HSPG2* transcript is circularized in bone marrow stromal cells and contains a heretofore unrecognized exon…………………………………………………………174
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADT</td>
<td>androgen deprivation therapy</td>
</tr>
<tr>
<td>ASW</td>
<td>artificial seawater</td>
</tr>
<tr>
<td>BM</td>
<td>basement membrane</td>
</tr>
<tr>
<td>CNS</td>
<td>central nervous system</td>
</tr>
<tr>
<td>CRE</td>
<td>cis-responsive element</td>
</tr>
<tr>
<td>DDSH</td>
<td>Dyssegmental Dysplasia, Silverman Handmaker type</td>
</tr>
<tr>
<td>DV</td>
<td>domain five</td>
</tr>
<tr>
<td>ECM</td>
<td>extracellular matrix</td>
</tr>
<tr>
<td>EHS</td>
<td>Engelbroth-Holm-Swarm</td>
</tr>
<tr>
<td>EMT</td>
<td>epithelial-mesenchymal transition</td>
</tr>
<tr>
<td>GAG</td>
<td>glycosaminoglycan</td>
</tr>
<tr>
<td>GS</td>
<td>Gleason score</td>
</tr>
<tr>
<td>HGPIN</td>
<td>high-grade prostatic intraepithelial neoplasia</td>
</tr>
<tr>
<td>HSV-TK</td>
<td>herpes simplex virus thymidine kinase</td>
</tr>
<tr>
<td>HSPG2</td>
<td>heparan sulfate proteoglycan 2</td>
</tr>
<tr>
<td>Ig</td>
<td>immunoglobulin</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>IL-1β</td>
<td>interleukin 1-beta</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>LDL</td>
<td>low-density lipoprotein</td>
</tr>
<tr>
<td>MET</td>
<td>mesenchymal-epithelial transition</td>
</tr>
<tr>
<td>miRNA</td>
<td>micro RNA</td>
</tr>
<tr>
<td>MMP</td>
<td>matrix metalloproteinase</td>
</tr>
<tr>
<td>NED</td>
<td>neuro-endocrine differentiation</td>
</tr>
<tr>
<td>NF-κB</td>
<td>nuclear factor kappa-light-chain enhancer of activated T cells</td>
</tr>
<tr>
<td>NMJ</td>
<td>neuromuscular junction</td>
</tr>
<tr>
<td>PAT</td>
<td>paralyzed arrest at embryonic two-cell stage</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PIN</td>
<td>prostatic intraepithelial neoplasia</td>
</tr>
<tr>
<td>RACE</td>
<td>rapid amplification of cDNA ends</td>
</tr>
<tr>
<td>SEA</td>
<td>sperm, enterokinase, and agrin</td>
</tr>
<tr>
<td>SJS</td>
<td>Schwartz-Jampel Syndrome</td>
</tr>
<tr>
<td>TAM</td>
<td>tumor-associated macrophage</td>
</tr>
<tr>
<td>TGFβ</td>
<td>transforming growth factor beta</td>
</tr>
<tr>
<td>TNF-α</td>
<td>tumor necrosis factor alpha</td>
</tr>
</tbody>
</table>
Chapter 1: General Introduction

The heparan sulfate proteoglycan 2 (HSPG2) gene encodes perlecan, a heparan-sulfate proteoglycan of basement membranes (BM) and pericellular matrices of various cells and tissues. The protein is essential for the development and maintenance of boundaries between tissue types, i.e. epithelium or vasculature from stroma. The protein functions through an array of protein-protein and protein-glycosaminoglycan (GAG) interactions.

Regulation of expression of the HSPG2 gene has been studied in various cell types, but never in the context of prostate cancer or associated stromal cell types. Regulatory mechanisms must support the protein’s conserved evolutionary function, but the evolutionary origins of HSPG2 may be complex, involving coalition of protein domains used in other contexts in simpler organisms. Perlecan is key to metazoan development, important to response to injury [Kinsella et al., 2003; Nikkari et al., 1994] and in cancer progression. This dissertation is a general study of evolution of the HSPG2 gene and its regulated expression in prostate cancer and associated stromal cells.

Structure of the HSPG2 gene and its conservation

The human HSPG2 gene is located on the reverse strand of chromosome 1 and spans 115,000 base pairs. There are 97 exons that encode the transcript, which is 14,327 base pairs in length in its primary form (figure 1.1A). The gene is conserved in all members of the animal kingdom studied to date, including C. elegans. Until this study, the presence of HSPG2 homologues has not been verified below nematodes on the evolutionary tree. In accordance with most findings, the most conserved parts of the gene are coding regions. In defiance of this normal trend, the 40 kilobase pair intron 1 of the HSPGP2 gene also is...
well conserved amongst vertebrates (green track, figure 1.1A). This is further discussed in chapter 3, where I propose this is part of the regulatory region of the HSPG2 gene.

The HSPG2 gene promoter that drives and modulates perlecan expression in vertebrates has been defined as the 2.7 kilobase (kb) long region lying immediately upstream of the start site of transcription [Iozzo et al., 1997]. The proximal promoter region has the features of a CpG island promoter, lacking a TATA box. There are three transcriptional start sites as determined by primer extension assay [Cohen et al., 1993], and a short 5′ UTR consisting of 80 base pairs. Putative cis-responsive elements (CREs) were identified previously in the HSPG2 promoter [Iozzo et al., 1997]. To better understand the evolutionary conservation of this promoter, we performed an in silico analysis using the PhastCons [Siepel et al., 2005] function of the UCSC genome browser, along with a series of complementary analyses using publically available programs. This analysis revealed evolutionary conservation in the noncoding region upstream of the HSPG2 gene amongst human, chimpanzee, rhesus monkey and mouse genomes, with much less conservation between the upstream regions of the HSPG2 gene in human and zebrafish genomes. The 20kb upstream region of the HSPG2 gene includes 78 putative CREs as identified by the presence of known conserved sequences that interact with known transcription factors. Nineteen of these conserved putative CREs reside in the proximal 2.7 kb. Upstream of this conserved region lie approximately three kb of non-conserved sequence, followed by the next highly conserved upstream region. This indicates that the majority of functional cis elements in the HSPG2 promoter likely are found within 2.7 kb upstream of the transcriptional start site. Although additional putative CREs can be identified in the promoter region, when investigating the human genome alone, I suggest
Figure 1.1 Conservation of the perlecan gene and protein.  (A) The perlecan gene is represented in schematic form.  Vertical hashes represent coding exons, color coded to represent what type of folding module they encode.  Numbering represents how many base pairs of the gene are covered; small hashes on this line represent 10,000 base pairs.  Green vertical lines below the gene schematic represent conservation based upon the PhastCons statistical model, using the perlecan gene of 46 vertebrates as comparison (http://genome.ucsc.edu/).  Asterisk notes the location of a pseudogene, RPL21P29, similar to ribosomal proteins.  (B) The perlecan protein is represented in schematic form.  Numbering on the top line represents amino acids at certain positions.  Hashes on the second black line denotes where the arbitrary perlecan domains begin and end.  Vertical hashes on the schematic represent exon-exon junctions, and rectangles represent independently folding modules, color coded as described at the bottom of the figure.  Below the schematic is displayed conservation of the protein amongst four vertebrates, human, mouse, chicken and zebrafish (denoted by H, M, C, Z next to the grey bars).  Green lines represent 4-way protein identity, yellow lines represent 2 or 3-way identity, and red lines indicate no match between species.  The 4 grey lines at the bottom represent the protein sequences of perlecan from each species, and how well they match to the human protein.  Note that the line in domain IV of the chicken and zebrafish proteins, unmarked by an attending rectangle, means there is in fact no protein at those locations (domain IV in zebrafish is much shorter than in humans).  From Farach-Carson et al. 2014.
that the nineteen conserved CREs proximal to the transcriptional start site of HSPG2 are the most likely to function in primary gene regulation to produce the primary perlecan transcript (figure 1.1A).

**Structure of the perlecan protein and its conservation**

The perlecan core protein historically has been described in terms of five functional domains (figure 1.2), each of which contains a group of distinct functional modules that can interact with a variety of other proteins and molecules present in the extracellular matrix (ECM) (Table I). These modules, especially those nearer the C-terminus, arose early in evolution, suggesting a critical function that was needed early to establish the boundaries of functional tissues. It has been suggested that intact perlecan with all functional modules linked together could serve as an extracellular scaffold large enough to hold together several lipid rafts in the plasma membrane of cells exposed to the ECM [Farach-Carson and Carson, 2007]. An analysis of the assembly of perlecan’s multi-modular structure in lower species compared to that in humans provides a unique way to examine the evolution of perlecan’s function in tissues of multicellular organisms.

The perlecan protein is composed of a chain of modular motifs homologous to those of other ECM proteins [Murdoch et al., 1992; Noonan et al., 1991]. This structural organization is conserved amongst vertebrates and invertebrates alike [Hummel et al., 2004; Noonan et al., 1991; Rogalski et al., 1993; Voigt et al., 2002]. The domains of perlecan are assigned based upon shared homology with other proteins (figure 1.2), with human perlecan comprising five domains and perlecan in nematodes comprising only domains II, III, and IV [Rogalski et al., 1993]. The nature of mammalian perlecan’s modular structure has been reviewed at length [Farach-Carson and Carson, 2007] Herein
is described the known and putative functions of each domain in perlecan, and the extent of knowledge concerning each domain’s contribution to the proposed barrier function of perlecan, with a focus on evolutionary insights.

The structure of perlecan protein domains

At some point during the evolution of organisms with multiple cell layers, it appears that selective pressure forced perlecan’s organization as a single several-hundred kiloDalton chain of modular repeating motifs. Speculation that the protein serves as an ECM scaffold is borne out by the variety of binding partners for the protein core and heparan sulfate chains. Recent work has added molecular details to show how perlecan links the two separate protein networks of collagen IV and laminin which together constitute the BM [Behrens et al.]. This work demonstrated binding of the laminin network by the perlecan core protein and binding of the collagen IV network by perlecan’s heparan sulfate chains, which corroborates previous studies [Brown et al., 1994; Tillet et al., 1994]. Despite this, heparan sulfate-dependent binding of collagen IV has little bearing on perlecan’s function in early metazoans, as to date reported invertebrate perlecan molecules lack domain I and thus most of the heparan sulfate attachment sites [Hummel et al., 2004], although these species have the machinery to make heparan sulfate [Toyoda et al., 2000]. Certain of perlecan’s interactions are crucial for maintaining tissue stability during development, and I will describe later in this dissertation cases where loss of or deficiencies in perlecan protein can abrogate the barrier function of the molecule.
Figure 1.2 Schematic of perlecan. Modular structure of perlecan chain depicted as train cars assembled in tandem behind a common “engine” that is unique to perlecan. The individual modules are self-folding, joined by linker sequences between. The locomotive depicts the signal sequence driving the train into the cellular secretory path for secretion into the extracellular space. The SEA module is envisioned as the coal car, and is near the region where the majority of the glycosylation occurs in domain I (sugars not shown). Features of note include the “split cars” such as the three laminin EGF-like modules split by a laminin IV type A-like module (red split by blue). Note the one Ig-like module at the end of domain II, with the rest in long domain IV.
Domain I

Perlecan domain I consists of a module that shares homology with the Sperm, Enterokinase and Agrin (SEA) fold of other ECM and cell surface proteins. The function that perlecan domain I plays in vertebrate tissue integrity and development cannot be overstated. A large part of this function is owed to perlecan’s GAG chains, typically a mixture of heparan sulfate and chondroitin sulfate chains usually with a majority of heparan sulfate, attached at three sites that immediately precede the SEA fold [Murdoch et al., 1992]. Domain I is the only domain of perlecan unique to the molecule – modules of domains II through V have high sequence homology to modules present in other ECM proteins [Murdoch et al., 1992]. These three GAG attachment sites share the consensus site S-G-D in mammalian perlecan. In Gallus gallus perlecan the third site is conserved, while the first two are modified to S-A-D and S-G-E, respectively. In Danio rerio, the SEA fold of domain I is conserved, but there exists no homology of the GAG attachment sites observed in tetrapods. Domain I is not present in D. melanogaster or C. elegans perlecan and appears to have been linked to the perlecan core protein later in evolution [Hummel et al., 2004].

Domain II

Human perlecan domain II consists of four repeats with homology to low-density lipoprotein (LDL) receptor and one folding module with homology to the immunoglobulin (Ig) V-set domain. This is the only isolated Ig module in the perlecan core protein. Each of these folding modules shares greater than 80% pair-wise identity with other mammalian perlecan proteins. Perlecan domain II in Gallus gallus and Danio rerio is structured similarly to mammalian domain II, and their folding modules share 50 to 75 percent homology with human perlecan. In D. melanogaster perlecan, or trol, domain II diverges
from the vertebrate molecule by containing twenty one LDL-receptor domains, and one additional I-set Ig repeat. The final two hundred and fifteen amino acids of *D. melanogaster* perlecan domain II share 43% pair-wise identity with human perlecan and the folding modules in this region are spaced very similarly to those in human perlecan. *C. elegans* perlecan, or Unc-52, domain II, is similar to that of *D. melanogaster*, contains two Ig folds, but in contrast to *D. melanogaster* and similar to vertebrate perlecan, contains only three LDL-receptor-like folds. The Unc-52 type III mutant occurs via transposon element insertion into the first Ig repeat of domain II, resulting in creation of a premature stop codon [Rogalski et al., 1995].

**Domain III**

Human perlecan domain III encompasses an interesting stretch of the protein, containing three laminin B domains that break up a longer stretch of eleven laminin EGF-like repeats. An interesting structural feature of the laminin B domains is that their insertions split three of the laminin EGF-like repeats, hinting that they were inserted sometime during evolution into what was likely a contiguous stretch of EGF-repeats. Also note that one of the laminin EGF-like repeats is shortened (the 4th). All vertebrates including *Danio rerio* share this splitting of the laminin EGF-like folding modules, and it appears to be the same in *D. melanogaster* perlecan. In *C. elegans* perlecan, these “split” laminin EGF-like modules also exist, but as described below, *C. elegans* domain III is shortened by one third. The similarities between human domain III and that of other vertebrates mirror that of domain II; domain III in mammals shares roughly 90% protein identity whereas *Gallus gallus* and *Danio rerio* share roughly 60% protein sequence identity. Overall domain III protein identity with invertebrates is lower, 32% in *D. melanogaster* and 36.8% in *C. elegans*. 
While protein sequence identity within folding modules can be as high as 50%, the *D. melanogaster* domain III appears to encode one less laminin-EGF repeat than human perlecans at the C-terminal, and *C. elegans* domain III is absent the 200 amino acid C-terminal stretch containing a laminin B module and several laminin EGF-like repeats. This variability no doubt provides clues to the complex evolution of this domain over the hundreds of millions of years that perlecans has existed in the metazoan genome.

Interestingly, in Unc-52, the homologue of perlecans in *C. elegans*, a type II mutation (st549) occurs in the second laminin B repeat of domain III – resulting in a premature stop codon. This mutation creates a terminal paralyzed arrest at embryonic two cell stage (PAT) phenotype characterized by immobility and arrested elongation at the two-fold stage of embryonic development [Rogalski et al., 1995], a phenotype attributed to incomplete development of body-wall muscle cells.

**Domain IV**

Human perlecans domain IV is the most homogeneous region of perlecans, and consists of twenty one tandem Ig domain repeats, numbered as repeats 2-22 in human perlecans. Many mammalian perlecans encode 21 Ig modules in perlecans domain IV, but mouse perlecans includes only fourteen repeats [Hopf et al., 1999], a result of splicing. A mouse splice variant form that added human Ig module-equi­valents 8-10 has been reported [Noonan and Hassell, 1993]. The number and composition of the Ig repeats in domain IV also varies throughout evolution. *Gallus gallus* perlecans contains eighteen Ig repeats, *Danio rerio* perlecans contains twelve, *C. elegans* perlecans contains fourteen, and *D. melanogaster* contains only ten. Protein identity for domain IV amongst mammalian species is near 90%, but in *Gallus gallus* protein identity is reduced to close to 60%, and
among Danio rerio and invertebrate perlecan there is less than 20% protein amino acid sequence identity when compared to human perlecan. All protein alignments and analytics were performed using the sequence analysis software package Geneious (Auckland, New Zealand) and GenBank reference sequences.

A colorized analysis of the individual sequences of the modules contained within the disulfide bonded regions of each the Ig repeats of human domain IV is shown in figure 1.3. The two bonded cysteines are shown in green. All repeats contain a common tyrosine two amino acids from the C-terminal cysteine, shown in pink, and a common invariant tryptophan about a third of the way into the sequence shown in black. Proline (yellow) also is a common feature, with minor variation seen. From visual inspection, it is immediately apparent that two patterns appear: repeats 2-5 and 16-22 characterized by a more variable pattern for the Igs nearer the domain termini, and a tighter and more conserved pattern within the central repeats 6-15. All of these central repeats contain a doublet of basic amino acids (dark blue) followed by a glycine couplet (gold). It is intriguing to speculate that two common ancestral Ig folding patterns account for this structure. The blocked sequence in repeat 21 is the sequence of the adhesive peptide reported previously [Farach-Carson et al., 2008].

Unc-52 mutations e1421, e1012, e669, e444, e998, st196 occur in domain IV [Rogalski et al., 1995], in exons that can be spliced out of alternative transcripts, rescuing the worm phenotype. This suggests that mutations in the Ig repeat domains can act as dominant negatives, and that having fewer repeats and a shorter domain IV is better than having misfolded proteins with regard to tissue structure and function. It is intriguing to speculate
that the length of the Ig repeat-rich region thus plays another function, such as specification of tissue stiffness or elasticity at the tissue border.
Figure 1.3 Colorized analysis of Ig-modules in domain IV showing two clear patterns. The amino acid sequences from the human perlecan core protein that lie in between the cysteines involved in disulfide bond formation of the Ig-modules in domain IV (green) were extracted, aligned, and colorized by polarity of amino acid (green – polar, gold – nonpolar, blue – basic, red – acidic). Highly conserved residues are given unique coloring (tryptophan – black, proline – yellow, tyrosine – pink). The terminal module sets (2-5) and (16-22) are similar to one another, and are split by (6-15) showing a distinct pattern (see text). The consensus sequence represents the modal amino acid at that position. Mean hydrophobicity is a representation of the overall hydrophobic nature of the amino acids at that position in all 21 repeats (higher, redder bars represent more hydrophobic residues at that position). The sequence logo readout indicates what amino acids are present at that position, and in what proportions. Finally the height of the bars in the sequence identity graph indicates the degree of identity at the position in these 21 sequences. From Farach-Carson et al. 2014.
Domain V

Human perlecan domain V contains three folding modules with homology to the laminin G chain, separated by two modules that share homology to EGF-like protein domains. Protein identity with vertebrate orthologues mirrors that of domain III, but in D. melanogaster and C. elegans, protein identity of folding modules is between fifteen and thirty-five percent. The C-terminal EGF-like fold is missing from C. elegans perlecan.

In mammals, the functions of this portion of the perlecan protein are associated with modulation of blood vessel growth and structural integrity [Mongiat et al., 2003b]. Ongoing work in several labs focuses on dissecting the complex structural functional relationships of the motifs in domain V that provide angiogenic control. Like a border patrol agent, endorepellin, one of the active fragments created by proteolytic cleavage of perlecan domain V [Gonzalez et al., 2005], was reported to inhibit four aspects of angiogenesis: 1) endothelial cell migration; 2) collagen-induced endothelial tube morphogenesis; 3) blood vessel growth in the chorioallantoic membrane and; 4) angiogenesis in Matrigel™ plug assays [Mongiat et al., 2003b]. Further work by this group showed that endorepellin requires both the α2β1 integrin and VEGFR2 to demonstrate its angiostatic activity [Goyal et al., 2011]. In brain, perlecan domain V (DV) promotes brain angiogenesis by triggering release of VEGF from brain endothelial cells as shown in an in vivo stroke model [Clarke et al., 2012]. Interestingly, in this same study, the LG3 portion of DV, which is reported to possess the majority of DV’s “angio-modulatory” activity outside of brain, binds poorly to α5β1 and induces less endothelial cell proliferation than does complete domain V [Clarke et al., 2012]. This combined work provides experimental strength to the notion that perlecan can act as a context driven “switch” that can, by virtue
of local proteolytic processing, patrol borders where endothelial growth and tissue penetration is under strict regulation.

**Perlecan function in normal tissues and the effect of HSPG2 loss**

The perlecan protein is key in maintenance of normal tissue function [Farach-Carson et al., 2013]. The protein is found in the BM of many tissues, providing protein scaffolding and signaling regulation in the dividing line between epithelium and stroma, or endothelium and the surrounding tissue. A number of cell types express perlecan in their pericellular matrix, including chondrocytes, bone marrow stromal cells and muscle. The loss or mutation of the *HSPG2* gene can be fatal or devastating to human health depending on the nature of the genetic aberration and the extent of haploinsufficiency.

**Perlecan in the basement membrane**

Perlecan binds directly to the other three major constituents of the BM [Murdoch et al., 1994], collagen IV, laminin, and nidogen/entactin (see table 1.1). The protein was first recognized in the secretome of the BM-synthesizing Engelbroth-Holm-Swarm (EHS) tumor cell line [Hassell et al., 1980], and was originally referred to as “large BM heparan sulfate proteoglycan” [Laurie et al., 1988] or variants thereof [Anderson and Fambrough, 1983] before its classification as *Heparan sulfate proteoglycan 2* [Dodge et al., 1991] or perlecan [Noonan et al., 1991]. The protein derives its name from its appearance upon rotary shadowing electron microscopy as beads on a string [Yurchenco et al., 1987]. Perlecan is the major proteoglycan component of BMs [Hassell et al., 1980; Paulsson et al., 1986] and as such acts as the reservoir for heparin-binding growth factors (see table 1.2), limiting their diffusion to act on cells on either side of the BM. Only through
degradation of the protein or its associated GAG chains can these signaling factors fully act on downstream cells [Whitelock et al., 1996], in some circumstances through a ternary complex requiring the heparan sulfate, signaling molecule and receptor protein [Chuang et al., 2010]. Perlecan’s degradation during wound healing or tissue remodeling may act to quickly introduce necessary mitogens and trophic factors to begin regenerative processes. In this model, perlecan acts as a quickly activated depot for latent signaling activities and adhesive motifs.

The BM maintains the polarity of epithelia [Schneider et al., 2006; Yu et al., 2005], and the various components of the BM are scaffolded by perlecan [Behrens et al.] in concert with nidogen [Fox et al., 1991]. In the prostate, BM proteins are expressed by the basal cell type of the epithelium [Gandellini et al., 2012]. This lends special importance to the role of basal cells in development and maintenance of the normal epithelium in the prostate. The general hypothesis is that these cells are responsible for depositing perlecan as well as the other BM factors in the prostate. Other tissues may have different expression patterns. Perlecan is expressed by parabasal cells in stratified oral epithelia [Ikarashi et al., 2004], and perlecan epitopes in the BM are only exposed after digestion with proteases. This study showed an increased and unregulated expression of perlecan during progression of oral squamous cell carcinoma. In a transgenic mouse model of perlecan knockout, the corneal epithelium is significantly thinner as a result of retarded growth of the epithelial progenitor cells. The hypothesis is that perlecan provides important growth signals to progenitor cells to form the mature cornea [Inomata et al., 2012].

Perlecan is found in the BM surrounding the myofibrils of skeletal muscle [Murdoch et al., 1994]. Perlecan binds directly to several proteins important to skeletal muscle function
including alpha-dystroglycan, acetylcholinesterase and myostatin (see table 1.2). The “rafting” or scaffolding function of perlecan is evident in its association with acetylcholine receptor plaques in a *Xenopus* skeletal muscle model of neuromuscular junctions (NMJ) [Anderson and Fambrough, 1983]. This acetylcholine receptor rafting is likely related to the binding of perlecan to acetylcholinesterase [Arikawa-Hirasawa et al., 2002b], which is not found at the NMJ of *Hspg2⁻/⁻* mice, consistent with the myotonic phenotype of perlecan mutants in mice [Stum et al., 2008] and humans [Arikawa-Hirasawa et al., 2002a; Mereu et al., 1969]. Perlecan is responsible for maintenance of skeletal muscle fiber size and composition [Xu et al., 2010] at least in part through promoting myostatin expression and signaling; *Hspg2*-deficient mouse models of perlecan result in significantly increased cross-sectional area of skeletal muscle fibers. Cardiac muscle [Sasse et al., 2008] and cardiovascular development in mice is dependent upon perlecan. In *Hspg2*-null mice, cardiac development is so severely impaired that very few (30%) fetuses survive to birth and all die perinatally.

The endothelial BM separates the endothelium from the surrounding tissue [Leung et al., 2000]. Endothelial cell interaction with pericytes in co-culture stimulates the cells to produce a mature BM and thus conclude vasculogenesis [Stratman et al., 2009]. Paracrine signals between pericytes and endothelial cells in mature vessels are presumably mediated by perlecan and its heparan sulfate chains. Aortic endothelial cells adhere to the perlecan core protein in culture through an integrin-mediated interaction modified by heparan sulfate [Hayashi et al., 1992] and synthesize perlecan and biglycan in response to TGFβ signaling [Kaji et al., 2000]. Perlecan aids in maintaining intact and functional [Tran et al., 2004] blood vessels by involvement in the BM, but also inhibits the formation of new blood
vessels when fragmented. The C-terminus of perlecan is sometimes known as endorepellin for its ability to repel endothelial cells and prevent angiogenesis [Mongiat et al., 2003b]. This activity is at odds with pro-angiogenic action of the intact molecule driven by heparan sulfate-bound fibroblast growth factor 2 signaling [Aviezer et al., 1994]. Thus endorepellin functions differently than the intact molecule after digestion of perlecan by proteases [Gonzalez et al., 2005]. The potent angiostatic activity of endorepellin shows promise in the treatment of ischemic stroke [Lee et al., 2011]. The blood-brain barrier depends on the integrity of BM surrounding capillaries, of which perlecan is a part [Deguchi et al., 2002].
### Table 1.1

Perlecan binding factors involved in tissue structure and function

<table>
<thead>
<tr>
<th>Binding Partner</th>
<th>Binding domain involved</th>
<th>Affinity (Kd&lt;sub&gt;app&lt;/sub&gt;)</th>
<th>Bioactivity</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>collagen type V alpha 4 (rat)</td>
<td>HS</td>
<td>unreported</td>
<td>Schwann cell binding</td>
<td>[Rothblum et al., 2004]</td>
</tr>
<tr>
<td>elastin</td>
<td>HS and protein core</td>
<td>unreported</td>
<td>elastic fiber assembly</td>
<td>[Hayes et al., 2011]</td>
</tr>
<tr>
<td>laminin</td>
<td>HS</td>
<td>5 nM</td>
<td>BM structure and integrity</td>
<td>[Ettner et al., 1998]</td>
</tr>
<tr>
<td>nephronecrtin (mouse)</td>
<td>HS</td>
<td>unreported</td>
<td>elastic fiber assembly</td>
<td>[Sato et al., 2013]</td>
</tr>
<tr>
<td>PRELP (mouse)</td>
<td>HS</td>
<td>3 nM</td>
<td>connective tissue/BM integration</td>
<td>[Bengtsson et al., 2002]</td>
</tr>
<tr>
<td>fibrillin-1</td>
<td>II</td>
<td>5-10 nM</td>
<td>connective tissue microfibril organization</td>
<td>[Tiedemann et al., 2005]</td>
</tr>
<tr>
<td>WARP</td>
<td>III - second Laminin-EGF repeat</td>
<td>23 nM</td>
<td>chondrocyte pericellular matrix</td>
<td>[Allen et al., 2006]</td>
</tr>
<tr>
<td>fibronectin</td>
<td>unknown</td>
<td>2 nM</td>
<td>pericellular matrix assembly</td>
<td>[Heremans et al., 1990]</td>
</tr>
<tr>
<td>collagen IV</td>
<td>IV</td>
<td>233 nM</td>
<td>basement membrane integrity</td>
<td>[Hopf et al., 1999; Whitelock et al., 1999]</td>
</tr>
<tr>
<td>nidogen-1</td>
<td>IV-1</td>
<td>10 nM</td>
<td>structural integrity of BM</td>
<td>[Hopf et al., 1999]</td>
</tr>
<tr>
<td>nidogen-2 (human)</td>
<td>IV-1 (mouse)</td>
<td>10 nM</td>
<td>structural integrity of BM</td>
<td>[Hopf et al., 1999]</td>
</tr>
<tr>
<td>alpha-dystroglycan</td>
<td>V</td>
<td>3 nM</td>
<td>epithelial cell polarity</td>
<td>[Talts et al., 1999]</td>
</tr>
<tr>
<td>alpha-2 – beta-1 integrin</td>
<td>V (mouse)</td>
<td>80 nM</td>
<td>anti-angiogenic signaling</td>
<td>[Bix et al., 2004]</td>
</tr>
<tr>
<td>alpha-5 - beta-1 integrin (mouse)</td>
<td>V</td>
<td>160 nM</td>
<td>cell adhesion, anti-angiogenic signaling</td>
<td>[Lee et al., 2011; Nystrom et al., 2009]</td>
</tr>
<tr>
<td>ECM1</td>
<td>V</td>
<td>unreported</td>
<td>basement membrane organization</td>
<td>[Mongiat et al., 2003a]</td>
</tr>
<tr>
<td>fibulin-2 (mouse)</td>
<td>V</td>
<td>196 nM</td>
<td>BM regulation</td>
<td>[Brown et al., 1997]</td>
</tr>
</tbody>
</table>
Table 1.2

Bound active molecules that can be released by perlecan turnover

<table>
<thead>
<tr>
<th>Binding Partner</th>
<th>Binding Domain</th>
<th>Affinity</th>
<th>Bioactivity</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>acetylcholinesterase</td>
<td>HS (avian)</td>
<td>unreported</td>
<td>NMJ function</td>
<td>[Peng et al., 1999]</td>
</tr>
<tr>
<td>activin A</td>
<td>HS</td>
<td>unreported</td>
<td>tissue growth and repair</td>
<td>[Li et al., 2010]</td>
</tr>
<tr>
<td>BMP-1</td>
<td>HS</td>
<td>573 pM</td>
<td>osteogenesis</td>
<td>[Decarlo et al., 2012]</td>
</tr>
<tr>
<td>BMP-2</td>
<td>HS</td>
<td>600 pM</td>
<td>osteogenesis/ chondrogenesis</td>
<td>[Decarlo et al., 2012]</td>
</tr>
<tr>
<td>FGF-1</td>
<td>HS</td>
<td>unreported</td>
<td>tissue growth and repair</td>
<td>[Melrose et al., 2006]</td>
</tr>
<tr>
<td>FGF-2</td>
<td>HS</td>
<td>0.6-2 nM; 12 nM</td>
<td>tissue growth and repair</td>
<td>[Chu et al., 2005; Knox et al., 2002]</td>
</tr>
<tr>
<td>FGF-9</td>
<td>HS</td>
<td>unreported</td>
<td>tissue growth and repair</td>
<td>[Melrose et al., 2006]</td>
</tr>
<tr>
<td>HB-EGF</td>
<td>HS</td>
<td>44 nM; 6.1 nM</td>
<td>tissue growth and repair</td>
<td>[Chu et al., 2005]</td>
</tr>
<tr>
<td>histone H1</td>
<td>HS</td>
<td>unreported</td>
<td>tissue growth and repair</td>
<td>[Henriquez et al., 2002]</td>
</tr>
<tr>
<td>IL-2 (mouse)</td>
<td>HS</td>
<td>unreported</td>
<td>inflammation; immune surveillance</td>
<td>[Miller et al., 2008]</td>
</tr>
<tr>
<td>LTBP-2</td>
<td>HS</td>
<td>unreported</td>
<td>tissue remodeling and repair</td>
<td>[Parsi et al., 2010]</td>
</tr>
<tr>
<td>Thrombospondin-1</td>
<td>HS</td>
<td>unreported</td>
<td>tissue development and growth</td>
<td>[Vischer et al., 1997]</td>
</tr>
<tr>
<td>VEGF165</td>
<td>HS</td>
<td>“high”</td>
<td>tissue growth /angiogenesis</td>
<td>[Zoeller et al., 2009]</td>
</tr>
<tr>
<td>WNTs/βWg</td>
<td>HS</td>
<td>0.01-0.1 nM</td>
<td>tissue development and growth</td>
<td>[Lin, 2004; Reichsman et al., 1996]</td>
</tr>
<tr>
<td>Hedgehogs (Hhs)*</td>
<td>HS/Core Protein</td>
<td>“high”</td>
<td>tissue development and growth</td>
<td>[Datta et al., 2006; Lin, 2004; Park et al., 2003]</td>
</tr>
<tr>
<td>TGF beta</td>
<td>HS/Core Protein</td>
<td>low nM</td>
<td>tissue growth and repair</td>
<td>[Cheifetz and Massague, 1989; Harrison et al., 2011]</td>
</tr>
<tr>
<td>LDL &amp; VLDL</td>
<td>II</td>
<td>“high”</td>
<td>cholesterol and lipid delivery</td>
<td>[Hummel et al., 2004]</td>
</tr>
<tr>
<td>FGF-7</td>
<td>III</td>
<td>60 nM</td>
<td>tissue growth and repair</td>
<td>[Mongiat et al., 2000]</td>
</tr>
<tr>
<td>FGF-18</td>
<td>III</td>
<td>27.8 nM</td>
<td>tissue growth/chondrogenesis</td>
<td>[Smith et al., 2007]</td>
</tr>
<tr>
<td>FGF-bp</td>
<td>III - second laminin-EGF repeat</td>
<td>18 nM</td>
<td>release of FGF</td>
<td>[Mongiat et al., 2001]</td>
</tr>
<tr>
<td>PDGF</td>
<td>III - second laminin-EGF repeat</td>
<td>8-25 nM</td>
<td>tissue growth/angiogenesis</td>
<td>[Gohring et al., 1998]</td>
</tr>
<tr>
<td>FGF-7</td>
<td>V</td>
<td>60 nM</td>
<td>tissue growth and repair</td>
<td>[Mongiat et al., 2000]</td>
</tr>
<tr>
<td>myostatin (mouse)</td>
<td>V (mouse)</td>
<td>11 nM</td>
<td>regulation of muscle mass</td>
<td>[Sengle et al., 2011]</td>
</tr>
<tr>
<td>Progranulin</td>
<td>V</td>
<td>1 µM</td>
<td>blood vessel growth</td>
<td>[Gonzalez et al., 2003]</td>
</tr>
<tr>
<td>VEGFR1</td>
<td>V</td>
<td>10.6 nM</td>
<td>VEGF signaling; angiogenesis</td>
<td>[Goyal et al., 2011]</td>
</tr>
<tr>
<td>VEGFR2</td>
<td>V</td>
<td>2.8 nM</td>
<td>VEGF signaling; angiogenesis</td>
<td>[Goyal et al., 2011]</td>
</tr>
</tbody>
</table>

* Extrapulated from genetic and diffusion data.
Perlecan in pericellular matrices

Perlecan is part of the pericellular matrix of chondrocytes [Kvist et al., 2008; Melrose et al., 2002; West et al., 2006]. The organization of cartilage is such that a BM would serve no purpose in the traditional sense of separating tissue layers, but perlecan is observed in the local matrix of developing chondrocytes and is thought to form a chondron “BM” of sorts. Articular cartilage demonstrates perlecan antibody reactivity in fetal human elbow and knee [Melrose et al., 2006]. Perlecan derived from cartilage is often substituted with chondroitin sulfate in place of heparan sulfate, lending the molecule decreased binding capacity for basic FGF and other heparan-binding growth factors [Melrose et al., 2006]. Chondroitin sulfate-substituted perlecan demonstrably facilitates collagen fibril assembly in vitro, indicative of the molecule’s importance to cartilage development. Primordial cartilage express the perlecan transcript in E15.5 mouse embryos [French et al., 1999], indicating that perlecan is an early but not initial part of the chondrogenic pathway, as collagen type II transcripts were detected as early as E12.5. Perlecan is found in pericellularly in growth plate cartilage in long bones [Melrose et al., 2006], but at the chondro-osseous junction is abruptly lost [Brown et al., 2008; Melrose et al., 2002]. This stark border between chondrocytes and trabecular bone implicates perlecan in the maintenance of the cartilage phenotype and inhibition of calcification. Proteoglycans are not found in the collagen I-rich matrix of mineralized bone [Chen and Boskey, 1985], but are found in the lacuna canalicular system, presumably supporting the function of osteocytes in regulating bone remodeling and homeostasis [Thompson et al., 2011]. Perlecan is the most abundant proteoglycan of the hematopoietic bone marrow [Schofield et al., 1999], and is expressed by bone marrow stromal cells.
Disorders associated with *HSPG2* mutation

The mutation or loss of *HSPG2* can have dramatic effects on human and animal health. *HSPG2* functional null humans and mice are not viable [Arikawa-Hirasawa et al., 2001], and die before childbirth or perinatally of a condition referred to as Dyssegmental Dysplasia, Silverman Handmaker type (DDSH)[Handmaker et al., 1977]. This is characterized by incomplete development of bone and cardiac muscle, resultant in anisospondyly, micromelia (or dwarfism), encephalocele, flat face and cleft palate. Less severe mutations result in a condition known as Schwartz-Jampel Syndrome (SJS) [Arikawa-Hirasawa et al., 2002a; Nicole et al., 2000]. SJS involves a spectrum of symptoms from mild to severe, but patients survive [Mereu et al., 1969; Schwartz and Jampel, 1962]. General myopathy including myotonia, chondrodysplasia, and blepharophimosis (droopy eyelids) are common features of SJS. Initially thought to be a recessive disorder [Fontaine et al., 1996], SJS patients have been identified carrying heterozygous *HSPG2* alleles [Arikawa-Hirasawa et al., 2002a], with mutations presumably acting in a dominant negative fashion. In mice, perlecan knockouts rarely survive to birth and survivors die perinatally. Embryos commonly suffer heart failure at E10.5 from cardiac muscle abnormalities resulting in hemopericardium [Costell et al., 1999; Sasse et al., 2008]. These cardiac aberrations are associated with incomplete BM formation in the cardiac muscle [Sasse et al., 2008]. Transposition of the great arteries is common in mouse embryos that survive past E10.5 [Costell et al., 2002], related to misregulated neural crest cell migration during development. These embryos develop exencephaly [Costell et al., 1999] and skeletal abnormalities reminiscent of DDSH. There exist *HSPG2* hypomorph mice, and mice with tissue-specific perlecan knockouts [Xu et al., 2010] or perlecan
mutations [Rossi et al., 2003] whose phenotype are less severe allowing the study of perlecan-deficient tissue. Zebrafish perlecan morphants have myopathic and cardiovascular irregularities [Zoeller et al., 2008]. Although the heart of these animals does contract, circulation is limited owing to incomplete vascular development. Muscular development and deficiency in movement of perlecan morphant zebrafish echoes the phenotype of the Unc-52 C. elegans perlecan mutants, discussed in the following section.

**Perlecan function in invertebrates gives clues to its evolutionary origin**

*u*nc-52

The “uncoordinated” or *Unc* mutants were first identified by Brenner in 1974 [Brenner, 1974], but *unc*-52 was not attributed to a mutation in the *C. elegans* perlecan gene until 1993 [Rogalski et al., 1993]. Many *unc*-52 mutants create new potential splice sites, or create frameshift mutations [Merz et al., 2003]. Loss of *unc*-52 results in a PAT phenotype similar to that seen in the loss of nematode collagen IV genes, characterized by loss of embryonic extension concomitant with body wall muscle dysfunction [Gilchrist and Moerman, 1992; Gupta et al., 1997]. Loss of functional Unc-52 protein is thought to disrupt formation of body wall muscle cell dense bodies, an analogue to focal adhesions [Waterston et al., 1980]. These dense bodies detach from the plasma membrane with age or if pressure is applied to the animal [Waterston, 1988]. In certain *unc*-52 mutants, electron microscopy reveals that as the animals age, foreign cell types invade the muscle cell layer, which should be contiguous [Waterston et al., 1980]. This is not unlike what is seen in the developing heart in the absence of perlecan [Costell et al., 1999; Sasse et al., 2008].
Additional effects of viable *unc-52* mutations include defects in gonad organogenesis. These defects can be suppressed independently of the body-wall muscle defects, indicating that these functions of *unc-52* are separate [Gilchrist and Moerman, 1992]. Loss of *unc-52* regulation of growth factor signaling may be responsible for gonad dysgenesis in mutants [Merz et al., 2003]. The Unc-52 M isoform is crucial for maintenance of wild-type gonadogenesis, loss of this specific isoform leads to aberrant growth factor signaling [Merz et al., 2003]. The barrier function of Unc-52 is apparent in the gonadal sheath rupture phenotype of certain mutant strains. This phenotype includes destruction of the BM surrounding the gonad, sometimes accompanied by “disappearance” of the sheath cells and leaking of oocytes into the body cavity [Gilchrist and Moerman, 1992]. It is unclear what drives this effect on sheath cell integrity, but it has been speculated that it is mediated by a loss of adhesion signals to the cells from the Unc-52-depleted BM. *Unc-52* mutation or depletion has been tied to developmental defects in the pharynx [Jafari et al.; Morck et al.].

*trol*

*D. melanogaster* perlecan is referred to as *terribly reduced optic lobes* or *trol*. Mapped to the X chromosome, *trol* was identified in a genetic screen for mutations affecting imaginal central nervous system (CNS) development [Datta and Kankel, 1992]. P-element insertion and chemical mutagenesis were used to create the mutant *D. melanogaster* strains in this study. Certain mutant alleles of *trol* result in vacuolation of the optic lobe neurons and misguided development of axons in the optic lobe, which occur to varying degrees of severity. These phenotypes are tied to a loss of proliferation of the precursor optic lobe neuroblast cells during the transition from second to third instar larvae, and to a decrease
in overall CNS cell proliferation. The functional consequence of this mutant is decreased performance in certain behavioral experiments that were created to test vision [Datta and Kankel, 1992]. One allele of *trol* affects the integrity of the neuropil in the ventral ganglion of third instar larvae: neuropil in these animals is nearly overtaken by neuron cell bodies [Datta and Kankel, 1992]. Once again, this is an example of compromised function of a critical barrier that allows cells to move into regions from which they are normally excluded.

Trol, in accord with the signaling role of perlecan in vertebrates, regulates signaling and diffusion of signals by the FGF and hedgehog molecules [Park et al., 2003]. This interaction is not clearly understood, as in vertebrates it relies upon heparan sulfate binding to heparin-binding growth factors, and trol protein includes neither GAG attachment sites nor a homologue to the vertebrate domain I. Signaling transforming growth factor beta and WNT, or wingless, in the *D. melanogaster* brain also depends upon trol expression [Lindner et al., 2007].

Trol has some roles in tissue development that seem to be related to cell adhesion and tissue polarity. There are three Arg-Gly-Asp (RGD) sequences in trol, and some evidence that they bind to the *D. melanogaster* integrin βV in larval hemocytes. This interaction has been suggested to mediate the phagocytosis of apoptotic cells [Nagaosa et al.]. Trol and dystroglycan partner to maintain follicle-cell epithelium polarity through BM interactions when flies are under energetic stress. In trol-negative systems this barrier is no longer functional, and dystroglycan is no longer sequestered in the basal membrane, in parallel with apical cell markers migrating to the basal membrane [Schneider et al., 2006]. This may contribute to reproductive dysfunction in these fly genotypes. Studies of matrix
metalloproteinase (MMP) function in developing flies revealed a function for MMP2 in spatially and temporally dynamic regulation of perlecan and collagen IV deposition in the BM. This barrier, when left intact by MMP2 knockout, prohibited the proper formation of the *D. melanogaster* air sac primordium, and was characterized by accumulation of collagen IV and perlecan [Guha et al., 2009]. BM-associated trol also defines organ shape in developing wing disc and ventral nerve cord [Pastor-Pareja and Xu]. In trol-knockout larvae, these organs have a narrow and constricted shape in comparison with wild-type organs. This provides yet another example of a key role for perlecan in setting cell boundaries during complex tissue morphogenesis.

**Evolutionary clues about perlecan and BM assembly**

The evolutionary origin of the perlecan gene remains speculative, and a multicellular animal genome that does not encode a perlecan-like gene has yet to be discovered. A telling insight into the evolution of the perlecan gene will be provided when an early genome that does not encode a perlecan gene is found. *C. elegans* Unc-52 is the oldest well studied perlecan gene, and *D. melanogaster* trol is the second-oldest well studied form of perlecan. Clues to the evolutionary organization of perlecan are found by comparing these ancient genes to the more modern perlecan genes of vertebrates, as described above in section 2. It is interesting to speculate that the inclusion of 18-20 more LDL receptor-like folding modules in *trol* represents a remnant of the ancestral form of domain II, and these additional LDL receptor-like modules were lost in vertebrate perlecan, but the absence of these additional repeats in the more simple *C. elegans* makes one wonder if this is so. Interesting too is the presence in Unc-52 of an additional Ig module in domain II, which is present in flies, but not vertebrates. It is unclear whether these differences occur
due to duplication of these LDL receptor-like modules in flies or negative selective pressure in the other clades. Also of note is the lack of O-glycosylation sites in domain I of zebrafish, *D. melanogaster*, and *C. elegans* perlecan. There are three O-glycosylation sites in domain I of chicken and mammalian perlecan molecules, at positions where heparan sulfate or chondroitin sulfate typically are added. The evolutionary pressure that led to these complex adaptations is unclear.

Of interest, the RPL21P29 pseudogene resides within the 40,000 base pair intron one of perlecan in all apes, but not in other mammals or any other species. This pseudogene is found 145 times in the human genome [Zhang et al., 2002]. Given that it is conserved in the perlecan gene only in apes, it seems to be a recent addition to this locus. My studies find variation among the numbers of Ig modules in domain IV (figure 1.3), suggesting that insertions or duplications of certain Ig modules in domain IV occurred commonly throughout evolution. Although the central Ig modules are more conserved relative to one other in human perlecan (Figure 1.3), they are not conserved in evolution (figure 1.1B), persisting in mouse perlecan, but not entirely present in chicken perlecan, and even more scarcely present in zebrafish perlecan. The reasons for this variation remain unknown, but the nature of perlecan molecules in more ancient animals is explored further in chapter 4.

**Prostate cancer biology**

**Epidemiology**

One in six American men will be diagnosed with prostate cancer, and one in thirty will die from the disease. It is the most commonly diagnosed cancer in American men as of 2013 [Siegel et al., 2013]. It is also the second most lethal cancer, behind only lung cancer. This
is reflective of the relative cure rate of locally restricted prostate cancer as compared with that of lung cancer. As with most cancers, the rate of incidence increases with age [Whittemore et al., 1991]. The increasing incidence with age is so great that some consider the disease inevitable – that the question is not whether you will develop prostatic neoplasia, but how malignant it will be [Zlotta et al., 2013]. This is related to the phenomenon of latent prostate cancer – which exists in asymptomatic men who die of other causes and is only found upon autopsy. Although lethal prostate cancer is highest among men in western countries, latent prostate cancer is a worldwide phenomenon [Zare-Mirzaie et al., 2012]. Ethnicity is a risk factor for prostate cancer: African-American men develop cancer 1.6 times more frequently than Caucasian men [Siegel et al., 2013]. The disease is much more common in the west than in Asia, leading some to hypothesize that the western diet predisposes men to the development of the disease, and there is some data supporting the role of hyperinsulinemia in tumor growth promotion in xenografts [Venkateswaran et al., 2007]. The prostate cancer incidence rate amongst Asian men who have moved to the United States is much higher than that of their counterparts in their home countries [Cook et al., 1999; Lee et al., 2007]. Variable results for association of obesity smoking and other risk factors with prostate cancer incidence have been reported but the association between prostate cancer mortality and obesity is strong [Calle et al., 2003]. In the late 1900s, increased surgical treatment of prostatitis and the advent of the PSA screen [Catalona et al., 1991; Wang and Kawaguchi, 1986] contributed to an increase in the perceived rate of prostate cancer incidence. The increase in prostate cancer incidence in the late 1900s is most likely due to increased screening and improved detection methods [Welch and Albertsen, 2009]. Although increased screening has elevated the number of diagnoses and
inarguably saved lives, a high percentage (23-42%) of cancers detected by PSA screening would never have produced symptoms during the patient’s lifetime [Draisma et al., 2009]. This has caused some concern about overdiagnosis of prostate cancer and unnecessary treatment leading to quality of life concerns.

Disease progression at diagnosis can be defined as local, regional or distant. Local disease is confined to the prostate, regional disease involves invasion through the prostate capsule or lymph node involvement, and distant disease involves metastasis of other tissues. When detected early, such that cancer remains local, prostate cancer is curable – radical prostatectomy [Han et al., 2001] or brachytherapy [Potters et al., 2005] can completely ablate the organ. The 5-year survival rates after acute treatment of local disease are near 100% [Brawley, 2012]. The quality of life concerns associated with prostatectomy are generally considered insignificant when compared to this survival rate [Miwa et al., 2013]. Between 1999 and 2006, diagnoses of regional disease and distant disease accounted for 12% and 4% of all diagnoses respectively. Once the cancer metastasizes to the lymph nodes and then most frequently the bone marrow, treatment efficacy drops significantly. The 5-year survival rate of patients with disseminated disease is below 30% [Pond et al., 2013]. This is indicative of the dearth of effective treatments for metastatic disease.

**Primary disease**

The normal prostate is comprised of three zones: the peripheral, central and transitional. Prostate cancer is most commonly derived from peripheral zone cells. The organ also can be described in terms of lobes, with posterior median and anterior lobes corresponding to the peripheral, central and transitional zones, respectively [McNeal, 1981]. The normal function of the gland is to secrete the fluid comprising a majority of the semen. This
function is supported structurally by a ductal system feeding towards the ejaculatory ducts. Normally, ductal epithelium is comprised of highly regular cell layers of luminal epithelial cells derived from basal epithelial cells [Schalken and van Leenders, 2003]. Certain basal cells have stem cell-like properties of replenishing the luminal epithelial cell layer [Garcia et al., 2007], and some believe they are the prostate cancer progenitor cells [Lawson et al., 2010] although this is somewhat contested [Choi et al., 2012; Okada et al., 1992; Parsons et al., 2001]. The relative lifetimes of basal stem cells (infinite) and luminal epithelial cells (limited) intuitively suggest that these cells are more likely to accrue genetic aberrations leading to hyperplasia. A relatively rare type of cell found in the luminal epithelial cell layer is the neuroendocrine cell, which secretes various signaling molecules to affect the surrounding cells in a paracrine manner [Bonkhoff et al., 1994].

Prostatic intraepithelial neoplasm

Prostate cancer is related to a pre-malignant neoplasm known as prostatic intraepithelial neoplasm (PIN), originally identified as intraductal dysplasia [McNeal and Bostwick, 1986]. High-grade (HG)PIN is believed to be the precursor to bona fide prostate cancer in many cases [Eminaga et al., 2013]. The morphological definition of PIN shares characteristics with invasive prostatic carcinoma. In the luminal epithelium there occurs an increase in nuclear size and nucleolar number, atypical nuclear polarity, cell crowding and multilayering as well as increased chromatin content, making the nuclei darker when stained [McNeal and Bostwick, 1986]. HGPIN has some predictive value for invasive prostate cancer on biopsy, particularly when there are numerous foci of HGPIN [Srigley et al., 2010]. Both primary prostate cancer and PIN are multi-focal in 75 percent of radical prostatectomies. Multifocal lesions may develop from a single clonal source [Boyd et al.,
2012] possibly via intraglandular dissemination, but once established these multiple sites surely develop and adapt independently [Arora et al., 2004]. This ultimately contributes to the heterogeneity of late-stage disease and will confound efforts to develop personalized treatments for disseminated prostate cancer. PIN and prostate cancer are differentiated by loss of the basal cell layer, early stromal invasion [Bostwick et al., 1993], and discontinuity of the BM. In essence, prostate cancer breaks the boundaries of the epithelial cell layer, and HGPIN stays within those boundaries. Stromal and immune cell recruitment are largely absent from PIN lesions, but are integral to prostate cancer progression.

The Gleason scoring system

The Gleason scoring system is used to classify tumors based upon morphology and histological characteristics [Delahunt et al., 2012]. Gleason grades 1 and 2 are defined by well-organized and normal-looking lumens. The later grades 3 4 and 5 are defined by decreasing organization of the ductal epithelium in the lumens, and involve stromal invasion by cancer cells. To obtain the Gleason Score (GS), two independent observers grade a tumor specimen and those grades are added together. For instance, if one observer grades a specimen a 3 and a second observer gives a specimen a grade of 4, the tumor has a GS of 7. The Gleason score is a useful prognostic indicator of long-term survival. In general, a score of 5 or less indicates a highly favorable outcome, and a score of 8-10 predicts a very low (30%) 5-10 year survival rate [Egevad et al., 2002].

Stromal reaction

As indicated by the Gleason scoring system, the involvement of surrounding tissue is associated with tumor progression. The microenvironment has a profound effect on
whether cancer cells proliferate or remain indolent [Zong et al., 2012] and these effects are only now becoming fully appreciated. The primary prostate tumor is thought to co-evolve with other cells they recruit to their microenvironment by paracrine interaction such as stromal fibroblasts and immune cell types [Giannoni et al., 2010]. Fibroblasts incorporated with the tumor may take on a phenotype similar to the myofibroblasts of wound healing, and are then known as reactive stromal cells or cancer-associated fibroblasts [Seemayer et al., 1979]. Complicating the diagnosis of myofibroblastic differentiation in the prostate is the presence of smooth muscle cells in the normal prostatic stroma, precluding the use of alpha smooth muscle actin as an immunohistochemical marker of reactive stroma [Tuxhorn et al., 2001]. Thus reactive stroma in the prostate can be identified by vimentin and fibroblast activation protein, as well as the ECM molecules collagen I and tenascin-C [Tuxhorn et al., 2002]. This protracted wound healing response coupled with the angiogenic and immune component of the tumor has earned cancer a dubious reputation as “the wound that never heals” [Dvorak, 1986; Schafer and Werner, 2008] The reactive stromal phenotype is characterized by cytoskeletal changes and a secretion of a collagen and proteoglycan-rich ECM known as desmoplasia.

Desmoplasia is thought to be an effort by the host-response mechanism to sequester the tumor while immune cells perform their duties of ridding foreign invaders from the tissue. This desmoplastic phenotype echoes that of the normal response to wound healing, but is co-opted by the cancer cells to foster growth of the tumor [Dakhova et al., 2013; Giri et al., 1999]. Reactive stromal cells in turn are prompted to perpetually attempt to heal the wound of the cancer in contrast to their apoptotic response to completed tissue remodeling following wound healing. During the desmoplastic reaction to oncogenesis in the prostate,
the collagenous matrix of the stroma becomes dysregulated and undergoes fundamental structural changes [Morrison et al., 2000]. Increased expression of the extracellular glycoproteins fibronectin and tenascin-C and the proteoglycan versican also are found in the reactive stroma of prostate cancer [Sonmez et al., 1995]. Collagen and fibronectin expression in reactive stroma are reminiscent of the provisional matrix of wound healing [Midwood et al., 2004; Ricciardelli et al., 1998; Xue et al., 1998]. There is a proposed modification to the prostate cancer staging system that includes a judgement of stromal involvement in the tumor. It is thought that there is a bimodal effect of stromal integration on patient prognosis – where there is no recruitment or an excess of recruitment, patient survival is very poor, but in the case of middling stromal involvement, patient survival is slightly better [Ayala et al., 2003]. This may reflect the salutary intent of the stromal reaction; when the body reacts to the tumor it can often contain it, but if the struggle goes on for too long, the tumor adapts to commandeer the stromal reaction to further its own growth.
Figure 1.4: The typical progression of prostate cancer. As prostate cancer progresses from early to late disease, cancer cells must first negotiate the matrix surrounding the primary site (A). At this stage, the basal cell layer (black arrow) becomes discontinuous, and the tumor cells acquire invasive capacity, including the ability to degrade BMs and the ECM of stromal tissue. After intravasation (black arrowhead), cancer cells then most often extravasate into and colonize the bone marrow (B). In the bone marrow the cells encounter a dynamic cellular milieu including osteoblasts (curved blue cells), stromal cells (blue cells surrounding tumor) and various hematopoietic cell types. Cells must degrade and migrate through (at least) three separate BMs (double black lines in A, B) in order to complete this process. Adapted from Shiozawa et al. 2008.
**Tumor-associated macrophages**

There are thought to be two classes of macrophages, derived from monocytes in response to various stimuli. M1 macrophages are associated with tumor suppression and are elicited in response to lipopolysaccharide or interferon-\(\gamma\). These cells have anti-tumoral activity through reactive oxygen species, interleukin-12 [Verreck et al., 2004] and TNF-\(\alpha\) secretion. M2 macrophages are the typical tumor-associated macrophage (TAM) and are observed in wound healing, recruited by the cytokine CCL2/MCP-1 [Bottazzi et al., 1983], or various interleukin molecules [Stein et al., 1992]. Thus the M2-polarized TAM represents yet another wound healing/host response mechanism co-opted to foster tumor growth and angiogenesis [Comito et al., 2013]. Pre-polarized monocytes can be recruited by prostate cancer cells or stromal cells through paracrine interactions. This results in one of the oft-bemoaned “vicious cycles” prominent in cancer progression, where cancer cells recruit fibroblasts and TAMs, which secrete mitotic factors for cancer cells, which results in more cell growth and recruitment.

**Prostate cancer metastasis**

**Epithelial to mesenchymal transition**

During prostate cancer progression, cancer cells gain the motility and invasive phenotype required to exit the primary tumor site and spread elsewhere in the body (figure 1.3). The loss of luminal organization as well as breakdown of the basal cell layer and cancer cell invasion involve a cascade of gene expression resulting in a phenotypic shift known as epithelial to mesenchymal transition (EMT) [Xu et al., 2006]. EMT is thought to be central to prostate cancer metastasis [Odero-Marah et al., 2008]. The transition increases both
motility and invasiveness of cancer cells as well as a decrease in expression of BM genes and their integrin receptors [Drake et al., 2010]. Downregulation of E-cadherin and upregulation of N-cadherin promote movement away from epithelia [Xu et al., 2006], as do changes in integrin expression patterns [Bisanz et al., 2005; Zheng et al., 1999]. Upregulation of various MMPs increases invasiveness by promoting the cancer cells’ ability to degrade the BM and the stromal ECM [Lovaas et al., 2013]. Upon reaching the metastatic site, tumor cells must undergo a reverse mesenchymal to epithelial transition (MET), because metastatic lesions often demonstrate an epithelial phenotype [Banyard et al., 2013]. The EMT phenotype is promoted by various factors in the tumor microenvironment and by adaptation of the cancer cells to their surroundings.

**Bone metastasis**

The predominant site of metastasis for prostate cancer is the bone marrow [Bubendorf et al., 2000; Shiozawa et al., 2008], particularly of the spine in the manner of Batson’s venous flow hypothesis [Batson, 1940]. The most common modes of detection for prostate cancer bone metastasis are PET scan or technetium 99m bone scintigraphy (bone scan). Bone scans detect integration of the radioactive technetium-labelled methylene diphosphonate into hydroxyapatite, not a direct indicator of cancer cells, but rather an indicator of rapid formation of calcified bone matrix [Morris et al., 2013]. PET scans detect metabolic uptake of radioactive isotopes of fluorine or carbon [Russell and Gacesa, 1988]. In theory, preferential uptake of these radioactive compounds allows for visualization of metastatic lesions in 3D reconstructions. At this stage of the disease, the metabolic switch (or Warburg effect) from oxidative phosphorylation to glycolysis-dominated ATP production increases the cancer’s effect on patient quality of life [Bongaerts et al., 2006; Watson et al.,
2013]. This is associated with cachexia in late-stage prostate cancer patients, as the cancer cells are consuming most of the calories taken in by the patient, a phenotype exacerbated by inflammatory conditions [Nakashima et al., 1998; Pfitzenmaier et al., 2003; Vaughan et al., 2013]. Unfortunately, after androgen deprivation therapy (ADT), metastatic cells frequently are no longer visible by PET scan, indicating some effect of ADT on prostate cancer cell metabolism [Oyama et al., 2001]. Once prostate cancer has colonized the bone marrow, there are few effective chemotherapeutic treatments, and the cancer has typically advanced to lethal disease.

Avenues of dissemination to the bone marrow include the circulation [Bekes et al., 2011], the lymphatic system [Roma et al., 2006] and perineural invasion [Hassan and Maksem, 1980]. There are a number of hypotheses to explain the predilection of prostate cancer cells to home to the bone marrow. One is that the sinusoidal blood vessels in the bone marrow slow the rate of blood flow during retrograde flow such that the cancer cells have an opportunity to bind the walls of the vessel and extravasate [Batson, 1940; Shibasaki et al., 1983]. This could facilitate binding of extracellular carbohydrate moieties to bone marrow endothelial cell E-Selectin [Barthel et al., 2013; Yin et al., 2010]. Work in a murine lung cancer model suggests that perlecan present in the endothelial BM acts as a cell adhesive molecule during intravasation [Tovari et al., 1997]. Several studies link CXCR4 expression [Akashi et al., 2008; Sun et al., 2003] by prostate cancer cells as the factor attracting them to the bone marrow, a site rich in the CXCR4 ligand SDF-1 [Conley-LaComb et al., 2013; Orimo et al., 2005]. Cancer cells that can adapt to the bone marrow microenvironment often remain dormant for extended periods of time, and may enter a differentiated state known as neuro-endocrine differentiation (NED) [Zhang et al., 2011],
associated with ADT [Ismail et al., 2002; Nordin et al., 2013; Zelivianski et al., 2001]. In this state, cells are resistant to apoptosis, but do not divide.

Once prostate cancer cells adapt to survive in the bone marrow, they most commonly form osteolytic lesions [Nemeth et al., 2002], which result in bone resorption and decrease in bone mass. This creates an environment rich in signaling molecules, the end result of which is yet another phenotypic shift, from osteolytic to osteoblastic metastasis [Shimazaki et al., 1992]. This shift includes a change in expression profile that is known as osteomimicry, wherein cancer cells often express proteins thought to be specific to osteoblasts [Huang et al., 2005]. The net effect of osteomimicry is another vicious cycle: prostate cancer cells express paracrine factors inducing osteoblastic differentiation [Koutsilieris et al., 1987] of mesenchymal stem cells, while these cells and release of factors previously trapped in calcified bone matrix effect mitosis of prostate cancer cells. All the while these osteoblastic lesions result in the creation of a disorganized bone matrix known as woven bone [Aoki et al., 1987]. ADT, most commonly the first line of treatment for disseminated prostate cancer, can induce an artificial osteoporosis [Hamilton et al., 2010]. This in addition to woven bone formation by osteoblastic metastasis leads to fractures [Greenspan et al., 2013]. Additionally, bony metastases often cause intractable pain (bone pain) attributed to various aspects of tumor burden but of unclear origin [Coleman, 2006]. Thus the bone metastasis of prostate cancer results in serious quality of life concerns.
Chapter 2

Introduction

Perlecan is found in the BM of most tissues [Hassell et al., 1980; Laurie et al., 1988] and the pericellular matrix of many tissues, including bone marrow [Melrose et al., 2002; Schofield et al., 1999]. Upregulation of perlecan during stromal activation in response to tumor progression has been observed in oral cancers, and in this work in prostate cancer [Ikarashi et al., 2004; Warren et al., 2014]. In this work, the various cell lines used to represent the stages of tumor progression and stromal cells which prostate cancer cells may interact with express the HSPG2 gene at varying steady-state levels. This will certainly have consequences on cell-cell interactions and cell behavior in culture, given the perlecan protein’s role in cell signaling and maintenance of ECM integrity [Behrens et al.; Savoré et al., 2005]. The bone marrow stromal cell lines HS-5 and HS-27a, the prostate stromal cell line WPMY-1 the early stage prostate cancer cell line LNCaP and the late stage prostate cancer cell line C4-2B are used here as models for each of these cell types in vivo.

The human HSPG2 proximal promoter has been studied in various cell lines and is activated by TGFβ signaling [Iozzo et al., 1997]. The distal region of the promoter is not completely understood and contains potentially functional elements conserved in the mouse genome. The signaling environment of prostate cancer is complex. To parse the signals responsible for increasing perlecan deposition in vivo will be an extraordinary task given cancer’s heterogeneity. Various signaling pathways induce the wound healing and reactive stromal phenotypes including TGFβ and inflammatory cytokines amongst others [Roberts et al., 1986; Shaw et al., 2009]. The overlapping phenotypes of myofibroblasts in wound healing and reactive stromal cells in cancer may be induced by similar signaling
cascades. In normal wound healing, stromal cells differentiate into myofibroblasts and deposit new ECM and close the wound. At this point, myofibroblasts enter apoptosis, although in the co-opted wound healing phenotype found in the reactive stroma of cancer, these myofibroblasts persist [Seemayer et al., 1979].

The data presented here represent general observations on HSPG2 transcription and perlecan protein secretion by the aforementioned cell lines and isolated primary bone marrow stromal cells. The HSPG2 5’ region is studied in detail using promoter-reporter assays and treatments with various potential activators of the HSPG2 gene as well as in co-culture experiments. Cell phenotypes and survival are observed during treatments with cytokines and growth factors and conclusions are drawn concerning the effect of inflammatory cytokines on stromal cell viability.

**Materials & Methods**

**Cell Culture**

For culture of cell lines, refer to chapter 3.

Primary bone marrow stromal cell isolates were extracted by aspiration through the posterior superior iliac spine, performed at Christiana Care Hospital (Newark, DE) during three collections from 2003, 2006, and 2009. Aspirates were suspended in DMEM and allowed to settle overnight at 37°C. The supernatant was removed and placed in a 150 cm² (Corning, Corning, NY) flask with DMEM supplemented with 10% (v/v) fetal bovine serum (Atlanta Biologicals, Flowery Branch, GA), penicillin (1000 units/mL), and streptomycin (1000 units/mL). These cells were grown to confluence, passaged once, and then frozen in liquid nitrogen until thawing in 2012. They were cultured in supplemented
DMEM, and passaged a maximum of five times. For treatments, cells were plated in 12-well plates and grown to 90% confluency, then incubated in serum-free medium during treatment.

**Treatments**

Cytokines and growth factors were used as follows. Cells were treated in their respective growth medium without serum or antibiotics. Tumor necrosis factor alpha (TNF-α) (Roche Applied Science – Indianapolis, IN - #11371843001), interleukin 1-beta (IL-1β) (Roche), transforming growth factor beta 1 (TGFβ1) ((R&D Systems Inc. – Minneapolis, MN. #100-B) were used at 10 ng/mL each. IGF-1(R&D Systems) was used at 200 ng/mL. Forskolin (Sigma Aldrich – St. Louis, MO) and the synthetic Hedgehog pathway agonist HhAg1.5 (Cellagen Technology - San Diego, CA) were each used at 10 μM. Interferon-γ (Roche) was used at a concentration of 200 units/mL. Beta-2 microglobulin was used at a concentration of 1 μg/mL. Vehicle controls were used appropriately for each treatment.

**Co-cultures and conditioned medium**

Co-culture experiments were performed directly. Transient transfection was performed as described in chapter 3, and cells were allowed to recover overnight in medium supplemented with 1% (v/v) fetal bovine serum. The following day, 200,000 cells were plated directly onto the transfected cell line. These cells were cultured for 24 h, and the luciferase assay was performed as described in chapter 3. Controls in these experiments were serum-free medium, or plating the identical cells on top of transfected cells.

Conditioned medium was collected from 150 cm² flasks with cells grown to 90% confluency. These cells were cultured in serum and antibiotic-free DMEM. Conditioned
medium was collected every 48 h for two weeks and stored at -80°C. These aliquots then were pooled and sterile filtered. During conditioned medium treatments, cells were transfected as described in chapter 3 and allowed to recover overnight in medium supplemented with 1% (v/v) fetal bovine serum. Controls in these experiments were conditioned medium derived from the same cell line, or serum-free medium.

**Cloning the HSPG2 upstream region**

The full-length HSPG2 promoter-reporter was created as described in chapter 3. Serial deletions were created by polymerase chain reaction (PCR) using the promoter-reporter plasmid as template, TOPO and restriction enzyme cloning using the same scheme as described in chapter 3 for the creation of the full-length promoter-reporter plasmid. Primers can be found in table A.1.

The distal 69 base-pair region responsive to TNF-α treatment was cloned using oligonucleotides with restriction enzyme sites adapted to either end. These oligonucleotides were ligated into the pGL4.74 plasmid and confirmed by sequencing. Oligonucleotide sequences can be found in the appendix.

Sequences were manipulated and presented in this and all aspects of this project using the Geneious software package 5.4.6 (Biomatters, Auckland, New Zealand).

**RNA isolation and qPCR**

For details regarding RNA isolation qPCR, refer to chapter 3.
Immunoblotting

For details regarding immunoblotting for the perlecan in conditioned medium of cultured cells, refer to chapter 3.

WST-1 assay for metabolically active cell content

For details regarding use of the WST-1 assay for measuring metabolically active cells in culture, refer to chapter 3.

Genomatix Matinspector analysis of the HSPG2 5’ region in human and mouse

The 2,565 base pairs 5’ of the HSPG2 gene in both the human and mouse genomes were downloaded from the Ensembl database. These sequences were uploaded to the Matinspector program [Cartharius et al., 2005]. The matrix family library version 6.3 was used, specifically the “All vertebrates” library. The results were then compared manually for the conservation of consensus binding elements for transcription factors. Elements in the human HSPG2 5’ region were deemed minimally conserved if there was encoded a corresponding element with 250 base pairs 5’ or 3’ of the analogous position relative to the start site of transcription in the mouse genome or highly conserved if there was encoded a corresponding element within 50 base pairs 5’ or 3’ of the analogous position relative to the transcriptional start site in the mouse genome.

Luciferase reporter assays

For details regarding the luciferase reporter assay, see chapter 3.
Results

*HSPG2 mRNA expression is highest in stromal cells and lowest in C4-2B cancer cells*

To determine at what level each of our cell lines express the *HSPG2* transcript, cells were grown to 90 percent confluence and incubated in serum-free medium for 48 h prior to collection of RNA samples. The relative expression levels of *HSPG2* were highest in the prostate stromal cell line WPMY-1, with a delta C$_t$ of 7.7 compared to the internal control beta actin, which had an average C$_t$ of 15 in every cell line. The largest Delta Ct values occurred in the HS-5 (11.09) and C4-2B (11.13) cell lines. Data are expressed in figure 2.1 as a percent of the beta actin transcript. WPMY-1 cells expressed the *HSPG2* transcript at a significantly higher rate than every other cell line in this study. HS-27a cells expressed *HSPG2* at a significantly (p<.01) higher level than all cell lines except WPMY-1 cells. LNCaP cell expression of *HSPG2* transcript was nearly twice as high but not significantly different than HS-5 or C4-2B cell expression. The primary bone marrow stromal cell sample number seven isolated in 2003 expressed the *HSPG2* transcript at by far the highest rate, at 1.7% of *ACTB*, nearly four times the rate as that of the WPMY-1 cell line and more than 20 times that of the prostate cancer cell lines. This expression in primary cells was little affected by treatment with cytokines or growth factors at 72 h.
Figure 2.1 Basal levels of HSPG2 mRNA in various cell lines. HSPG2 mRNA was measured in these cell lines (A) and primary bone marrow stromal cells (B) using qPCR with beta actin as an internal control. Cells at 90 percent confluence were changed to serum-free medium for 48 h (A) or 72 h during treatment of primary cells (B) prior to collection of RNA samples. Data are expressed as the percent of beta actin expression. Numbering over columns represents Delta Ct values when compared with the beta actin internal control. Beta actin Ct values were invariably between 15 and 16. Bars, mean +/- SEM, n = 3. ***, p<.001 vs WPMY-1. *, p<.05 vs WPMY-1.
Primary bone marrow stromal cells’ response to TNF-α and TGFβ1 treatment varies in concordance with cell number

Primary bone marrow stromal cells collected from iliac spine aspiration were cultured in serum-free medium supplemented with TNF-α, TGFβ1, a combination of the two or control for 72 h. The conditioned medium was collected and assayed by immunoblot targeted at perlecan. Each cell sample secreted perlecan in abundance but only a subset of the samples increased perlecan secretion during treatment with growth factors and cytokines. Cells derived from a prostate cancer patient designated “7” in 2003 and cells extracted from a normal patient in 2005 both significantly increased perlecan secretion during combination treatment with TNF-α and TGFβ1 (figure 2.2A). TNF-α treatment alone caused an increase in perlecan secretion by the “7” cells but not the normal cells from 2005. Other cell isolates demonstrated a trend to increasing perlecan secretion, but none were significant. Data from all of these cell isolates were averaged (figure 2.2B) and the trend towards increasing perlecan secretion during TNF-α treatment alone or in combination with TGFβ1 was significant once these data points were compiled.

Treatment with TNF-α and/or TGFβ1 decreases cell number in stromal cell lines but increases cell number in primary bone marrow stromal cell isolates

Although perlecan secretion increased in the conditioned medium of cells treated with TNF-α with or without the addition of TGFβ1, these treatments concomitantly increase cell number in each of these primary cell isolates after 72 h. In each primary cell isolate (figure 2.3A) there was an increase in cell number in the combined treatment of TNF-α and TGFβ1, although this was not significant in two of the samples. In three of eight samples, TGFβ1 alone increased cell number and in one sample TNF-α alone increased cell number.
The response was variable among samples, but in each case where perlecan protein secretion increased after 72 h in these cell isolates, there was an accompanying increase in cell number.

Because TNF-α and TGF-β treatments are used routinely in chapter 3 and affect perlecan secretion and transcription, it was pertinent to perform assays for cell number following typical treatment protocols. These assays were done in the cell lines that increase HSPG2 promoter activity in response to these treatments, namely, the HS-27a, WPMY-1 and LNCaP cells. The WST-1 assay is a measure of cell metabolic activity, which we use as a proxy for cell number. In each case, values were normalized to the vehicle control. After 48 h treatment with TNF-α, there was seen a decrease in cell number in each cell line. This difference was significant in the WPMY-1 cells. After 48 h treatment with TGFβ1, there was also a decrease in cell number in both the HS-27a and WPMY-1 stromal cells, but not LNCaP cells. This decrease was significant in the WPMY-1 cells. Combination treatment decreased cell number in all three cell lines, but was not significant in any of them.
Figure 2.2 Perlecan secretion varies during cytokine treatment in different primary bone marrow stromal cell populations. Primary bone marrow stromal cells isolated from bone marrow aspirates were treated for 72 h in serum-free medium with indicated cytokine and growth factor combinations. The resulting conditioned medium was immunoblotted for perlecan protein. Each sample cell population was studied independently (A) and the results for all of the cell lines were averaged (B). Data are normalized to the vehicle control. Bars, mean +/- SEM. At least three replicates were performed for each treatment condition. **, p<.01 vs vehicle control. *, p<.05 vs vehicle control.
Figure 2.3. Cell number is affected by treatment with TNF-α and/or TGFβ1. Cells were treated for 48 hours with as indicated. WST-1 reagent was added to culture medium and absorbance was measured after 1 hour. Data is plotted as absorbance in treated cells normalized to that of untreated cells. Primary bone marrow stromal cell isolates were tested to produce data in (A) and cell lines were used to produce data in (B). Error bars, mean +/- SEM. At least three replicates were performed for each condition. *, p<.05 vs control. **, p<.01 vs control. ***, p<.001 vs control.
**In silico analysis of HSPG2 promoter reveals conserved consensus elements**

An analysis using the Matinspector program revealed a number of consensus transcription factor binding elements encoded by the upstream region of the HSPG2 gene that are conserved in both humans and mice (figure 2.4). The 2618 base pairs 5’ of the transcriptional start site were analyzed. Putative transcription factor binding elements that were found in both the human and mouse upstream regions within 500 base pairs of each other relative to their respective start sites of transcription (minimally conserved elements) are highlighted in blue in figure 2.4. Highlighted in pink are those sites which occurred in the mouse and human promoters within 100 base pairs of each other relative to their respective start sites of transcription (highly conserved elements). Green annotations represent the primer sequences used to amplify the HSPG2 promoter serial deletions used in later figures. The 5’ UTR shown in this figure represents the region included in the promoter-reporter construct. The area within 400 base pairs of the start site of transcription showed a great deal of conservation, matching highly conserved elements nearly exactly. Important to chapter 3 of this dissertation, the NF-κB consensus elements found nearly 2500 base pairs upstream of the transcriptional start site are highly conserved. There is a concentration of minimally conserved elements in the region from 1400 to 1800 base pairs upstream of the transcriptional start site. There were a number of predicted elements not highlighted in this figure that also were found in the mouse gene upstream region, but were not found within 500 base pairs of the analogous element in the human promoter.
Figure 2.4 Conserved CREs in the region 5’ of the HSPG2 gene. The human and mouse promoters were analyzed using an in silico approach as described. The human sequence is shown above, with the 5’ upstream region denoted in green. The 5’ UTR, included in the promoter-reporter constructs used in this study, is denoted in grey, and exon 1 of the HSPG2 transcript is denoted in red. Primer sequences used to amplify the serial deletion constructs are denoted in dark green. Putative transcription factor binding elements predicted by our analysis are shown in blue and pink. Pink annotations denote transcription factor binding elements located at precisely the same location relative to the start site of transcription upstream of both the mouse and human gene, or at a maximum 50 base pairs up or downstream from the same location relative to the start site of transcription. Blue sequences denoted elements which were conserved upstream of both the mouse and human gene, at a maximum of 250 base pairs up or downstream from the same location relative to the start site of transcription.
**HSPG2 promoter-reporter constructs used in this study**

Various *HSPG2* promoter-reporter constructs were created for systematic analysis of active elements in the promoter region. The “full-length” promoter (figure 2.5 A) included the 5’ UTR as well as 2,618 base pairs upstream of the start site of *HSPG2* transcription. The various conserved *cis* elements identified in figure 2.4 were studied by creating systematic serial 5′ deletions of the promoter-reporter construct (figure 2.5 B, C, D) and by inserting potentially active regions of the promoter upstream of a heterologous promoter in a different reporter plasmid backbone (figure 2.5 E). Each of these constructs is used the following data. The SMAD3 element found over 1000 base pairs upstream of the start site was not among the minimally conserved elements illustrated in figure 2.4, but was predicted in the human upstream region.
Figure 2.5 *HSPG2* promoter-reporter constructs used in this study. The “full-length” *HSPG2* promoter-reporter (A) includes the region up to 2.6 kilobase pairs upstream of the start site of transcription. Serial deletions of 2.4 kilobase pairs (B), 1.7 kilobase pairs (C), and 1.3 kilobase pairs (D) were constructed to specifically eliminate various putative transcription factor binding elements. Putative transcription factor binding elements conserved in the mouse and human genome as indentified in figure 3 are denoted by rectangles containing their various identifiers. A heterologous (Herpes simplex virus) promoter-driven reporter construct was modified (E) by upstream insertion of the 50 base pair region encoding two putative NFkB transcription factor binding motifs found in the *HSPG2* promoter.
**HSPG2 2.6 kb promoter-reporter has the highest baseline activity in WPMY-1 and HS-27a stromal cells, and lowest in LNCaP cells.** The full-length HSPG2 promoter-reporter construct was co-transfected into stromal and cancer cell lines with the thymidine kinase promoter-driven *Renilla* luciferase internal control. Baseline promoter-reporter activity levels were drawn from vehicle controls performed during treatments with various cytokines and growth factors. These data are meant to reflect relative transcriptional activity in these cell lines (figure 2.6A). The lowest reporter activity was recorded in LNCaP cells (13 times greater than control), followed by C4-2B and HS-5 (52 and 59 times greater than control) and HS-27a and WPMY-1 (both at 242 times greater than control). It is worth noting that this number does not directly translate from relative mRNA levels in these cell lines. These data reflect luciferase assays performed using responsive HS-27a cells. When these cells have passed beyond their window of responsiveness, their baseline HSPG2 promoter-reporter activity drops significantly.

Serial deletions of the upstream region of *HPSG2* were created to systematically analyze promoter-reporter responses to stimuli. In each of these deletion constructs, promoter-reporter activity was diminished for the HS-27a to below 25, 15 and 40 times greater than the *Renilla* control in the 2.4, 1.7 and 1.3 kb promoter-reporters, respectively (figure 2.6 B,C,D). WPMY-1 cell activity was reduced similarly in deletion constructs, with the exception of the 1.3 kb promoter-reporter (C), where activity was significantly (70 times greater than control) higher than activity in other cell lines. HS-5 cells always exhibit the least activity with any of the serial deletion constructs. LNCaP cell activity is only modestly decreased in the 2.4 kb deletion construct. C4-2B cell activity decreases to 8 times greater than control in the
Figure 2.6 Baseline promoter-reporter activity levels are highest in stromal cell lines. Luciferase activity was measured in the full-length promoter-reporter (A), 2.4 kb promoter-reporter (B), 1.7 kb promoter-reporter (C), and 1.3 kb promoter-reporter (D). Untreated cells were assayed for luciferase activity levels 48 h after transfection. Promoter-reporter luciferase activity is divided by activity of the co-transfected control, a thymidine kinase promoter-driven Renilla luciferase. Error bars, mean +/- SEM. At least three replicates were performed for each condition. In (A) *, p<.05 vs HS-27a, WPMY-1, **, p<.01 vs HS-27a, WPMY-1, ***, p<.001 vs HS-27a, WPMY-1. In (B), ***, p<.001 vs HS-27a. In (C) ***, p<.001 vs WPML-1. In (D), ***, p<.001 vs HS-27a, WPMY-1, C4-2B.
2.4 kb deletion, but rises to 30 and 40 times greater when using the 1.7 or 1.3 kb deletion constructs.

The 50 base pair region of the *HSPG2* promoter containing the putative NFκB binding elements is insufficient to increase expression driven by a heterologous promoter.

To determine if the region housing the NFκB consensus elements of the *HSPG2* promoter was sufficient to drive the response to TNF-α treatment, this 50 base pair region was inserted upstream of the heterologous herpes simplex virus thymidine kinase (HSV-TK) promoter driving a *Renilla* luciferase reporter. WPMY-1 cells were tested for their response to TNF-α when co-transfected with this construct and a CMV-driven firefly luciferase reporter construct. There was a significant (p<.001) decrease in *Renilla* luciferase activity when treated with TNF-α, down to 20% of the activity of the samples treated with vehicle control. This indicates that while the NFκB elements are necessary (figure 2.7) for TNF-α induced activation of the *HSPG2* promoter, they are not sufficient to drive this activation.
Figure 2.7 The 50 base pair region containing the consensus NFκB elements is insufficient to drive TNF-α responsiveness. Luciferase activity was measured in an HSV-TK promoter driven Renilla luciferase vector, with the 50 base pair region of the perlecan promoter containing two consensus NFκB elements inserted upstream of the promoter (see figure 4). Cells were treated with or without TNF-α for 24 h, starting 24 h after transfection and then harvested for luciferase assays as described in Materials and Methods. Promoter-reporter luciferase activity is divided by activity of the co-transfected control, a CMV promoter-driven firefly luciferase, and normalized to the same activity index of the vehicle control. Error bars, mean +/- SEM, n = 3. ***, p<.001 vs control.
Treatments with various cytokines have no effect on 2.6 kb promoter-reporter activity in many cell lines.

There exist several conserved transcription factor binding motifs in the HSPG2 promoter (figure 2.8) and the full-length promoter-reporter construct was used to test their responsiveness to their typical signaling molecules. Stromal and prostate cancer cell lines did not respond to IL-1β, Beta-2 Microglobulin (β-2M), the synthetic hedgehog pathway agonist HhAg1.5, IGF-1, IFNγ, or the cAMP pathway activator, forskolin, at the concentrations indicated in Methods. IL-1β was reported previously to stimulate HSPG2 expression in human umbilical vein endothelial cells [Reine et al., 2012], β-2M and forskolin both should in theory activate cAMP and thus CREB, IGF-1 in some cell contexts activate ELK-1. IFNγ is reported to decrease HSPG2 expression in a colon carcinoma cell line [Sharma and Iozzo, 1998]h, and the Hedgehog pathway is thought to interplay with the reactive stroma phenotype. As stated, none of these treatments had any effect on HSPG2 promoter-reporter activity, although IFNγ in combination with TNF-α induced promoter-reporter activity. This response is little different from that triggered by TNF-α treatment alone, as will be expounded upon in chapter 3.
Figure 2.8 HSPG2 promoter activity is unresponsive to many treatments. Luciferase activity was measured in the full-length HSPG2 promoter-reporter. Untreated cells were assayed for luciferase activity levels 48 h after transfection. Treatments were performed for 24 h in serum-free medium 24 h after transfection. Treatments were performed at concentrations indicated in methods. Promoter-reporter luciferase activity is divided by activity of the co-transfected control, a thymidine kinase promoter-driven Renilla luciferase, and normalized to the same activity index of the vehicle control. Error bars, mean +/- SEM. At least three replicates were performed for each condition. *, p<.05 vs control.
Co-culture of cancer and stromal cell lines or their conditioned medium has minimal effect on HSPG2 promoter-reporter activity.

The overarching hypothesis of this work is that cell-cell interactions in the tumor microenvironment initiate tumor progression and stromal reaction, and these phenotypes include increased HSPG2 expression. To pursue this hypothesis, stromal cells and cancer cells were co-cultured after transfection with the HSPG2 promoter-reporter construct (figure 2.9). There was a modest but insignificant decrease in luciferase activity in C4-2B cells transfected with the reporter and co-cultured with HS-27a cells. The only significant activity change in any of the transfection/co-culture combinations was when the LNCaP cells stably transfected with the empty vector control (LN-Neo) cells were co-cultured with WPMY-1 cells transfected with the reporter (B). In this case, the luciferase activity level was significantly higher than that of controls. In cells treated with conditioned medium from their counterpart cell line, the only response seen again was in the conditioned medium from the LN-Neo stable cell line (C). Although not significant, when reporter-transfected HS-27a cells were incubated with LN-Neo stable cell conditioned medium, there was a nearly 2-fold increase in reporter activity. Somewhat surprising is the minimal effect of HS-5 co-culture or conditioned medium on LNCaP and C4-2B cells. Although these conditions elicited the NED transdifferentiation effect seen in previous reports [Chang et al., 2014; Zhang et al., 2011], there was no effect on HSPG2 promoter-reporter activity.
In various replicates, stromal cells treated with TNF-α in combination with TGF-β1.

**Figure 2.9 HSPG2 promoter is unresponsive to physical co-culture and conditioned medium.** Luciferase activity was measured in the full-length HSPG2 promoter-reporter. Cells were plated onto transfected cells 24 h after transfection, and luciferase activity was measured 24 h after plating. The labels on the abscissa of panels A and B indicate the cell line that was plated onto the transfected cells, which are indicated by the legend. In panel C, the abscissa labels denote what cells were used to condition the medium that was used in these experiments. Conditioned medium was added 24 h after transfection, for 24 h. Conditioned medium was created as described in Methods. Values were normalized to the control, which was serum-free medium. Promoter-reporter luciferase activity is divided by activity of the co-transfected control, a thymidine kinase promoter-driven Renilla luciferase, and normalized to the same activity index of the vehicle control. Error bars, mean +/-SEM. At least three replicates were performed for each condition. ***, p<.001 vs control.
Treatment with TNF-α induces a morphological change in stromal cell types

In various replicates, stromal cells treated with TNF-α in combination with TGF-β1 underwent a dramatic change in cell-cell organization. This change is characterized by creating clearings in the well and cells mounding on top of one another in roughly circular patterns (figure 2.10). In HS-5 cells, this occurred independently of treatment, the driving factor being removal of serum from the medium. It is possible that removal of serum predisposes HS-27a and WPMY-1 cells to this change as well, as some small clearings start to form in vehicle treatments of these cells. The organizational change is at its most noticeable in these cells during TNF-α and TGFβ1 combination treatment. While TNF-α alone can drive a weak approximation of this phenotype during serum-free treatment, the full phenotype was only observed when TGFβ1 treatment was combined with TNF-α.
Figure 2.10 Stromal cell line cell-cell organization shifts under TNF-α and TGFβ1 combination treatment. Cells were treated for 24 h with the combination of TNF-α (10 ng/mL) and TGFβ1 (10 ng/mL) during promoter-reporter luciferase transfection experiments. Before harvesting, cells were viewed with a 10X objective. Scale bar, 100 μm.
Discussion

Signaling molecules in cancer and stromal cell lines

This work represents an analysis of expression of the human HSPG2 gene and protein in various cell lines in response to various cytokines, growth factors and other cell lines. Under baseline conditions, primary bone marrow stromal cells express the perlecan transcript at four times the level of the WPMY-1 stromal cells, which express perlecan at the highest rate of all of the tested cell lines. The WPMY-1 and HS-27a stromal cells, in turn, express the perlecan transcript at significantly (6-10 fold) higher levels than the HS-5 bone marrow stromal cell line, or the LNCaP and C4-2B prostate cancer cell lines. Unstimulated LNCaP cells express roughly twice as much HSPG2 transcript (relative to ACTB mRNA) than do C4-2B cells. Nonetheless, C4-2B cells still rely on expression of HSPG2 to form tumors in mouse models and to grow in an anchorage-independent environment [Savoré et al., 2005].

The fibroblastic response during normal wound healing is to migrate to the site of the wound, initiate differentiation to a myofibroblastic phenotype and begin remodeling of the ECM while contracting to close the wound [Darby et al., 1990]. After this process is complete, these cells will under normal circumstances enter apoptosis [Desmouliere et al., 1995]. In our system, the WPMY-1, HS-27a and LNCaP cell lines all decreased in cell number following 48 h treatments with TNF-α or a combination of TNF-α and TGFβ1, and WPMY-1 stromal cells decreased in cell number following TGFβ1 treatment alone. This is in accord with the apoptosis of fibroblasts following induction of the wound healing phenotype, and increased secretion of ECM and signaling molecules. Primary bone marrow stromal cells, on the other hand, are unaffected or slightly increase in cell number.
following treatment with TNF-α, TGFβ1 or a combination of the two. This was unexpected, and may reflect a difference between primary stromal cell isolates and the clonal population of the HS-27a bone marrow stromal cell line. The HS-27a cell line represents a very specific subpopulation of the bone marrow stromal niche, and it seems that their response to signaling by cytokines and growth factors is much more acute than that of the bone marrow stromal cell population as a whole. This slight increase in cell number following treatment with TNF-α, TGFβ1 or a combination of the two occurred in almost all of the primary bone marrow stromal cell samples that were tested, and was closely correlated with an increase in perlecan secretion into the conditioned medium when tested in parallel.

**Characterization of the HSPG2 gene regulatory region**

The HSPG2 gene regulatory region is in all probability much larger than appreciated before this study was begun. Recent data produced by the ENCODE consortium details the chromatin activation state of the entire genome [Qu and Fang, 2013]. This data is publically available to view through the UCSC genome browser. The data for the HSPG2 gene can be found in appendix B figure 1 and is described in detail in the discussion of chapter 3. In this study, the analysis of the HSPG2 5’ region was performed using the Genomatix Matinspector software. This region of the human gene was compared to the analogous area of the mouse gene. Meticulous search of the resulting data identified both minimally and highly conserved consensus elements for transcription factor binding. Of interest are the highly conserved NFκB consensus elements in the most distal part of the 5’ region. These elements are further examined in chapter 3. While identifying conserved motifs is not a guaranteed predictor of regulatory elements, in this case these highly
conserved NFκB elements, in fact, turned out to regulate HSPG2 expression. Further analysis and understanding of HSPG2 transcriptional regulation will be a referendum on the utility of identifying minimal versus high conservation in the manner demonstrated here. A comparison of minimally conserved elements as presented here with ChIP-seq data available through the ENCODE data set of the UCSC genome browser revealed that none of our minimally conserved elements were identified as active during the ENCODE project. This is also true of the NFκB site studied in chapter 3, so it is possible that there are regulatory sites not identified by the ENCODE project.

Sequential 5′ deletion of the promoter-reporter constructs gave insight into the relative activity of the upstream region of the HSPG2 gene. The “full-length” promoter-reporter included the 2618 base pairs 5′ of the start site of transcription as well as 79 base pairs of the 5′ UTR prior to the start codon inserted upstream of the firefly luciferase gene. This construct demonstrated baseline activity mirroring that of the relative HSPG2 mRNA expression in the prostate and bone marrow stromal as well as prostate cancer cell lines. Stromal cell lines HS-27a and WPMY-1 exhibited the highest baseline promoter-reporter activity while the HS-5 bone marrow stromal cell line exhibited far less activity. The only discrepancy between the full-length promoter-reporter activity and relative HSPG2 mRNA expression was between C4-2B cells and LNCaP cells. In this case, C4-2B cells exhibited higher promoter-reporter activity than did the LNCaP cells, although they express less HSPG2 mRNA than LNCaP cells. This could mean that although this isolated sequence is sufficient to support the general trend of expression of the HSGP2 gene, there are additional functional elements up- or downstream of the start site of transcription not included in this construct. Sequential 5′ deletion of the construct affected the activity of the promoter-
reporter; sequential deletions demonstrated relative activity levels different than that of baseline mRNA expression. This suggests that the distal region of the promoter (beyond 2.4 kilobase pairs upstream) contains functional elements crucial to transcriptional control of HSPG2. While the highly conserved NFκB consensus elements are necessary for activation of the promoter by TNF-α signaling, the insertion of these elements was not sufficient to drive this response out of the context of the full-length promoter. This means that these elements are activated in concert with other functional elements in the HPSG2 promoter to drive HSPG2 mRNA and protein expression in response to TNF-α treatment.

The approach of this study was to use the promoter-reporter assay as a convenient screen to detect activation of the HSPG2 gene by signaling molecules potentially important to prostate cancer progression. This proved valid for the determination of TNF-α’s effect of HSPG2 expression as described in chapter 3, but did not reveal the effects of any other treatments or co-cultures on HSPG2 expression as detailed in this chapter. Other treatments reported to affect HSPG2 expression such as IL-1β, TGFβ1 or IFNγ showed no response at the HSPG2 promoter-reporter level in any of these tested cell lines. Other treatments hypothesized to affect HSPG2 expression such as β-2M, IGF-1, the synthetic patched agonist HhAg1.5, forskolin or co-culture with complimentary cells also had no effect on activity. While this sort of screen may work for ideal circumstances such as, in this case TNF-α signaling, results may be confounded by incomplete promoter isolation or interference of other inhibitory elements which, when out of the chromosomal context, obscure the signal that would be detected by directly measuring steady-state mRNA levels. In addition, as described in the discussion of chapter 3, the construct here does not contain the full HSPG2 regulatory region, as that region potentially spans a 50 kilobase region up
and downstream of the transcriptional start site. Without using transgenic mice to illustrate an appropriate expression pattern driven by DNA elements with regulatory control over HSPG2 expression, the full promoter of HSPG2 cannot be fully identified.

*Cell-cell interaction during cytokine treatment*

A distinct cell-cell organization shift occurs during treatment of the stromal cell lines with TNF-α and TGFβ1. This shift is an exaggeration of what occurs during removal of serum in these cell lines, including increased spindle cell-like morphology, creation of clearings in the well by cells forming mounds, and alignment of cell directionality. It is possible that this phenotype echoes that of the wound healing or reactive stromal myofibroblast *in vivo*. Quantitative PCR or immunofluorescence for myofibroblast markers during treatment with signaling factors will be an informative method for characterizing this activated phenotype in culture. Various targets such as α smooth muscle actin or the *Col1α1* gene are accepted markers for the myofibroblast phenotype.
Chapter 3

As published in The Journal of Cellular Biochemistry

Transcriptional Activation by NFκB Increases Perlecan/HSPG2 Expression in the Desmoplastic Prostate Tumor Microenvironment

Curtis R. Warren¹, Brian J. Grindel¹, Lewis Francis², Daniel D. Carson¹,³, Mary C. Farach-Carson¹,⁴

¹Department of Biochemistry and Cell Biology, Rice University, Houston, Texas.
²Institute of Life Science, College of Medicine, Swansea University, Swansea, UK.
³Biochemistry and Molecular Biology, M.D. Anderson Cancer Center, Houston Texas.
⁴Department of Bioengineering, Rice University.
Abstract

Perlecan/HSPG2, a heparan sulfate proteoglycan typically found at tissue borders including those separating epithelia and connective tissue, increases near sites of invasion of primary prostatic tumors as previously shown for other proteins involved in desmoplastic tissue reaction. Studies of prostate cancer cells and stromal cells from both prostate and bone, the major site for prostate cancer metastasis, showed that cancer cells and a subset of stromal cells increased production of perlecan in response to cytokines present in the tumor microenvironment. *In silico* analysis of the *HSPG2* promoter revealed two conserved NFκB binding sites, in addition to the previously reported SMAD3 binding sites. By systematically transfecting cells with a variety of reporter constructs including sequences up to 2.6 kb from the start site of transcription, I identified an active *cis* element in the distal region of the *HSPG2* promoter, and showed that it functions in regulating transcription of *HSPG2*. Treatment with TNF-α and/or TGFβ1 identified TNF-α as a major cytokine regulator of perlecan production. TNF-α treatment also triggered p65 nuclear translocation and binding to the *HSPG2* regulatory region in stromal cells and cancer cells. In addition to stromal induction of perlecan production in the prostate, I identified a matrix-secreting bone marrow stromal cell type that may represent the source for increases in perlecan in the metastatic bone marrow environment. These studies implicate perlecan in cytokine-mediated, innate tissue responses to cancer cell invasion, a process I suggest reflects a modified wound healing tissue response co-opted by prostate cancer cells.
**Introduction**

Prostate cancer progression is marked by increasing interaction of stroma with the tumor, and an accompanying phenotypic shift of associated stromal cells first in the prostate then at sites of metastasis [Dakhova et al., 2009]. Activated stromal cells involved in such desmoplastic reactions are variously named cancer-associated fibroblasts, tumor-associated fibroblasts, or reactive stromal cells [Olumi et al., 1999]. A hallmark of desmoplastic stroma in cancer [Coulson-Thomas et al.] is deposition of a fibrotic matrix that closely mimics the wound healing response including the heparan sulfate proteoglycan, perlecan, encoded by the gene *HSPG2* [Yamazaki et al., 2004]. The parallels between the actions of the fibroblasts in wound healing with the persistence of reactive stromal cells during disease has led some to describe cancer as “the wound that never heals” [Schafer and Werner, 2008]. Inflammatory cell recruitment is crucial to the process of wound healing [DiPietro, 1995], and is inextricably woven into tumor progression in prostate cancer [Comito et al., 2013]. Cytokines and growth factors produced at the wound [Kohyama et al., 2004; Siegbahn et al., 1990] or the tumor [Shaw et al., 2009] recruit and activate stromal cells whose role is to produce ECM to encase the site of the perceived wound. The ECM deposition profile of proteins present in reactive stroma includes fibronectin, collagens, and various proteoglycans [Brown et al., 1999; Lagace et al., 1985]. Heparan sulfate proteoglycans, such as perlecan, facilitate growth factor delivery during tissue remodeling or repair [Jung et al., 2013; Zcharia et al., 2005] in addition to filling various extracellular scaffolding [Farach-Carson and Carson, 2007], adhesive [Chen et al., 2005], and boundary setting [Farach-Carson et al., 2013] roles that establish tissue function.
Perlecan is a large (~200 nanometers, 400-800 kDa) [Farach-Carson and Carson, 2007] heparan sulfate proteoglycan found in all BMs [Yurchenco et al., 2002]. It is particularly abundant in the bone marrow, where it is the predominant heparan sulfate proteoglycan [Schofield et al., 1999], and in cartilage, where it resists vascular invasion [Brown et al., 2008]. The perlecan core protein has a modular structure that shares homology with other ECM proteins [Murdoch et al., 1992] while the attached heparan sulfate chains function as a reservoir for growth factors useful in wound healing [Brown et al., 2008; Savore et al., 2005; Yang et al., 2005]. The scaffolding function of the core protein [Behrens et al.; Farach-Carson and Carson, 2007] reinforces the barrier function of the BM – important to denying invasion and metastasis of carcinoma *in situ* [Terranova et al., 1986]. Perlecan helps promote prostate cancer cell viability [Savoré et al., 2005] and is part of the reactive stroma gene expression profile seen in some cancers [Sabit et al., 2001]. *HSPG2* gene expression also increases in bone after fracture [Wang et al., 2006]. Given that bone marrow is both a perlecan-rich environment [Schofield et al., 1999] and the predominant site of prostate cancer metastasis [Bubendorf et al., 2000], it is of interest to study regulation of *HSPG2* gene expression in the context of bone marrow stromal cells, as well as in prostate stromal cells near sites of tumor. Because prostate cancer cells encountering tumor stroma undergo EMT [Zhau et al., 2008], and EMT is associated with expression of ECM [Freire-de-Lima et al., 2011], perlecan production by the prostate cancer cells themselves also is of interest.

A previous study examined the role of local cytokines such as TGFβ or TNF-α on *HSGP2* expression in the tumor microenvironment [Iozzo et al., 1997], a finding not yet examined in prostate cancer although the role of these cytokines in this disease is clear [Dayyani et
al., 2011; Yang et al., 2010]. The proximal promoter region of the perlecan gene contains a functional binding site for SMAD3 [Iozzo et al., 1997]. In the context of bone metastasis, I previously found that TGFβ signaling is dynamically regulated in the bone marrow tumor stroma through down regulation of the TGFβ co-receptor endoglin, which also reduces signaling through SMAD2/3, suggesting that it is unlikely that SMAD3 elements in the perlecan promoter are directly responsible for increases in perlecan expression in the desmoplastic tumor microenvironment [O'Connor et al., 2007]. Both TGFβ and TNF-α can recruit and activate immune and stromal cells at primary and secondary tumor sites [Andrades et al., 1999; Postlethwaite and Seyer, 1995], triggering further increases in inflammatory cytokines and increasing ECM production in an attempt to heal the perceived wound. TNF-α, produced by many immune cells, induces apoptosis in the human bone marrow hematopoietic cell line, HS-5, and in some cancer cells, but not in the HS-27a bone marrow stromal cell line [Byun et al., 2005; Sumitomo et al., 1999]. TNF-α signaling, acting through NFκB, activates transcription of many target genes including that encoding the ECM protein, tenascin [Nakoshi et al., 2008], but its ability to activate HSPG2 gene expression has never been examined. In this study, I systematically examined HSPG2 expression in the tumor microenvironment and by prostate and bone marrow stromal cells in the presence of cytokines, TNF-α and/or TGFβ. I present evidence for inflammatory cytokine-mediated stimulation of HSPG2 expression and perlecan secretion, with implications for wound healing and cancer progression.
Materials & Methods

Immunohistochemistry

Formalin fixed, paraffin embedded prostate cancer and normal adjacent tissue sections were processed as previously described [Tuxhorn et al., 2002] with the following modifications. Antigen retrieval was achieved by incubating the tissue in 0.1 mg/mL bovine testicular hyaluronidase (Sigma-Aldrich - St. Louis, MO) and tissue was blocked with Background Punisher (BioCare Medical, Concord, CA). Anti-perlecan HSPG2 Prestige Rabbit polyclonal antibody (Sigma) was diluted 1:100 in antibody diluent (Ventana Medical Systems - Tucson, AZ) and incubated overnight at 4°C. Tissue was incubated with the Diagnostic Biosystems (Pleasanton, CA) Polymer Penetration Enhancer and incubated with Anti-mouse/Rabbit PolyVue HRP (Diagnostic Biosystems). Tissues were mounted in Cytoseal 60 (Fisher Scientific, Hampton, NH).

Cell Culture

The bone marrow stromal cell lines, HS-5 and HS-27a, the prostate cancer cell lines, LNCaP and C4-2B, and the prostate stromal cell line, WPMY-1, were cultured as previously described [O'Connor et al., 2007; Webber et al., 1999]. Although isolated from the same bone marrow aspirate, the hematopoietic stem cell-supporting HS-5 cell line secretes many growth factors and cytokines, and the HS-27a cell displays a structural secretory profile, including rich production of ECM components [Roecklein and Torok-Storb, 1995]. The prostate stromal cell line, WPMY-1, (ATCC), which expresses α-smooth muscle actin characteristic of reactive stromal cells [Webber et al., 1999], was used to represent stromal cells at the site of the primary tumor. LNCaP cells originate from a
lymph node metastasis of prostate cancer and C4-2B cells are adapted to growing in the bone marrow [Thalmann et al., 2000].

For all treatments, cells were plated into 6-well plates and allowed to grow for 48 hr (until 90% confluent). Media was then changed to 1% (v/v) serum and antibiotic-free medium overnight, and then treatments initiated the next day. Recombinant TNF-α produced by E. coli (Roche Applied Science – Indianapolis, IN - #11371843001) was stored at -20°C at 10 ng/µL stock solution, and cell treatments used a final concentration of 10 ng/mL. TGFβ1 (R&D Systems Inc. – Minneapolis, MN. #100-B) purified from human platelets was reconstituted in 4 mM HCl and 0.1% (w/v) bovine serum albumin (BSA) at 10 ng/mL. TGFβ1 treatments were carried out at 10 ng/mL. BSA was dissolved in phosphate buffered saline (pH 7.4) at 1 mg/mL to serve as the vehicle control.

**Plasmid Construction**

Genomic DNA was purified from HS27a cells using the DNeasy Blood & Tissue kit (Qiagen - Valencia, CA). The 2627 and 2,393 base pair upstream region (hereafter referred to as the 2.6 and 2.4 kb HSPG2 promoter constructs) of the HSPG2 proximal promoter was amplified by genomic polymerase chain reaction (PCR) (for primer pairs see appendix A table I). This PCR product then was cloned into pCR®2.1 TOPO (Life Technologies). This insert ligated into the promoterless pGL4.10 firefly luciferase reporter vector (Promega - Madison, WI). Site directed mutagenesis was performed using the Quickchange II Site-Directed Mutagenesis Kit (Agilent – Santa Clara, CA) according to the manufacturer’s instructions (for primers see appendix A table I).
Transient Transfection

Cells were transiently transfected with luciferase constructs using Lipofectamine™ 2000 reagent (Invitrogen) according to the manufacturer’s instructions. A ratio of 1 µg of total DNA to 2 µL (for HS-27a and HS-5 cells) or 3 µL (for WPMY-1, C4-2B, and LNCaP cells) of Lipofectamine was used for each well in six-well culture plates. For experiments using the perlecan promoter constructs, the control Renilla luciferase was driven by the thymidine kinase promoter in a pHRG-TK vector (Promega). For experiments using insertions of HSPG2 promoter elements into pGL4.74 with the HSV-TK promoter driving expression of the Renilla luciferase, firefly luciferase expression driven by the CMV promoter was used as the control luciferase.

Luciferase reporter assay

The Dual-Luciferase® Reporter Assay system (Promega) was used to measure luciferase activity in transiently transfected cells, according to the manufacturer’s instructions.

Quantitative PCR

Quantitative PCR was performed as previously described [Savore et al., 2005]. Reverse transcription reactions were performed using the QScript cDNA Supermix kit (Quanta Biosciences, Gaithersburg, MD). The final concentration of primers (sequences found in appendix A table I) was 200 nM (HSPG2) and 400 nM (ACTB). IQ SYBR Mix (Quanta Biosciences) was used according to the manufacturer’s instructions, to equal half of the final reaction volume. The thermal cycling program was as follows: 95°C for 3 min, then 40 cycles of the following two steps: 95°C for 30 seconds, 63°C for 45 seconds. Relative
transcript levels were determined by the Pfaffl method, using ACTB as a constitutively expressed gene control.

**Immunoblotting for secreted perlecan protein**

Dot blotting was performed to analyze accumulation of extracellular perlecan as previously described [Savore et al., 2005]. For dot blotting, equal volumes of medium were loaded into each well. Medium was diluted 1:10 in phosphate buffered saline (PBS), and 100 µL of the diluted solution were loaded into each well. Perlecan antibody A7L6 (Santa Cruz Biotechnology, Santa Cruz, CA) was diluted 1:5,000 in blocking buffer and incubated with the membrane overnight at 4°C while shaking.

**WST-1 assay for metabolically active cell content**

Assays were performed in parallel with conditioned medium collection. First, conditioned medium was removed for use in dot blotting for perlecan secretion. Serum- and phenol red-free medium was returned to the well, with 10% (v/v) WST-1 reagent (Roche) added. Cells were incubated for 1 hr at 37°C and then medium was transferred to a 96-well plate, where dual wavelength measurements were taken at 440 nm (measurement wavelength) and 610 nm (reference wavelength). Per the manufacturer’s instructions, only metabolically active cells will metabolize this salt at the mitochondrial membrane, resulting in a spectrophotometric shift in the medium. This allows determination of a ratio of metabolically viable cells during treatment compared to cells treated with the vehicle control.
**TNF-α enzyme-linked immunosorbent assay (ELISA)**

Cells were grown to 80% confluence in normal growth medium, then medium was removed and cells were prepared for medium collection during which time they typically reached confluency. Cells in 6-well culture plates continued to grow for 24 hrs in 1 mL serum-free medium, and this conditioned medium was collected and assayed for TNF-α production. At the same time, the cells were trypsinized from the plates, and counted using a hemocytometer. TNF-α levels were assayed and quantified using the human TNF-α ELISA kit (Thermo Scientific) according to the manufacturer’s instructions.

**Immunostaining for NF-κB**

Cells were grown in 8-well chamber slides (NUNC – Thermo Scientific) in normal growth medium and subsequently treated in serum-free medium supplemented with TNF-α (10 ng/mL) for 30 min at 37⁰. Cells were rinsed with PBS and fixed with paraformaldehyde. After rinsing, they were permeabilized with Triton X-100 (0.2% v/v) at room temperature for 10 min, blocked with 3% (w/v) BSA for 40 min at room temperature then incubated with anti-p65 (Santa Cruz Biotechnology – sc109) in a 1:200 dilution at 4⁰C overnight. Slides then were washed and incubated with a goat anti-rabbit Alexa-Fluor 488 secondary antibody (Invitrogen) in a 1:200 dilution, DAPI (Invitrogen) in a 1:500 dilution, and Phalloidin 633 (Invitrogen) at a 1:500 dilution for 40 min at room temperature. Slides were mounted in SlowFade Gold (Invitrogen) and imaged using a LSM 710 confocal microscope (Zeiss, Oberkochen, Germany).
Chromatin Immunoprecipitation (ChIP)

The Chromatrap® kit (Porvair Sciences – Leatherhead, Surrey, England) was used for ChIP assays. LNCaP cells were grown to ~70% confluency and serum-withdrawn overnight, then treated for 24 h with TNF-α (10 ng/mL) or vehicle control prior to collection of chromatin. Fixation, DNA shearing and ChIP were performed following the manufacturer’s instructions. Antibodies included the anti-p65 SC-109X antibody (Santa-Cruz), as well as the control antibodies anti-polII SC-9001 (Santa Cruz) and anti-H3K4me3 39159 (Active Motif – Carlsbad, CA). Primer sequences used to amplify various gene regulatory regions are included in appendix A table 1. The NFKBIA gene promoter served as a positive control for p65 immunoprecipitation, and the GAPDH promoter was used a positive control for RNA polymerase II immunoprecipitation. The HBB gene was the negative control for nonspecific pulldown by either the antibody or adsorption by the column beads.

Statistics

All reporter assays, quantitative PCR and dot blotting were analyzed using a one-way ANOVA followed by Tukey’s post-test. Tissue microarray data were analyzed by the Cedars Sinai statistics group funded through the NIH/NCI P01CA098912. Stain concentration was normalized using square root transformation. These values were analyzed by ANOVA and two-sample test.
**Results**

**Perlecan accumulates in desmoplastic stroma of PCa.** Figure 3.1A shows immunostaining for perlecan in paraffin-embedded sections of PCa (right panel) and normal adjacent tissue (left panel) from the same patient. In sections of normal tissue, perlecan is expressed diffusely in the stroma. In this section containing Gleason grade 4 cancer, stromal cells surround and infiltrate the tumor, a hallmark of reactive stroma. Perlecan immunostaining surrounding these stromal cells is intense, noted by the rich brown color, in contrast to what is seen in normal stroma. The increase in perlecan deposition in tumor stroma versus normal stroma was statistically significant in this
Figure 3.1 Perlecan accumulates in the reactive stroma of prostate cancer. Perlecan staining (brown) with nuclear counterstain (blue) in Gleason grade 4 prostate cancer (PCa, right panel) and normal adjacent tissue (Normal, left panel) from the same patient. Increased deposition of perlecan is observed in the regions surrounding PCa when compared to regions of normal morphology. Quantification of perlecan staining in the tissue microarray is shown in figure 1B. This microarray included 34 sections of normal prostate and 23 sections of prostate cancer. Perlecan staining in normal stroma was compared to perlecan staining in tumor stroma in each of these sections. Perlecan staining was significantly higher in tumor stroma than in normal stroma with a p value of 0.0018. Scale bar = 100 µm.
representative microarray (p=.0018, figure 3.1B) comprising tissue from fifty-seven patients. This finding adds perlecan to the growing list of ECM proteins deposited into the desmoplastic stroma formed at primary sites of prostate cancer.

**Perlecan secretion and HSPG2 mRNA expression increase in response to cytokine treatment.** Cells representing tumor and stromal compartments of the tumor microenvironment were treated with TNF-α and TGFβ, alone and in combination, and perlecan production and HSGP2 gene expression were assessed. Because of the size and heterogeneity of the perlecan proteoglycan (~900 kDa) western blots are very difficult to use for quantification of perlecan protein and dot blotting is the preferred method.

When cell populations of the prostate stromal cell line, WPMY-1, were treated with TNF-α and TGFβ (figure 3.2A), I observed a dramatic loss of viable cells compared to controls (see figure 2.3). For those surviving WPMY-1 cells, there was a clear increase in perlecan production on a per cell basis in response to TNF-α, but not TGFβ, treatment. In HS-27a bone marrow stromal cells, perlecan secretion was increased in response to TNF-α treatment alone, and even more so when cells received both TNF-α and TGFβ (figures 3.2A,B). Thus, for both cells resembling the stromal cells present in the prostate (WPMY-1) or bone marrow (HS-27a), TNF-α increased perlecan secretion within 48 hr of treatment.

During the course of these experiments, I observed that the stromal cell response to TNF-α in terms of perlecan protein production, mRNA levels and reporter activity diminished over time with increasing passage number, e.g. see response level in figure 3.4 vs response level in figure 3.6. These non-transformed stromal cells appear to change phenotypically
in cell culture during prolonged passage, making this response difficult to measure once outside the responsive window. For this reason, experiments were performed routinely with lower passage (<20) cells.

LNCaP prostate cancer cells increased perlecan protein secretion during 48 hr treatment with TNF-α (10 ng/mL - 588 pM) but not TGFβ (10 ng/mL - 781 pM) (figure 3.2A,B), concentrations consistent with those used in other studies [Iozzo et al., 1997; Schmidt et al., 2006; Yagi et al.]. Per cell secretion of perlecan by LNCaP cells was not increased further by TGFβ either alone or when TNF-α was present. LNCaP cells treated with TNF-α or TNF-α combined with TGFβ increased HSPG2 mRNA levels after 24 hr, which returned to basal levels by 72 hr (figure 3.2C). The addition of TGFβ magnified the mRNA response at 24 hr, but this effect was not seen at the level of absolute protein secretion at either 24 or 48 hr (figure 3.2D). As with the mRNA levels, protein secretion returned to basal levels by 72 hr. Thus for LNCaP cells, TNF-α is able to acutely and transiently increase perlecan production with increases in mRNA preceding changes at the protein level that are best seen within 48 hr of treatment.
Figure 3.2 Expression of *HSPG2* mRNA and perlecan increase in HS-27a and LNCaP cells during TGFβ1 and TNF-α treatment. Medium was collected and perlecan content assessed by quantitative dot blot as described in Materials and Methods (A). The fold change in perlecan secretion shown in (A) was divided by the live cell number to determine perlecan secretion/cell (B). A time-course treatment was performed using LNCaP cells and cells were collected every 24 hr for 72 hr. Steady state *HSPG2* mRNA levels were assessed (C) by qRT-PCR as described in Materials and Methods. Conditioned medium was collected from LNCaP cells over the same time periods, and protein assessed by dot blot as above (D). Data is presented as the ratio of densitometric value of treated medium over the value for vehicle controls set to 1.0 (A,B,D) or Delta delta Ct (C). Bars represent the mean +/- SEM, *, p<.05 vs control, **, p<.01 vs control, ***, p<.001 vs control. At least three biological replicates were performed for each condition.
The *HSPG2* promoter has conserved NFκB binding elements more than 2.4 kb upstream of the start site of transcription. *In silico* analyses of the *HSPG2* promoter revealed multiple potential transcription factor binding sequences within 2.6 kb upstream of the start site of transcription including those for NFκB and Smad3 (figure 3.3). This 2.6 kb of upstream promoter sequence plus 79 bp of the 5′ untranslated region of the *HSPG2* transcript were cloned into the reporter vector used for the majority of these experiments (figure 3.3A). Given the responses to TNF-α seen in a variety of cells, it was interesting to find two conserved putative binding sites for NFκB that were located more than 2.4 kb upstream of the transcriptional start site. These two consensus sequences are separated by 28 bp which were both deleted to produce the 2.4 construct (figure 3.3B). To determine if these sites are functional NFκB sites, each was mutated singly (figure 3.3C,D) then in tandem (figure 3.3E) to create three new promoter constructs, labeled 2.6mut1, 2.6mut2, and 2.6mut1,2. As expected, two conserved putative SMAD3 binding sites also were present, one site more than 1 kb upstream of the transcriptional start site and one site approximately 200 bp upstream of the transcriptional start site.
Figure 3.3 Conserved NFκB consensus elements are present in the HSPG2 promoter. The schematic depicts the HSPG2 promoter that was used in the luciferase promoter-reporter assays. The full-length reporter used in these studies (A) extends more than 2.6 kb upstream of the transcriptional start site, and includes two putative NFκB binding sites in the most distal 200 base pairs. The 2.4 kb deletion construct (B) was created by removing the distal 200 base pairs of the 2.6 kb promoter. Directed mutants of the putative NFκB sites include mutation of NFκB 1 (C), NFκB 2 (D) or both NFκB sites (E).
Bone marrow stromal, prostate stromal and prostate cancer cell lines secrete very low levels of TNF-\(\alpha\).

Although it is expected that the bulk of the TNF-\(\alpha\) in the tumor microenvironment is produced by infiltrating immune cells, I sought to measure the baseline levels of cytokine produced by the various cancer and stromal cells used in this study (see figure B.2). Over 24 hr, HS-27a, WPMY-1 and C4-2B cells each secreted TNF-\(\alpha\) at levels near 40 pg/mL (2.35 pM), a concentration near the lower limit of detection for the assay (figure B.2 A). LNCaP cells secreted slightly more at \(~60\) pg/mL (3.5 pM), and HS-5 cytokine-secretory cells secreted \(~70\) pg/mL TNF-\(\alpha\) (4.1 pM). Cells were counted in order to calculate per cell secretion of TNF-\(\alpha\), and each of these cell lines secreted less than 10 fg/mL TNF-\(\alpha\) per cell per day (figure B.2 B). These very low levels of TNF-\(\alpha\) secretion were well below those required to increase expression of perlecan. The 24 hr baseline levels of TGF\(\beta\) produced by these lines was previously published: prostate cancer cells (0.4–4.8 pM, 0.01-0.12 pg/mL); bone marrow stromal cells (2.5–7 pM, 0.06-0.175 pg/mL), or prostate fibroblasts (5 pM – 0.125 pg/mL) [Chung et al., 2009]. The endogenous production of TNF-\(\alpha\) and TGF\(\beta\) was more than 100-fold lower than the concentrations of these cytokines used in our routine treatments.

The \textit{HSPG2} promoter is activated by TNF-\(\alpha\) in a dose-dependent manner. In untreated cells, the 2.6 kb promoter construct shown in figure 3.3A demonstrated the highest level of activity in the prostate stromal cell line, WPMY-1, and the bone marrow stromal cell line, HS-27a, and the lowest activity in the prostate cancer cell line, LNCaP (figure 2.6). Responsiveness of the \textit{HSGP2} 2.6 kb promoter to TNF-\(\alpha\) treatment was measured by luciferase assay using HS-27a and HS-5 cells (figure 3.4A). These data revealed that
promoter activity steadily increased during treatment of HS-27a cells with up to 10 ng/mL (588 pM) TNF-α in HS-27a cells. At concentrations above 10 ng/mL, no further increase in promoter-reporter response was observed (data not shown). In contrast, HS-5 cells did not respond to TNF-α at any concentration, consistent with the lack of effect on perlecan mRNA and protein levels (data not shown). I used these cells as a negative control for NFκB activation in immunofluorescence assays and reporter assays. I also studied the response of the HSPG2 promoter to a variety of other growth factors and cytokines (IL-1β, insulin-like growth factor 1, β-2 microglobulin, receptor activator of NFκB ligand, interferon γ, and a synthetic Hedgehog pathway activator Hh-Ag1.5) as well as co-culture with other prostate cancer and stromal cell lines. None of these treatments or co-cultures provoked a significant perlecan response in any of the cells tested here (data not shown).

The HSPG2 promoter is active in stromal and prostate cancer cells. Next, TNF-α or TGFβ (each at 10 ng/mL) were added singly or in combination to relevant prostate and stromal cell lines transfected with the 2.6 kb HSPG2 promoter (figure 3.4B). TNF-α treatment increased HSPG2 promoter activity in HS-27a cells and in the WPMY-1 prostatic stromal cells, but not in the growth factor and cytokine producing bone marrow stromal cells, HS-5, used as a negative control. TNF-α addition to prostatic cancer cell lines, LNCaP and C4-2B, also increased reporter activity, but to a lesser extent than seen in the responsive stromal cell lines. TGFβ alone did not increase promoter activity in any of these cell lines to a statistically significant extent. When added in combination with TNF-α, TGFβ increased the HSPG2 promoter activity in the two responsive stromal cell lines, but not in the cancer cells. Interestingly, the stromal cell response to TNF-α increased
in the presence of TGFβ whereas prostate cancer cells showed no additional response to TGFβ.

**TNF-α treatment of HS-27a and WPMY-1, but not HS-5, stromal cell lines increases NFκB nuclear localization.** Given the differential responses of the 2.6 kb promoter in HS-27a, WPMY-1 and HS-5 stromal cells, I sought to determine if this was attributable to inherent cell-specific differences in transcriptional context for the promoter or to differences in the ability to respond to the cytokine itself. As seen in figure 3.5, cells in which the *HSPG2* promoter responded to TNF-α (HS-27a, WPMY-1) also showed nuclear translocation of NFκB in response to the cytokine. In contrast, HS-5 cells which did not show *HSPG2* promoter responsiveness to TNF-α also failed to show a nuclear translocation of NFκB. The antibody blocking peptide and secondary antibody-only background controls were negative for fluorescence, indicating that this signal is specific to p65 (figure B.3).
Figure 3.4 *HSPG2* promoter activity increases in response to TNF-α in a dose-dependent manner. Cells were transfected with the 2.6 kb reporter plasmid shown in 3A, and treated for 24 hr with TNF-α at the indicated final concentrations (A), before measuring luciferase activity in lysed cells as described in the text. Concentrations above 10 ng/ml did not further increase activity (not shown). For experiments shown in (B), cells were treated with TGFβ1 (10 ng/ml), TNF-α (10 ng/ml) or a combination of the two, then luciferase activity was measured as described in Materials and Methods. Data are reported as the fold change observed between cytokine treated and vehicle treated cells. Bars represent the mean +/- SEM. *, p<.05 vs .01, control, **, p<.001 vs .1, .01, control, ***, p<.001 vs 1, 0.1, 0.01, control. At least three biological replicates were performed for each condition.
Figure 3.5 TNF-α treatment triggers nuclear translocation of NFκB in HS-27a and WPMY-1, but not HS-5, cells. Cells were treated with TNF-α (10 ng/mL) in serum-free medium for 1 hr, permeabilized, and then stained with anti-NFκB as described in Materials and Methods. Cells were imaged by confocal microscopy. Red, NFκB. Green, phalloidin. Blue, DAPI. Scale bars, 50 µm.
**HSPG2** promoter responsiveness to TNF-α resides more than 2.4 kb upstream of the start site of transcription in the more distal NFκB binding site. The constructs shown in figure 3.3B were used to dissect the nature of the *cis* element(s) in the HSPG2 promoter that are responsible for the TNF-α response. As shown in figure 3.6, neither treatment with TNF-α nor TGFβ nor a combination of the two activated the 2.4 kb promoter-reporter in any of the four cell lines that demonstrated responsiveness of the 2.6 kb promoter (figure 3.6A). In these experiments, the full-length promoter construct responded to TNF-α similarly to that seen in figure 3.4 (data not shown).

To identify the functional *cis* element(s) in the distal promoter region, the three mutant constructs shown in figure 3.3 (2.6mut1, 2.6mut2, 2.6mut1,2) were transfected into the same four cell lines and responses to cytokines measured. In both stromal cell lines, mutation of NFκB site 1 (2.6mut1) or mutation of both NFκB sites completely abrogated the response to TNF-α (figure 3.6B). In contrast, mutating NFκB site 2 alone (2.6mut2) did not decrease the promoter response, and actually increased the promoter response over that of the wild-type promoter in each of the cell lines (figure 3.6B). The mutation of both NFκB sites (2.6mut1,2) decreased responsiveness to TNF-α in the prostate cancer cell lines LNCaP and C4-2B but mutation of NFκB site 1 alone (NFκBmut1).
Figure 3.6 Deletion of the distal 200 base pairs of the \textit{HSPG2} promoter or mutation of NFκB consensus elements eliminates responsiveness to TNF-α and TGFβ1. Cells in culture were transfected with the 2.4 kb promoter seen in Figure 3B, or constructs in which the two NFκB sites were mutated as described above and responses to each cytokine alone (10 ng/ml) or in combination (each at 10 ng/ml) were measured at 24 hr of treatment. Panel A shows that removal of the entire region containing both sites eliminates responses to cytokines in all four cell lines shown. Panel B shows that \textit{2.6\textsubscript{mut1}} or \textit{2.6\textsubscript{mut1,2}}, but not \textit{2.6\textsubscript{mut2}}, loses activity, hence response to TNF-α is associated with the more distal of the two putative NFκB sites. Bars represent the mean +/- SEM. **, p<.01 vs wild-type, ***, p<.001 vs wild type. At least three biological replicates were performed for each condition.
The NFκB HSPG2 regulatory region is bound by p65.

Chromatin immunoprecipitation (ChIP) was performed to validate binding of NFκB to the elements identified above. LNCaP cells were treated with TNF-α for 24 h. After immunoprecipitating with anti-p65 antibody, the HSPG2 promoter was significantly enriched compared to either the internal background control (HBB gene promoter) or relative to cells treated with vehicle alone. Representative endpoint PCR is shown in figure 3.7A and quantitative PCR is shown in figure 3.7B. Quantitative PCR data is presented as a percentage of the input chromatin control. The column binding control mock immunoprecipitation and no template controls are shown in the replicate of endpoint PCR, and demonstrate very little amplification. Chromatin from two endometrial cancer cell lines were used as positive controls for NFκB activation, and the HSPG2 regulatory region was immunoprecipitated in these samples as well. Collectively these data demonstrate that NFκB binds to the same region identified in the reporter assays described above.
**Figure 3.7 p65 binds the HSPG2 5′ regulatory region in LNCaP cells during TNF-α treatment.** Chromatin immunoprecipitation was used to detect p65 binding the perlecan promoter. The HSPG2 regulatory region is significantly enriched compared to the background control (the HBB promoter) in LNCaP cells treated with TNF-α. Endpoint PCR (panel A) indicates the HSPG2 regulatory region is amplified when p65 is pulled down in the chromatin of TNF-α treated cells, but not in that of cells treated with the vehicle control. This amplification was quantified using quantitative PCR. The HSPG2 regulatory region was also pulled down in Ishikawa and Hec-50 endometrial cancer cells, used as positive controls for NFκB activation. Three biological replicates were performed using LNCaP cells, and one replicate was used for each of the endometrial cancer cell lines.
**Discussion**

Metastatic tumors may be perceived by the host defense system in similar ways as certain pathogens or foreign bodies that result in encapsulation of the tumor with dense matrix that resists tissue penetration. The large proteoglycan, perlecan, is a key component of this host defense, and has evolved a unique multi-domain structure that allows it to perform a variety of functions in the desmoplastic stroma [Farach-Carson et al., 2013; Sabit et al., 2001; Terranova et al., 1986]. Here I examined the actions of cytokines present in the tumor microenvironment that are most likely to contribute to perlecan production. I found that prostate cancer cells, prostate stromal cells, and a structural subpopulation of bone marrow stromal cells are programmed to increase *HSPG2* mRNA expression and perlecan production in response to pro-inflammatory cytokine stimulus. While TGFβ, a key factor in connective tissue growth, plays a major role in the stromal response to cancer and induction of fibrosis [Ronnov-Jessen and Petersen, 1993], our work suggests that TNF-α is the primary effector of increased *HSPG2* expression in the invasive prostate cancer microenvironment both in the prostate and at sites of metastasis.
Figure 3.8 Model for *HSPG2* induction by tumor and immune cell interaction. In our proposed model, tumor recruitment of immune cells including macrophages results in inflammatory cytokine release. This amongst other cancer cell-derived factors induces stromal activation. Part of this phenotypic program is perlecan secretion.
While resting stroma has low levels of perlecan, tumor stroma shows high levels of perlecan deposition. In the context of the bone microenvironment, it is interesting that the stromal cell line with a phenotype supporting the structure of the bone marrow microenvironment, represented by HS-27a cells, rather than the cell line with a phenotype supporting hematopoietic progenitor cells, represented by HS-5 cells, were the ones most responsive to TNF-α induction of \textit{HSPG2}. Bone marrow stromal cells with the structural phenotype (HS-27a) also are resistant to TNF-α induced cell death (figure 2.3), in contrast to the secretory bone marrow stromal cell line HS-5 [Byun et al., 2005]. This finding is consistent with these cells being the primary producers of perlecan-rich matrix at sites of bone metastases. Additional perlecan deposition may be provided by the cancer cells themselves, especially if they have undergone EMT. This latter notion is supported by our findings that the 2.6 kb perlecan promoter is five times more active (at basal levels compared to internal controls) in C4-2B cells derived from bone-metastasis than in the parental LNCaP line, despite the fact that basal \textit{HSPG2} mRNA levels are similar in these cell lines [Savore et al., 2005].

The highest basal activity of the 2.6 kb promoter was found in the two stromal lines, WPMY-1 and HS-27a. Deletion of the distal 200 base pairs of the promoter decreased baseline reporter activity in all cell lines (ten-fold reduction in stromal cells; five-fold reduction in prostate cancer cells), highlighting the importance of this previously unexplored region in regulation of \textit{HSPG2} transcription. The upstream 2.5 kb of the \textit{HSPG2} promoter was examined previously and determined to contain all of the elements of the core promoter region [Iozzo et al., 1997]. Since this early report, the human genome
has been better annotated and it now made sense to reexamine the promoter using currently available information and tools. Using the ENCODE data track available on the UCSC genome browser (genome.UCSC.edu) I was able to make new inferences concerning the location of the HSPG2 promoter (figure B.1). Using the H3K27Ac mark track (part of the ENCODE data set), it is evident that there are activated histones (associated with activated enhancer elements) located throughout the 40 kilobase intron 1 of the human HSPG2 gene, as well as up to 7 kilobases upstream of the start site of transcription, well beyond the core promoter region denoted by the H3K4me3 histone mark. This suggests that complete regulation of the HSPG2 gene cannot be fully understood until transcriptional regulatory activities in this entire region are appreciated in cell-specific context. It also may explain why addition of the distal 200 base pairs to the promoter construct used in this study increased baseline reporter activity so greatly: important regulatory sequences exist in this heretofore unstudied upstream region.

Along with other inflammatory cytokines, TGFβ and TNF-α both have been implicated in modulation of stromal cell activity at tumor sites and in other pathologies [Rodriguez Perez et al., 2011; Ronnov-Jessen and Petersen, 1993]. While TGFβ is produced locally by stromal cells and other connective tissue cells, TNF-α is secreted by nearby tumor-associated macrophages that appear in the tumor microenvironment in response to tissue stress from the perceived wound. The HSPG2 gene promoter regulating perlecan expression responds to TGFβ [Iozzo et al., 1997], and I expect that local secretion of TGFβ maintains a basal level of perlecan production in stromal and connective tissue cells through an autocrine signaling loop.
NFκB activity plays a role in prostate cancer progression and formation of reactive stroma, but direct effects of this transcription factor on perlecan production at the level of the promoter have never been reported. Here I showed that a consensus NFκB binding sequence in the HSPG2 promoter is a key driver of the TNF-α response in bone marrow stromal and prostate cancer cells. I also showed that in stromal cell lines that respond to TNF-α treatment by increasing perlecan expression, NFκB localizes to the nucleus, whereas in the unresponsive stromal cell line, HS-5, there was no nuclear translocation of NFκB in response to TNF-α treatment. While the reason for the lack of TNF-α responsiveness in the HS-5 cell line is unclear, it nonetheless accounts for the lack of effect on perlecan expression in these cells.

It is of note that the per-cell increase in perlecan secretion in the WPMY-1 cells that survived treatment with TNF-α was much greater than the increase in perlecan secretion in the whole population which was dying off throughout the experiment (figure 3.2 panels A & B). This distinction may have important implications in the formation of reactive stroma during cancer invasion. It is appreciated that a subpopulation of stromal fibroblasts reacting to wounding can differentiate into myofibroblasts, which increases their desmoplastic ECM deposition profile [Tuxhorn et al., 2002]. Future work might establish whether the subset of stromal cells in the prostate producing large amounts of perlecan are indeed myofibroblasts.

Figure 3.8 shows a model that explains the up-regulation of HSPG2 in the stromal environment. Our data indicate that none of these stromal or cancer cell types can be the source of TNF-α in the microenvironment. Tumor associated macrophages are known to secrete inflammatory cytokines such as TNF-α, and are integral components of tumor
progression. This activity in concert with stromal cell autocrine TGFβ signaling could activate stromal cells to secrete perlecan and perform the other functions of cancer-associated fibroblasts. This model may further inform strategies for targeting reactive stroma in prostate cancer.
Chapter 4

As prepared for submission

Evolution of the *perlecan/HSPG2* gene and its activation in tissue regeneration
Abstract

The Heparan sulfate proteoglycan 2 (HSPG2)/perlecan gene is ancient and conserved in all triploblastic species. Its functions are crucial to the integrity of cell boundaries in tissue and its large (~900 kDa) modular structure has prompted speculation about the evolutionary origin of the gene. The gene’s conservation amongst basal metazoans is unclear. After the recent sequencing of their genomes the basal metazoan organisms Nemastella vectensis (the cnidarian) and Trichoplax adhaerens (the placozoan) have respectively become favorite models of tissue regeneration and multicellularity. More ancient basal metazoan phyla include the poriferan and ctenophores, whose evolutionary relationship has recently been clarified. Our data indicate that the HSPG2 gene is conserved in both the placozoan and cnidarian genomes, but not in those of the ctenophores or poriferans. The ancestral form of the gene in T. adhaerens is encoded as two separate but genetically juxtaposed genes which house all of the constituent pieces of the mammalian HSPG2 gene. These genetic constituents are found in separate genes of the poriferan Amphimedon queenslandica, indicating a possible intronic recombinatory mechanism for assembly of the HSPG2 gene. Expression of the gene was activated during tissue regeneration of N. vectensis and normally in the polarized epithelium of the body column. These data indicate that the HSPG2 gene evolved concurrently with the development of differentiated cell types separated by acellular matrices.
**Introduction**

The mammalian *HSPG2/perlecan* gene product has a dual functional role in maintaining tissue boundaries and providing a growth factor depot for wound healing [Farach-Carson et al., 2013]. The modern gene *HSPG2* encodes the perlecan proteoglycan, an extracellular heparan sulfate proteoglycan of high (~600 kDa) molecular weight and composed of 48 modular units sharing homology with other ECM proteins [Murdoch et al., 1992]. Remarkably, the core protein is subject to minimal alternative splicing in humans, despite having some 50 introns [Lord et al., 2013]. The human core protein consists of 4,370 amino acids and is modified with 3-4 heparan or chondroitin sulfate chains before secretion into the extracellular space, requiring a hefty metabolic investment by the cell manufacturing it. Perlecan performs a variety of signaling, adhesive and extracellular scaffolding roles [Farach-Carson and Carson, 2007; Knox and Whitelock, 2006; Melrose et al., 2008], such that mutation of the gene has pleiotropic effects; loss of the perlecan gene is lethal in humans and mice [Arikawa-Hirasawa et al., 2001]. The origin of perlecan is unknown, but the complexity of this gene begs such study. *HSPG2* orthologues have been studied in vertebrates, arthropods, and nematodes [Datta and Kankel, 1992; Hummel et al., 2004; Rogalski et al., 1993]. Mutations of nematode *HSPG2* (*unc-52*) are pleiotropic, although the gene was first recognized for its maintenance of muscle function [Rogalski et al., 1995], which provides clues to the function of the perlecan protein in its older incarnations. A goal of this study was to use the latest publicly available genomes of the basal metazoans *Amphimedon queenslandica* [Srivastava et al., 2010], *Trichoplax adhaerens* [Srivastava et al., 2008] and the cnidarian *Nematostella vectensis* [Putnam et al., 2007] as well as the ctenophore *Mnemiopsis leidyi* [Ryan et al., 2013] and the
choanoflagellate *Monosiga brevicollis* [King et al., 2008] to ascertain the oldest living member of the family of *HSPG2* orthologue-encoding metazoans, and to infer the evolutionary events that occurred to build this large highly conserved modern proteoglycan.

The multicellular animals most closely related to unicellular organisms are sponges and ctenophores [Ryan et al., 2013]. It is commonly thought that the placozoan, *Trichoplax adhaerens* was derived from a common organism with sponges [Srivastava et al., 2008]. These animals are simpler than sponges in that they have fewer cell types and are very small (1-3mm in diameter) [Grell and Benwitz, 1974]. They are by many measures closer relatives of cnidarians and bilaterians than are sponges, but this evolutionary relationship is by no means decided upon [Dohrmann and Worheide, 2013]. Choanoflagellates are the sister group to the metazoans, and the recently sequenced genome of the choanoflagellate *M. brevicollis* encodes few ECM-like proteins, but does contain genes for membrane receptors [Carr et al., 2008; King et al., 2008]. Given the uncertainties in understanding the evolutionary relationships of these primitive species, I hypothesized that the study of assembly of the complex *HSPG2* gene orthologues might offer unique insights into species relationships and early evolution of multicellularity.

The cnidarian *N. vectensis* is used as a model of tissue regeneration and of diploblastic development [Passamaneck and Martindale, 2012]. Its ease of cultivation in the lab and small size make it an ideal model system. This radially symmetrical creature can completely regenerate its oral end over two weeks after lateral bisection, providing an opportunity for the visualization and molecular characterization of the regenerative process [Reitzel, 2006]. *HSPG2* and orthologues are activated during wound healing of humans,
as part of the reorganized ECM in the various stages of wound reconstruction [Zhou et al., 2004], a process that involves both the core protein and the heparan sulfate cargo of heparin binding growth factors [Saksela et al., 1988]. The cnidarian epithelium secretes a basal lamina ECM that separates epithelial and connective tissues, as evidenced by transmission electron microscopy [Tucker et al., 2011]. To examine the early origins of perlecan in wound healing, I examined the presence of the perlecan orthologues in this tissue, specifically its appearance during regeneration of oral structures.

In this study, I present evidence for the evolutionary precursors to an HSPG2 orthologue perl in A. queenslandica, as well as sequence for T. adhaerens perl, which is encoded as two separate genes in the same region of the placozoan genome, and using this information suggest an evolutionary relationship in early multicellular organisms. I speculate on evolutionary pressures leading to the evolution of the perl gene and demonstrate its presence during the regeneration of N. vectensis oral structures.

**Methods**

**Perl gene Prediction**

Genomic scaffold and transcript sequences were downloaded from the Joint Genome Institute (http://www.jgi.doe.gov/) [King et al., 2008; Putnam et al., 2007; Srivastava et al., 2008; Srivastava et al., 2010]. tBLASTn with a low complexity filter, BLOSUM62 matrix, and conditional compositional score matrix adjustment was used to compare human perlecan peptide sequences to transcript sequences of the experimental animal. This identified genomic regions that had high similarity to some portion of the human HSPG2 gene. Many BLAST hits occurred on the same scaffold. In this case, independently folding
protein modules from the human perlecan peptide sequences were aligned (using a BLOSUM62 cost matrix) with all translated frames of this scaffold. *Trichoplax adhaerens* and *Nematostella vectensis* genomic regions that aligned with a >20% peptide homology with human Ig modules or >30% peptide homology with all other human perlecan modules were inputted into Genewise WISE2 (http://www.ebi.ac.uk/Tools/psa/genewise/) [Birney et al., 2004] to predict exon-intron boundaries of this putative *perl* gene. The nucleotide sequence of the entire identified genomic region was used as the DNA input for Genewise WISE2, and the concatenated peptide sequence of the regions that aligned with sufficiently high homology to the human perlecan modules was used as the protein input for Genewise WISE2. All tBLASTn and alignments were performed using the Geneious software package 5.4.6 (Biomatters - Auckland, New Zealand).

**Trichoplax adhaerens Culture**

*T. adhaerens* were fed on *Porphyridium* algae (Carolina Biological Supply – Burlington, NC) maintained in Artificial Seawater (ASW) – formulated from 35g of Reef Crystals Reef Salt (Foster and Smith, Rhinelander, WI) in 1L of ddH20. This ASW was carefully maintained at 32 parts per trillion, and supplemented with 1mL Micro Algae Grow (Florida Aqua Farms Inc, - City, St) to make seawater medium. *Porphyridium* subcultures were established in glass petri dishes one week prior to seeding with *T. adhaerens* organisms. Petri dishes with established *T. adhaerens* cultures had fresh 32ppt seawater medium added every week. New *T. adhaerens* cultures were established every week. *Porphyridium* monocultures were filtered every two weeks with Whatman #1 filter paper (GE Healthcare, Little Chatford, United Kingdom) and had fresh seawater medium added weekly. Ambient room temperature used for culturing the algae and placozoans was 22°C.
**Nematostella vectensis Culture**

*N. vectensis* were purchased from the Marine Biological Laboratory (Woods Hole, MA). The animals were maintained in ASW diluted 1:3 in ddH$_2$O. They were fed *Artemia nauplii* (Aqua Medic, Loveland, CO) and their 1:3 ASW was refreshed at least once weekly. *Artemia* were hatched in aerated 1:3 ASW for 48 h under a heat lamp. Animals were clonally propagated by lateral bisection. Animals were starved for three days, and immobilized in 7% MgCl$_2$ for 5 min prior to bisection with a one-sided razor near the physal end of the organism, and regeneration assays were performed according to the published standard [Bossert et al., 2013].

**Ribonucleic Acid Isolation**

RNA was obtained by starving an established *T. adhaerens* culture overnight in Artificial Seawater (ASW) to remove the majority of the *Porphyridium*. These dishes were rinsed the following day in ASW, and TRIzol® reagent (Life Technologies, Carlsbad, CA) was added directly to the dish. The animals were then removed using a cell scraper and RNA was extracted following the manufacturer’s instructions. *N. vectensis* were starved for 3 days prior to RNA isolation. Animals were extracted from their dish, rinsed with 1/3 ASW and immersed in 1 mL TRIzol® reagent. Animals were disrupted using a tissue homogenizer and RNA was extracted as published previously [Stefanik et al., 2013b].

**5’ Rapid Amplification of cDNA Ends (RACE)**

Primers for RT-PCR and RACE PCR were generated from alignment matches between the genomic sequences of the experimental animal and human perlecan modules. RACE PCR
reactions were conducted with the 5’ RACE System (Life Technologies - Carlsbad, CA) according to the manufacturer’s instructions. PCR products were visualized and purified using agarose gel electrophoresis and then cloned using a TOPO cloning kit from Invitrogen and One Shot Top 10 chemically competent E. coli from Invitrogen. Sequencing was done through Lone Star Labs.

**Phylogenetic Analysis of Perl**

The protocol set forth in [Hall, 2013] was followed for the creation of a phylogenetic tree using MEGA6 [Tamura et al., 2013]. The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model [Jones et al., 1992]. The tree with the highest log likelihood (-94037.8169) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using a JTT model. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 2.3678)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 3.3210% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 76 amino acid sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and
ambiguous bases were allowed at any position. There were a total of 1959 positions in the final dataset. The figure was assembled using the Figtree phylogenetic tree production program (http://tree.bio.ed.ac.uk/software/).

**Whole-mount In Situ Hybridization of T. adhaerens**

Trichoplax were allowed to adhere to superfrost slides (Thermo Fisher) for two h, and then fixed in Ladowsky fixative (50% EtOH (v/v), 3.7% formaldehyde (w/v), 4% (v/v) glacial acetic acid overnight at -20°C. Slides were washed two times for five min in PBT and treated with proteinase K (10 µg/mL) in PBS for two minutes at 37°C in a humidified chamber. Slides were then washed in PBS + .1% (w/v) Tween 20 with no agitation, and post-fixed in 4% (w/v) paraformaldehyde, 2% (w/v) glutaraldehyde for 1 h at room temperature in a humidified chamber. Tissue was then dehydrated in an ascending concentration of methanol in PBT, and stored overnight at -20°C. Tissue was rehydrated in PBT and washed in 50% hybridization buffer 50% PBT at 55°C. In-situ hybridization and riboprobe preparation were the same as above in *N. vectensis in situ* hybridization, excepting that tissue was treated with 20 µg/mL RNAse A in 2X SSC, 0.1% Tween 20 (v/v) for 30 min at 37°C in a humidified chamber after hybridization.

**In Situ Hybridization of Sectioned N. vectensis**

*N. vectensis* were starved for three days prior to fixation. In the case of regeneration studies, animals were starved three days prior to bisection, and bisection was performed as described above. Animals were immobilized in 7% (w/v) MgCl₂ in 1/3 ASW for five min. Animals were fixed in 10 mL of the following solution: 60% EtOH, 30% 37% formaldehyde, and 10% glacial acetic acid overnight at 4°C on an orbital shaker, and
embedded in paraffin. Animals were oriented in sectioned longitudinally. *In situ* hybridization was performed as previously described [Etchevers et al., 2001]. DIG-labeled Riboprobes (see appendix A) were synthesized using the T7 Megascript kit using probe sequence inserted into a pCR®4 vector as template (Life Technologies).
Results

HSPG2 is conserved in N. vectensis

The perl gene (an HSPG2 homologue) was localized in the genome of the cnidarian *N. vectensis* using tBLASTn-alignment (figure 4.1). The gene covers the 3’ end of scaffold 182 and the 5’ end of scaffold 196. This suggested that these two scaffolds are connected by a gap in the sequencing data. Sequences with homology to LDL receptor, Immunoglobulin, laminin B and G chains, and EGF were found in an order identical to that of domains II, III, IV, and V of perlecan in bilaterians. No SEA (domain I) homologue or O-glycosylation sites were found upstream of *N. vectensis* perl domain II. This finding necessitated the search for an HSPG2 homologue in the genome of the simpler and more ancient species *T. adhaerens*.

perl is conserved in T. adhaerens

The perl gene was found in the placozoan *T. adhaerens* using tBLASTn-alignment (figure 4.1). The gene is encoded on scaffold 6 of the current genome assembly. All five protein domains are included in the *T. adhaerens* perl gene, in an order identical to that of HSPG2 in humans. The *T. adhaerens* domain I includes a Sperm protein, Enterokinase, and Agrin (SEA) module which was until now thought to be found exclusively in perlecan of mammals and birds. The *T. adhaerens* perl SEA module includes the G-SVV motif so critical to the autocleavage of these modules in other proteins. Whereas the SEA module of perlecan in mammals and birds does not contain this autocleavage motif, it is intriguing to speculate that this ancestral perlecan molecule may have undergone autocleavage while in the secretory pathway. *T. adhaerens* perl domain II is structured exactly as that of
human \textit{HSPG2}, with three LDL receptor-like repeats followed by one immunoglobulin (Ig) domain. Domain III of \textit{T. adhaerens} perl includes one additional set of Laminin B- and Laminin EGF-like modules when compared to human perlecan, but is very similar. Domain IV of \textit{T. adhaerens} perl encodes roughly 36 total Ig modules, making it quite a bit larger than the domain IV of human perlecan which only includes 21 Ig modules. Domain V of \textit{T. adhaerens} perl encodes both the laminin G and EGF-like modules similar to other perlecan proteins.

\textit{T. adhaerens perl is split into two separately transcribed genes.}

5’ RACE PCR reveals that \textit{T. adhaerens perl} is transcribed as two separate genes (figure 4.1), which although located only thirteen kilobase pairs apart on the chromosome, have maintained separate promoters in this organism. There are two start sites of transcription for both \textit{T. adhaerens perl 1} and 2 revealed by 5’ RACE PCR. All efforts to connect these two genes by PCR-based methods were fruitless, indicating that these two products are indeed transcribed separately. This may indicate that the two separate \textit{N. vectensis} transcripts corresponding to perlecan may in fact be separate genes as I was also unable to connect those transcripts by PCR-based methods.

\textit{Perl is not conserved in Amphimedon Queenslandica, Mnemiopsis Leidyi or Monosiga Brevicollis}

\textit{Perlecan}-like motifs were identified by tBLASTn-alignment in \textit{A. queenslandica} (figure 4.1). There are genes encoding numerous IgG repeats, as well as multiple LDLR modules and laminin motifs. These sequences are often clustered together in groups of multiple repeats. The regions encoding the would-be domains of perlecan are encoded on separate
scaffolds. This means that at the very least they are far apart (at minimum 2,500 kilobase pairs), if not on different chromosomes. Without better coverage or direct RNA-seq, it is impossible to state with certainty that these genes are not transcribed together, but our best efforts have not convincingly demonstrated that these motifs are encoded together. The starting point for intronic recombination-based formation of a unified perl gene appears to the region rich in Ig modules, where there is one Ig encoded 44 kb upstream of a region of 37 predicted Ig modules in close proximity. This spacing is similar to that of unified perl in all younger organisms including T. adhaerens perl, where domain II houses one Ig module, and then 12 kb downstream are encoded 38 Ig modules.

Very few perlecan-like motifs were identified by tBLASTn-alignment in the transcriptome of the ctenophore Mnemiopsis leidyi or the genome of the unicellular choanoflagellate Monosiga brevicollis. As suggested by recent publications, M. brevicollis encodes few ECM gene homologues making this finding unsurprising. Our analysis of the M. brevicollis genome found only one perlecan-like IgG module, three LDLR modules, and four laminin EGF-like modules on different scaffolds and hundreds of thousands of base pairs apart. There is no single gene encoding a perlecan-like protein in M. brevicollis. M. Leidyi is a part of the ctenophore family which, as recently described [Ryan et al., 2013], evolved quite separately from the rest of the metazoan tree. The ctenophore transcriptome is well-covered by publicly accessible RNA-sequencing data. Sequences encoding the human perlecan protein modules BLASTed to 132 separate transcripts, but most of these only included one type of folding module. Some hits included multiple folding motifs as shown in figure 4.1, but none of these transcripts encode perlecan homologues. Homologues for several laminin chains, fibulin, Notch, protocadherin, attractin and
neurexin were found encoding multiple types of perlecan-like modules. This many hits to the various folding modules of the perlecan protein suggests that although the Ctenophores evolved separately from the sponge-eumetazoan branch of the animal kingdom, extracellular protein motifs had already multiplied in the genome at the time of their speciation.
Figure 4.1 Perlecan orthologues or constituents in model organisms. The presumptive pre-perlecan constituents of *A. queenslandica* (A) and perlecan proteins of *T. adhaerens*, *N. vectensis*, *C. elegans*, and *Homo sapiens* are shown in schematic form to demonstrate structural differences. Several of the scarce matches for perlecan modules in the *M. brevicollis* and *A. queenslandica* genome are shown in schematic form. The number of unique matches for individual or paired perlecan folding modules in the *M. leidyi* transcriptome are shown in subscript next to each schematic module. The dashed line between *T. adhaerens perl 1* and *T. adhaerens perl 2* indicates that these two genes, although only thirteen kb apart on the chromosome, are separately transcribed. *N. vectensis perlecan* is split by a sequencing gap (dashed line), and runs to the 3’ terminus of scaffold 182, and the 5’ end of scaffold 196. Blue rectangles represent parts of the transcript amplified by RACE PCR and sequenced, all other parts of the protein are predicted by tBLASTn-alignment. Schematics are not to scale, gene lengths are denoted at the right end of each diagram.
**HSPG2/Perlecan and the evolution of basal metazoa**

Figure 4.2 shows a phylogenetic tree produced using the peptide sequence of 76 perlecan homologues to show that these peptide sequences support the conventional evolutionary paradigm, with *T. adhaerens* being more ancestral than *N. vectensis*, and derived from some other common ancestor. This supports the general, albeit contested, view that *T. adhaerens* evolved prior to cnidarians and after sponges. One hundred bootstrapping generations were performed to test the results of our analysis, and the node separating *N. vectensis* and *T. adhaerens* from other invertebrates crossed the generally accepted threshold for evaluation by bootstrapping (0.7). As additional evidence of the accuracy of our analysis, the relationships amongst other perlecan peptide sequences displayed here follow the general evolutionary paradigm for the animal kingdom (shown in brackets). Peculiar though is the presence of the SEA module in domain I of *T. adhaerens perl* (figure 4.1). In this way and in the structure of domain II, *T. adhaerens perl* 1 is actually structurally more similar to mammalian perlecan than to that of *D. melanogaster* or *C. elegans*. This reflects the excess of LDL receptor modules in *D. melanogaster* perlecan (known as terribly reduced optic lobes) addition of and Ig module to *c. elegans* unc-52 and the lack of a defined SEA in both organisms. PFAM analysis indicates that the domain I of *D. melanogaster* and nematode perlecans encodes a module of insignificant similiarity to the SEA fold.
Figure 4.2 Molecular phylogenetic analysis of 76 perlecan homologues. MUSCLE alignment followed by maximum likelihood modeling resulted in this phylogeny of perlecan in 76 representative organisms. Bootstrapping values are demonstrated by a color gradient on branches, denoted in the legend. Species segregated along accepted evolutionary lines. The *N. vectensis* and *T. adhaerens* node (bootstrapping value .76) indicates more evolutionary time along the *T. adhaerens* (red box) branch than along that of the *N. vectensis* branch.
Figure 4.3 \textit{T. adhaerens perl 1 expression in the medial cell layer}. \textit{In situ} hybridization (A, blue) was used to characterize the expression pattern of \textit{perl 1} of \textit{T. adhaerens} in whole mount. Fluorescently-labeled Phalloidin (B – green) was used to detect filamentous actin, and DAPI (B- blue) demonstrates the location of nuclei in the animal in this z-stack projection. An acellular layer (B - white arrowheads) rings the outer layer of the animal. Algae stain indiscriminately in both assays, denoted by red arrows in A and B. Scale bar, 20 μm.
**T. Adhaerens perl 1 is expressed in the medial cells**

*In-situ* hybridization of *T. adhaerens perl 1* generates a staining pattern in the medial cell layer ringing the outer edge of the animal. A projection created by DIC microscopy shows no expression in the lower epithelium, fiber syncytium or upper epithelium (figure 4.3A). Algae stain indiscriminately, which makes recognizing the expression pattern somewhat more difficult. The medial cell layer expressing perlecan appears to be bordered by a transient acellular layer which forms between the medial cells and the rest of the body. Phalloidin and DAPI signals identify both filamentous actin and nuclei in the animal, and in many cases an acellular layer is clearly visible (figure 4.3B). Phalloidin signal (filamentous actin) is only detected in the lower epithelium, potentially because focal adhesions are formed between the glass slide and the cells of that layer.

**N. vectensis perl is expressed in the ectoderm of the scapus and the ectoderm of the mesentery, but not in the ectoderm of the tentacles.**

*In-situ* hybridization of *N. vectensis perl* reveals that the transcript is expressed throughout the ectoderm in of the scapus and physus, as well as in select cells of the mesentery (figure 4.4). Although there is a basal lamina of the gastroderm, there is little staining for the perlecan transcript in the gastroderm of the scapus. The acellular mesoglea is evident in high magnification images of the scapus (figure 4.4B), where it appears that the ectoderm expressed the *perl* transcript. The *perl* transcript signal in the mesentery follows no detectable pattern although it appears to be a part of the expression profile of distinct cells (figure 4.4C). There is little staining in the outer cell layer of the tentacles, but is expressed by select cells of an interior cell layer (figure 4.4 D,E).
Figure 4.4 *Perl* is expressed in the ectoderm of the scapus, irregularly in the mesentery, and irregularly in the inner cell layer of the tentacles of *N. vectensis*. An adult cnidarian (A) was tested for *perl* mRNA expression. The ectoderm of the body wall has consistent *perl* mRNA signal in the ectoderm, whereas the gastroderm is mostly negative for *perl* mRNA. The acellular mesoglea is clearly visible (black arrowhead, B). The mesentery expresses *perl* mRNA in a sporadic fashion, with pigmented cells largely not expressing *perl* (C). Tentacles do not express *perl* in the ectoderm but do sporadically in the inner cell layer (D,E). Scale bars, A – 0.5mm, B-E – 40 μm.
During *N. vectensis* regeneration, *perl* expression is activated at stage 2.5 in the regenerating pharynx and tentacles.

*Perl* mRNA expression during *N. vectensis* regeneration was studied using *in-situ* hybridization (figure 4.5). Organisms were fixed and sectioned at stages one through five of the regenerative process. At stages 1 and 2, *N. vectensis* *perl* mRNA is expressed much as it is in the intact organism. There is copious mRNA signal in the ectoderm of the physus and scapus, as well as individual cells in the physal end of the mesentery left over from bisection of the animal. At stage 2, elongation and compression of the physus and would-be scapus are evident, as is the “pumpkin head” phenotype of a slight depression at the oral end of the organism. At stage 2.5-3, activated tissue is evident as the oral end of the organism, and *de novo* *N. vectensis* *perl* mRNA signal is visible. Tentacle bud cells transcribe *N. vectensis* *perl* mRNA at this stage. At stage 4, tentacles are elongating and, in contrast to stage 5, or intact tissue, *perl* mRNA is resident in the tentacle ectoderm. By stage 5, this expression has receded, and tentacles and pharynx are free of *N. vectensis* *perl* mRNA expression.
Figure 4.5 Perl is expressed in the regenerating oral region and tentacles at stage 2.5 through stage 4-5. Animals were bisected midway through the mesentery region (A – adapted from Stefanik et al. 2013). At stage 1 and early stage 2 (C-D), perl transcript is found in the ectoderm of the body wall, as would be expected. At this stage, the wound has simply closed, and no oral structures are being regenerated. Activated tissue is visible in the stage 2.5-3 in the tentacle bud (E- black arrow) as well as in stage 4-5 (F – black arrowhead) animal in the regenerating tentacles (black arrows). A fully regenerated animal that was bisected 14 days prior was used as a regenerated control (B).
Discussion

By this analysis, the oldest HSPG2 orthologue perl is encoded by the genome of T. adhaerens. It exists on the chromosome in nearly the same format as in bilaterians, but is split into two genes. T. adhaerens perl encodes an SEA domain, which distinguishes this protein from D. melanogaster trol [Datta and Kankel, 1992], C. elegans unc-52 [Rogalski et al., 1993], and D. rerio pln [Zoeller et al., 2008] in that the perlecan protein in these animals is not predicted to contain an SEA fold. The SEA domain is commonly thought to have a regulatory activity on glycosylation during post-translational modification. Unlike HSPG2 in mammalian organisms, T. adhaerens perl does not encode consensus GAG attachment sites in the N-terminal domain I, although the T. adhaerens genome does encode genes with homology to the heparan sulfate synthetic machinery (Appendix C). The rest of T. adhaerens perl appears to exist in much the same organization as that of human HSPG2, in that regions encoding its various folding domains are organized in the same N to C order. Our sequencing reveals that there are at least 38 Ig folds between the domain IV orthologues of perl 2 and perl 3, 17 more than the 21 encoded by domain IV of the human gene, and more than the most found in any HSPG2 orthologue in the vertebrate tree. If these two genes eventually became one gene, as they must have since perl is one gene in bilaterian organisms, then it is a straightforward deduction to see that intronic recombination between T. Adhaerens perl 1 and perl 2 could have created one unified gene expressing close to 21 Ig folds in domain IV.

Exon shuffling or intronic recombination were key to the evolution of ECM and membrane proteins [Hynes, 2012]. As discussed in previous publications and included in figure 4.1, the A. queenslandica genome contains genes encoding orthologues of all of the constituent
pieces of the perlecan protein, but they are located far from one another in the genome. There is a single region of the A. queenslandica genome that houses nearly 40 repeats homologous to Ig modules, and it is possible that this, in addition to the other regions encoding homologous modules, is from where the T. adhaerens perl gene originated. Our analysis of the T. adhaerens perl indicates that nearly every piece of the gene could have been assembled from exons of other previously extant homologous genes.

Phylogenetic analysis of HSPG2/perl transcripts reveal that the T. adhaerens perl is indeed the most ancient of these genes, in agreement with the commonly accepted view of T. adhaerens’ place on the metazoan tree. Using a similar method to create a phylogenetic tree using T. adhaerens perl 1’s domain I and II, in comparison with the domain I and II peptide sequences from the other well-studied HSPG2/perl genes, T. adhaerens is actually placed closest on the tree to vertebrates (data not shown). This is attributed to the aberrant nature of domain I and II of other invertebrates: D. melanogaster trol has included many repeats of the LDLR module, while C. elegans unc-52 has incorporated an additional Ig module into the beginning of domain II and no SEA module of domain I. There can be two explanations of this, one being that parts of the T. adhaerens perl gene have remained unaltered for far longer than would be expected, and the other that placozoans developed these parts of the perlecan transcript through convergent evolution. Given the vastly different demands on placozoans compared to mammals and birds, it seems unlikely that these parts of the T. adhaerens perl gene evolved convergently. Certain of the information put forth in this tree regarding well-established evolutionary relationships is false (see S. purpuratus grouping with protostomes, nematodes grouping with vertebrates and certain nodes in the mammal tree having a low bootstrap score). Still the core
observation of this analysis, that *T. adhaerens perl* is more diverged from their common ancestor than that of *N. vectensis*, remains valid.

Although one paper [Ozbek et al., 2010] suggests that a perlecan homologue is found in the publically accessible genome of the hemoscleromorph sponge *Oscarella Carmela*, our search for this homologue found many laminin, LDL receptor, Titin, and hemicentin homologues but only one curious transcript which encoded a protein identified by BLASTp as a match for HSPG2. Without closely investigating this transcript one would accept that there exists some orthologue of *HSPG2* in the *O. carmela* genome, but PFAM analysis of the primary structure of this encoded protein reveals that this product differs greatly from perlecan, including Von Willebrand A modules interspersed amongst only 5 Ig modules and 4 LDLRa modules. Since there are no other significant matches which are identified by BLAST, this transcript and putative protein are most likely a product unique to sponges (analysis in appendix D).

I propose that *perl* first appeared in the common ancestor of *T. adhaerens* and *N. vectensis*, not in the common ancestor of sponges and placozoans. The presence of *perl* complicates previously published conclusions illustrating by electron microscopy the apparent lack of an organized ECM in placozoan tissue, but recent proteomic data supports our findings, in that perlecan-like peptides are found in the placozoan [Ringrose et al., 2013]. It is thought that *M. brevicollis* encodes a number of protein domains common to extracellular proteins, but that these organisms do not secrete complex, multimodular ECM proteins similar to those found in multicellular animals [Williams et al., 2014]. It is hypothesized that the choanoflagellate secretome serves to mediate interactions between the organism and its environment [King et al., 2008]. It is possible that the slime layer of *T. adhaerens* evident
on the upper epithelium comprises the bulk of their secreted proteins. If the *T. adhaerens* perlecan is secreted apically instead of basally, this would explain the lack of visible ECM in electron microscopic imaging. The acellular layer identified during these *in situ* and fluorescent imaging analyses indicate that there may be a heretofore unappreciated acellular layer ringing the animal, between the medial cell layer and the bulk of the animal’s body. It is possible that this acellular layer was missed by previous studies [Grell and Benwitz, 1974; Guidi et al., 2011] due to sectioning longitudinally *versus* looking at the organism from a top-down, latitudinal perspective.

Perlecan in vertebrate animals is critical for proper development and maintenance of tissue architecture [Arikawa-Hirasawa et al., 1999; Arikawa-Hirasawa et al., 2001]. In addition the protein is deposited during the later wound healing phase of various model systems [Yamazaki et al., 2004]. To determine if this function is conserved in simple animals, I chose to use the regeneration model *N. vectensis*. Detection of the *N. vectensis perl* gene using our tBLASTn-alignment method is confounded by a gap in the genomic construct – the 3’ end of the gene corresponding to the end of domain IV is cut off at the 3’ end of one scaffold and the 5’ end of another. Although I sequenced 4,400 base pairs, RACE PCR amplification of the *N. vectensis perl* transcript is incomplete, and I was unable to bridge the gap in these scaffolds. Thus it is uncertain whether *N. vectensis perl* is encoded as multiple genes, as in placozoans, or is assembled as one gene. The spatial organization of the perlecan protein modules follows the same order as that of all other perlecan proteins. There is no SEA domain encoded as part of the *N. vectensis perl*, which corresponds with the perlecan protein of *D. melanogaster*, nematodes, and fish. Our *in situ* hybridization analysis of longitudinally sectioned *N. vectensis* adults demonstrates the expression pattern
of *perl*. In fully developed animals the *perl* transcript is expressed in the ectoderm of the scapus, and sporadically in the mesentery and inner cell layer of the tentacles. *De novo* *perl* transcription is induced during stage 3 and 4 of regeneration, i.e. when oral structures such as the tentacles and pharynx begin to reform. This expression is lost from the tentacles once regeneration is fully accomplished. I hypothesize that the protein is needed to ensure proper tissue architecture during regrowth. Negative controls for *in situ* hybridization are presented in figure A.2. Hybridization with sense control riboprobes against *N. vectensis perl* (figure A.2 panel A), *T. adhaerens perl 1* (data not shown) and *T. adhaerens ACTB* (data not shown) all stained every cell in these animals. It is unclear why this occurs, although it is possible that there are antisense transcripts in these animals given the small size of their genomes (the *T. adhaerens* genome is 50 million base pairs in size). To account for possible residual alkaline phosphatase activity (which would result in chromogenic background stain in this assay), a no-probe control was performed during *T. adhaerens in situ* hybridization. This control (data not shown) resulted in no color development, indicating that there is little residual alkaline phosphatase activity in these animals after fixation.

This study identifies a possible origin of the *HSPG2/perl* gene in the poriferan phylum and the first assembled perlecian-like gene *perl* in the genome of the placozoan *T. adhaerens*. I have demonstrated expression of the gene in both placozoans and cnidarians, and demonstrate its activation during regeneration of lost tissue in *N. vectensis*. It is our hope that this sheds insight into the studies of animal multicellularity and tissue regeneration.
Chapter 5

Summary, General Discussion & Future Directions

These studies were designed to increase our understanding of the evolution and expression regulation of the HSPG2 gene, using the model systems of prostate cancer and stromal cell lines, as well as the basal metazoans Trichoplax adhaerens and Nematostella Vectensis. These two studies, while disparate in focus and methodology, are unified by the study of and fascination with the complexity of the perlecain gene and protein.

Summary: Perlecain expression in prostate cancer and stromal cells

The prostate cancer microenvironment is constantly changing throughout progression of the disease [Tuxhorn et al., 2001; Tuxhorn et al., 2002]. Increasing recruitment of stromal immune and vascular components to the tumor are typical of Gleason grade 4 and 5 disease in the primary site. Upon metastasis to the bone marrow cancer cells encounter a dynamic environment with fluctuating growth factor and matrix as bone remodeling occurs. Adaptation to the bone marrow microenvironment and tumor cell-bone marrow stromal cell interaction results in osteoblastic/osteosclerotic lesions, driven by tumor cell osteomimicry. The cure rate for metastatic prostate cancer is near 30 per cent [Pond et al., 2013]. This is a reflection of the heterogeneity and adaptation of diseased cells, and there are generally few avenues of recourse for these patients. Signaling through heparan sulfate may be a key regulatory component for growth of certain prostate cancers [Savore et al., 2005], and the microenvironment certainly can have powerful effects of cancer growth [Zong et al., 2012]. Understanding the regulation of the HSPG2 gene amidst the dynamic
signaling environment may be informative for a common regulatory pathway controlling the expression of other extracellular factors in cancer and stromal cells.

The first half of this dissertation was designed to investigate the regulation of the \textit{HSPG2} gene in the context of prostate cancer and associated stromal cells. The major findings of these studies are as follows:

1. The \textit{HSPG2} regulatory region was studied using promoter-reporter assays in prostate cancer, prostate stromal, and bone marrow stromal cell lines. The promoter-reporter responded to TNF-\(\alpha\) treatment in HS-27a bone marrow stromal cells, WPMY-1 prostate stromal cells, LNCaP and C4-2B prostate cancer cells. It did not respond to TNF-\(\alpha\) treatment in HS-5 bone marrow stromal cells. This difference between the two bone marrow stromal cell lines is interpreted as a product of the phenotypic rift between these two cell lines, HS-5 representing a secretory subtype and HS-27a representing a matrix-secreting subtype of the bone marrow stromal cell population. The effect of TNF-\(\alpha\) on reporter activity in the HS-27a and WPMY-1 stromal cell lines could be amplified by the addition of TGF\(\beta1\) to the culture medium, although TGF\(\beta1\) alone did not stimulate the promoter in these cell lines. The promoter-reporter constructs were used as a screening tool for the effect of various signaling molecules on \textit{HSPG2} promoter activity in these cell lines, and no other molecules elicited an effect as potent or widely relevant as did TNF-\(\alpha\). This reveals an inherent weakness in using artificial reporter constructs as a proxy for measuring gene regulation by signaling molecules. Nevertheless this promoter-reporter construct was designed correctly to identify this mechanism of regulating \textit{HSPG2} gene expression.
2. Serial deletions of the promoter-reporter constructs demonstrate activity of the distal 200 base pairs of the *HSPG2* regulatory region included in the full length reporter construct. The conserved NFκB consensus sites found in this region suggested that NFκB was indeed responsible for the increase in promoter-reporter activity during TNF-α treatment in these cell lines, and site-directed mutagenesis was used to test this hypothesis. Promoter-reporter constructs with a mutated NFκB consensus element no longer responded to TNF-α treatment. The data suggests that the more distal of the two putative NFκB consensus elements is active, whereas the element more proximal to the transcriptional start site is inactive in these cell lines. Deletion constructs also vary in activity levels under baseline culture conditions, suggesting that there are hidden regulatory elements in the region between 1.3 and 2.5 kb upstream from the start site of transcription. Insertion of the distal 200 kb upstream of a HSV-TK promoter-driven luciferase reporter does not increase the expression of this luciferase during TNF-α treatment, indicating these NF-κB elements are necessary but not sufficient to increase *HSPG2* expression in response to TNF-α treatment.

3. NFκB undergoes nuclear translocation during TNF-α treatment in the stromal cell lines HS-27a and WPMY-1, but not the bone marrow stromal cell line HS-5. This is in accord with the trend for *HSPG2* promoter-reporter response to TNF-α treatment and further implicates NFκB in the regulation of *HSPG2* expression.

4. The perlecan gene is regulated by the TNF-α-NFκB signaling axis in prostate cancer, prostate stromal, and bone marrow stromal cell lines. This regulation involves NFκB binding the distal region of the *HSPG2* promoter after 24 h of TNF-
α treatment. This binding was validated by chromatin immunoprecipitation using the LNCaP cell line. NFκB also bound the HSPG2 promoter in the endometrial cancer cell lines Ishikawa and Hec50, which were used as controls for demonstrable NFκB activation following TNF-α treatment in combination with interferon-γ.

5. HSPG2 transcription is increased in the LNCaP cells after 24 h of treatment with TNF-α in culture and tapers at 48 h. LNCaP cell secretion of perlecan is elevated at 24 and 48 h of treatment but tapers at 72 h. Similar increases in protein secretion were detected in HS-27a and WPMY-1 stromal cells on a per-cell basis. These cell lines demonstrated markedly decreased cell number following treatment with TNF-α and TGFβ1, suggesting that a subset of these cells survive treatment and secrete perlecan at high levels.

6. Perlecan is expressed in reactive stroma of prostate cancer, and only diffusely in the stroma of normal prostate tissue. This was demonstrated through immunohistochemistry of sectioned tissue from prostatectomy. Staining was most intense in sections containing cancer, whereas sections free of cancer were more commonly only diffusely stained by the perlecan antibody.

7. The basal transcription levels of HSPG2 as a percentage of β-actin are highest in the stromal cell lines HS-27a and WPMY-1, and significantly lower in the stromal cell line HS-5 and the prostate cancer cell lines C4-2B and LNCaP. HSPG2 transcription by HS-27a and WPMY-1 is dwarfed by that of primary bone marrow stromal cells, which is roughly five to six times higher as a percentage of Beta-actin.
8. Primary bone marrow stromal cells respond to TNF-α by increasing *HSPG2* transcription and perlecan secretion, but the response is highly variable sample to sample. This may reflect differences in metastatic state of these patients or even sex of the patient, as unfortunately the vital data of these patients is unavailable. These samples were collected between four and ten years ago, and it is possible that they are no longer responsive.

This project ended in the conclusion that *HSPG2* expression is regulated by NFκB binding the 5’ regulatory region of the gene in response to TNF-α expression. This mechanism was demonstrated in various cancer and stromal cell types. The TNF-α-NFκB signal transduction axis may effect numerous phenotypic changes in the stromal reaction or tumor progression pathways. The evidence that stromal cell lines decrease in cell number following treatment with TNF-α and/or TGFβ1 reflects the apoptosis of myofibroblasts following the completion of wound healing [Desmouliere et al., 1995; Iredale et al., 1998]. The role of inflammation and reactive stroma in disease progression is increasingly evident [Comito et al., 2013; Mangaonkar et al., 2012]; these studies demonstrate that *HSPG2* is yet another target gene in the signaling milieu of the reactive stromal microenvironment. The persistence of cancer cells through these treatments *in vitro* is indicative of cellular adaptation to the inflammatory environment of progressive prostate cancer. The adaptation to increase perlecan expression in response to inflammatory signaling could be a self-preservation mechanic. By increasing the presence of perlecan protein in the environment, these cells could be providing the signaling factors necessary (heparan sulfate and heparin-binding growth factors) for their own survival.
Summary: Perlecan in basal metazoans

The evolution of multicellularity and the loss of tissue regenerative capacity during evolution are critical events in the history of the metazoan tree. *Trichoplax adhaerens* is the sole species of the phylum placozoa [Eitel et al., 2013], a creature of roughly 4 to 5 cell types and separate cell layers [Guidi et al.]. Although it has no organized ECM its genome encodes a number of homologues for ECM genes from bilaterian organisms [Srivastava et al., 2008]. The convention is that the *Trichoplax* diverged from sponges as the sister group to cnidarians, but this is somewhat disputed [da Silva, 2007; Dellaporta et al., 2006]. The *Trichoplax* genome encodes a number of ECM molecules that the sponge does not [Srivastava et al., 2010]. This is in conflict with mitochondrial genomic evidence that *Trichoplax* is more similar to the choanoflagellate than the sponge [Dellaporta et al., 2006]. In addition, *Trichoplax* has a much simpler body plan than the sponge [Leys and Hill, 2012]. The cnidarian *Nematostella vectensis* is a more complex organism than the *Trichoplax* and is used as a model of tissue regeneration [Passamaneck and Martindale, 2012]. It is favored for its ease of cultivation in the laboratory. The genomes of these organisms were recently sequenced and as such provide ready models for the study of the origins of the perlecan gene and its involvement in tissue regeneration [Putnam et al., 2007; Srivastava et al., 2008; Srivastava et al., 2010]. A summary of the major findings of this portion of this dissertation are as follows:

1. Perlecan homologues *T. adhaerens perl* and *N. vectensis perl* have been identified in these recently sequenced genomes. The genetic pieces of these *perl* homologues can be identified in the genomes of the poriferans *A. queenslandica* and *O. carmela*
as well as the ctenophore *M. leidyi* although a single gene homologous to *HSPG2* cannot be identified in the genome of this animal. The choanoflagellate *M. brevicollis* is, as expected, even more bereft of these *HSPG2*-progenitor genes. *T. adhaerens perl* is expressed as two separate genes separated by only hundreds of base pairs on one chromosome. The presence of these *perl* genes puts the evolutionary origin of the perlecan protein at the common ancestor of sponges and placozoans, over 600 million years ago.

2. The phylogenetic trees created using various perlecan genes support the placement of placozoans as the sister group to sponges and cnidaria near the base of the metazoan evolutionary tree. This is not surprising given the fact that perlecan is not assembled in its entirety in these animals. One surprising aspect of *T. adhaerens perl* is the presence of an SEA module in the domain I region of the encoded protein. This separates *T. adhaerens perl* from *unc-52* and *trol* and even *hspg2* of zebrafish in that perlecan in these organisms does not include an SEA domain, although perlecan of mammals and birds does.

3. *N. vectensis perl* is expressed in the regenerating oral structures of the cnidarian. It appears to be developmentally important for the regenerating tentacles but is eliminated upon full formation of the structure. Normally it is expressed in the outer layer of the ectoderm, but not the gastroderm. It is also expressed normally in the mesentery of the organism, but only in select cells, without a discernable pattern.

This study was exploratory in nature. Upon the sequencing of the genomes of *T. adhaerens, A. queenslandica*, and *N. vectensis*, two previously opaque issues became
interrogable [Putnam et al., 2007; Srivastava et al., 2008; Srivastava et al., 2010]. One question was the evolutionary origin and age of the $HSPG2$ gene. The evolutionary origins of the $HSPG2$ gene are now clear: the constituent pieces of the gene, the nominal domains I, II, III, IV, and V, were assembled separately, possibly in the genome of the common ancestor of sponges and ctenophores, before they were ever associated as one transcript in a perlecan-like gene. The gene was assembled for the first time in the common ancestor between the poriferans and the placozoans, if indeed this is the evolutionary relationship of these two animals. $T. \text{adhaerens perl}$ includes every piece of the gene as we know it in modern mammals, and is strikingly similar to the human gene except for its division into two separate transcribed genes juxtaposed on the chromosome. It is unclear how $T. \text{adhaerens}$ perlecan maintains the SEA module of domain I whereas other organisms except mammals have lost this module.

The other major question addressed by this study was the activation of perl in tissue regeneration. The question of perlecan’s role in tissue regeneration was within reach before this study. Various model systems for regeneration exist, but none were readily available until I began cultivating $N. \text{vectensis}$ for the purpose of sequencing their RNA in search of a transcribed $HSPG2$ homologue. Now it is clear that $HSPG2$ transcription is activated in the regeneration of the oral structures of the cnidarian, and then repressed to its normal levels upon full regeneration. The cnidarian offers a robust and easily manipulable system for the study of tissue regeneration. It is a valuable tool that can be used to draw useful parallels to regeneration in more complex animals. Given the presence of the BM in the body wall of $Nematostella$ and differentiated cell types, establishment of organized ECM after bisection of the cnidarian is fertile ground for future investigations.
The integral finding of both of these studies is the activation of perlecan in tissue regeneration of the cnidarian and reactive stroma of prostate cancer, both processes related to wound healing. It is important to note that the function of perlecan in tissue regeneration must be separate from the proteoglycan’s action as a reservoir for heparin-binding growth factors given that there are no O-glycosylation sites in the *N. vectensis perl*. This insight could inform thoughts on perlecan’s role in wound healing in vertebrates and by extension in the reactive stroma of cancer. Considering the structural role of the perlecan protein as more ancient and fundamental than the growth factor-delivery role of the perlecan proteoglycan may be a useful perspective when considering activation of the gene in pathologic states in humans.

**Future Directions for the study of perlecan in prostate cancer**

p65 binding to the perlecan promoter in various cancer cell lines implicates the NFκB and inflammatory cytokines in the cancer microenvironment as modulators of ECM composition. Reactive stroma in the prostate may be regulated by inflammatory cytokines as well [Barron et al., 2010]. These findings raise many questions. Two major avenues of research would best continue these studies.

A further characterization of TNF-α, NFκB and perlecan co-localization in prostate cancer tissue specimens is key to fully establishing this phenomenon as a meaningful event in cancer progression *in vivo*. This will likely be variable and only found in a subset of patients, particularly those with high levels of inflammatory infiltration into the tumor. The cell type culpable for inflammatory induction of perlecan secretion by tumor and stroma alike through TNF-α signaling is most likely to be the M1 macrophage. Although M2 macrophages are the typical tumor-associated macrophage type [Sica et al., 2006], M1
macrophages can be recruited to tumors [Pantano et al., 2013]. M1 macrophages secrete TNF-α whereas M2 macrophages do not typically do so [Bogels et al., 2012; Carswell et al., 1975]. Investigation of this cell-cell interaction could be accomplished through co-culture of monocytes and stromal or cancer cells. Three-dimensional tissue culture models under development currently by other members of this lab will be invaluable tools for creating truly relevant in vitro models of tumor biology [Gurski et al., 2009]. Recreation of HSPG2 expression induction in a three dimensional model system using multiple cell types could be a powerful demonstration of this signaling pathway’s relevance to prostate cancer biology. The minimal response of the HSPG2 promoter-reporter constructs to co-culture of stromal cells and cancer cells (reported in chapter 2) could be blamed on the lack of a macrophage-like cell type in this culture system. Addition of this cell type to the co-culture could activate the stromal or cancer cell expression of perlecan driven by the TNF-α-NFκB signaling axis.

Another possible avenue of research is the study of NFκB in tumor-stroma interaction in general. There may be a connection between inflammatory signaling and reactive stroma recruitment, or induction of the desmoplastic phenotype. Collagen I, fibronectin, tenascin-C and versican promoters should be analyzed for conserved NFκB consensus elements to build some evidence for studying their regulation by this signaling axis. A screen of their production by western blot or quantitative PCR during TNF-α treatment is a simple method for testing this phenotype. Another gene that is coordinately upregulated with proteoglycans in other contexts is tumor necrosis factor-stimulated gene 6 (TSG-6) [Stove et al., 2000]. The TSG-6 protein interacts with hyaluronan [Heng et al., 2008] and other ECM molecules such as matrix metalloproteinases [Glant et al., 2002]. This gene is also
stimulated by TNF-α in certain cell contexts [Lee et al., 1990], and may be a part of a common cell program induced by TNF-α treatment along with perlecan.

The shift in cell-cell organization in stromal cells during TNF-α treatment that was observed in chapter 2 could be a bona fide phenotypic shift. It is possible that these cell lines are differentiating to reactive stromal cells or myofibroblasts, which could be identified by α smooth muscle actin, vimentin or fibroblast activation protein expression [Tuxhorn et al., 2002]. The creation of clearings in the dish and favoring cell-cell contact over cell-substrate contact could be related to increased perlecan expression, as coating tissue culture dishes has the effect of discouraging attachment and promoting cell-cell attachment in prostate cancer cell lines (Grindel et al. in prep). If this phenotype does turn out to be a bona fide reactive stromal response, RNA-seq could be a useful tool for investigating the entire phenotype, and differences between that of prostate stromal cells and bone marrow stromal cells.

Future directions for the study of perlecan and ECM in basal metazoa

Trichoplax adhaerens is an easily manipulable organism and is useful for the study of multicellularity. It is one of the simplest, if not the simplest animal, and has differentiated cell types. Its lack of organized ECM as identified through transmission electron microscopy is perplexing [Guidi et al., 2011]. These animals express T. adhaerens perl and encode genes homologous with various other ECM proteins. They secrete some substance to their dorsal surface, creating a slime layer thought to provide protection [Grell and Ruthmann, 1991]. Given the lack of microscopic evidence for an ECM it is possible that this dorsal slime layer is the product of their ECM molecules, including perlecan. If this is the case, then perlecan evolved as a barrier not between cells, but perhaps between
animals and their environment. Whole mount in situ hybridization using differential interference contrast or confocal microscopy (with fluorescent probes or antibodies) could identify exactly which cell type is expressing perlecan. This would allow more informed and useful speculation as to the function of perlecan in these animals. Morpholinos and dsRNA have been used to manipulate the trichoplax proteome [Jakob et al., 2004]. Knocking down perlecan in this way also could be informative, but no method for this technique has been established. Transgenic placozoans have not been created but would be a useful tool to circumvent the inherent weaknesses of using morpholinos. Because morpholinos are not heritable they will eventually be lost as the animals reproduce, diluting the effect of gene knockdown. This will not be the case with clonal transgenic or knockout placozoans.

Beyond the study of perlecan lie a host of unanswered questions concerning placozoan biology. The nature of their reproduction [Eitel et al., 2011], regeneration following fragmentation, stem cell population, digestion of food organisms, intercellular signaling, chemotaxis, maintenance of polarity, ECM, glycome, proteome, lipidome and transcriptome are all largely unstudied. Given the placozoans basal position on the metazoan tree all of these systems are of fundamental importance to the evolution of multicellular characteristics. The lack of understanding of *Trichoplax* sexual reproduction poses problems for the development of a classical linkage map. In the new era of genomics this will probably be solved using bioinformatics but this work has yet to be completed.

The cnidarian *Nematostella vectensis* has proven a robust model for tissue regeneration [Bossert et al., 2013]. Easily cultivated in the lab, many organisms can be grown and regenerated in a matter of months to create a clonal population. Sexual reproduction
creates many more organisms much more quickly [Stefanik et al., 2013a]. The same approach to perlecan knockdown suggested in *T. adhaerens* should be considered for *N. vectensis*. In this case, since embryos can be readily produced, transgenic animals have already been created [Renfer et al., 2010]. Thus the method already exists and should be taken advantage of. The tissue regeneration phenotype of perlecan knockout *N. vectensis* would be informative, but general muscle dysfunction similar to that found in *unc-52* nematodes is predicted [Rogalski et al., 1995]. This could interfere with feeding. When capturing *Artemia* the animals contract violently in both the body column and the tentacles. Another part of the nematode phenotype in perlecan mutants is the rupturing of the body wall following mechanical pressure, indicative of loss of the barrier function of the perlecan molecule [Gilchrist and Moerman, 1992]. Perlecan mutant animals should be closely analyzed for dysmorphic tissue.

The lack of GAG attachment sites in invertebrate perlecan and indeed perlecan in zebrafish means that the molecule evolved independently of the function of heparan sulfate [Rogalski et al., 1993; Zoeller et al., 2008]. These two molecules were joined later in evolution, indicating that their functions may have been needed in the same locations. An interesting line of inquiry is to determine the major function of heparan sulfate in the basal organisms *T. adhaerens* and *N. vectensis*. The genomes of both of these animals encode the proteins necessary for the synthesis of heparan sulfate [Feta et al., 2009]. Knocking these genes out and examining the phenotypes may echo the phenotype of perlecan knockouts in these animals, or it may have completely separate effects. These experiments will shed light on the nature of both the evolution of GAGs and perlecan.
The findings of research into perlecan have generally suggested the protein functions in the maintenance of borders between cell layers or between cells and their surroundings [Farach-Carson et al., 2013]. *HSPG2* expression activation during wound healing processes is probably dedicated to the re-establishment of cellular boundaries. The common thread in these studies is *HSPG2* expression during tissue remodeling. My studies have demonstrated the presence of a CRE binding the NFκB transcription factor in the regulatory region of the *HSPG2* gene that is activated in response to inflammatory cytokines, which can be a part of the extracellular milieu in tumor progression. Tissue regeneration of oral structures in *N. vectensis* includes a transient upregulation of *perl* transcription, potentially of importance in tissue patterning of regenerating tentacles. Perlecan’s myriad functions make understanding the protein’s key role in these processes difficult to parse. Successful mechanistic work in the future will clarify perlecan’s precise role in the maintenance and establishment of tissue boundaries.
References


confirmation of linkage to chromosome 1p, evidence of genetic homogeneity and reduction of the SJS locus to a 3-cM interval. Hum Genet 98:380-5.


Passamaneck YJ, Martindale MQ. 2012. Cell proliferation is necessary for the regeneration of oral structures in the anthozoan cnidarian Nematostella vectensis. BMC Dev Biol 12:34.


### Appendix A

#### Table A.1

Oligonucleotide primers and riboprobes used in this study

<table>
<thead>
<tr>
<th>Primer</th>
<th>5' -&gt; 3' Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.6 kb <em>HSPG2</em> Promoter F</td>
<td>GGGCCCGAGGTAGGGAACAGAC</td>
</tr>
<tr>
<td>2.4 kb <em>HSPG2</em> Promoter F</td>
<td>GCTGTGGAGGCTGCTCCTCTATCAG</td>
</tr>
<tr>
<td>5' UTR <em>HSPG2</em> Promoter R</td>
<td>GCCCCCGCAGCGCCTCTC</td>
</tr>
<tr>
<td>NFκB Mut1 Mutagenesis primer F</td>
<td>GAAGCCTTTTCTCATCGGACAAAAAGAAACCAAGTGCAACCTGGGTACTC</td>
</tr>
<tr>
<td>NFκB Mut1 Mutagenesis primer R</td>
<td>GAGTACACCAGGGTTGCACCTGGGTACTTCTCTCCTACGAGTAAAGGCTTC</td>
</tr>
<tr>
<td>NFκB Mut2 Mutagenesis primers F</td>
<td>GCCAACCTGGGTACTCTCCCTCAAATAAGAAGACCTGGGTGCTGT</td>
</tr>
<tr>
<td>NFκB Mut2 Mutagenesis primers R</td>
<td>CACAGCCACCAGGGCATTTCTCTTTATTTGAGGAGAGTACACCAGGGTTTG</td>
</tr>
<tr>
<td><em>HSPG2</em> qPCR F</td>
<td>AGCATGGACGTTGGGTGTG</td>
</tr>
<tr>
<td><em>HSPG2</em> qPCR R</td>
<td>GGCCTGCGTTGGTAGCCGT</td>
</tr>
<tr>
<td><em>ACTB</em> qPCR F</td>
<td>GATGAGATGGCATGGCTT</td>
</tr>
<tr>
<td><em>ACTB</em> qPCR R</td>
<td>CAACCTCACCAGGGTTC</td>
</tr>
<tr>
<td><em>HSPG2</em> NFκB site qChIP F</td>
<td>GAAGCCTTTTCTCATCGGACA</td>
</tr>
<tr>
<td><em>HSPG2</em> NFκB site qChIP R</td>
<td>AATTTCTGGTGCTGCTCCAAT</td>
</tr>
<tr>
<td><em>HBB</em> promoter qChIP F</td>
<td>CTGGTGGGGTGAATTCTTTGC</td>
</tr>
<tr>
<td><em>HBB</em> promoter qChIP R</td>
<td>AGTCCAAGCTAGGCCCTTT</td>
</tr>
<tr>
<td><em>GAPDH</em> promoter qChIP F</td>
<td>TCGACAGTCAGCCGCATCT</td>
</tr>
<tr>
<td><em>GAPDH</em> promoter qChIP R</td>
<td>CTAGCCTCCGCGGTCTT</td>
</tr>
<tr>
<td><em>NFKBIA</em> promoter qChIP F</td>
<td>GCTCAGGGTTTAGGCTTCTT</td>
</tr>
<tr>
<td><em>NFKBIA</em> promoter qChIP R</td>
<td>CTGGCGTGGGGATTCTCTT</td>
</tr>
<tr>
<td><em>N. vectensis in situ</em> probe R</td>
<td>GAGAACATTGGCCTCGGTG</td>
</tr>
<tr>
<td><em>N. vectensis in situ</em> probe F</td>
<td>CGGGCGGAACACTCCTCCCAAGG</td>
</tr>
<tr>
<td><em>T. adhaerens in situ</em> probe F</td>
<td>AGGCAGCAATCGTCTGCTG</td>
</tr>
<tr>
<td><em>T. adhaerens in situ</em> probe R</td>
<td>AGAGGCTTTTGCAGTTCTG</td>
</tr>
</tbody>
</table>
Figure A.1 Riboprobes synthesized for *T. adhaerens* and *N. vectensis* in situ hybridization. Schematic representations of *T. adhaerens perl 1* (A), *T. adhaerens perl 2* (B), and *N. vectensis perl* (C) transcripts. Red arrows represent coding exons. Grey bars represent protein domain encoded by the labeled exons. Orange rectangles show probes used in this study, dark red probes were also cloned but not used in experiments. Green bar in (C) shows the two sequencing scaffolds to which these sequences are assigned.
Figure A.2 Negative controls for *in situ* hybridization of *N. vectensis*. Sectioned *N. vectensis* were hybridized with sense control riborobes to demonstrate background signal levels. (A) A sense probe directed at the same sequence as the anti-sense probe used for characterization of *perl* expression in chapter 4. Blue staining represents positive hybridization. (B) A sense riboprobe synthesized from the *G. gallus TGFB3* transcript demonstrates little binding to the adult *N. vectensis*. Images taken with a 5X objective.
Appendix B

HSPG2 gene ChIP-seq data from ENCODE

Figure B.1 Histone modifications of the 5' region and intron 1 of the HSPG2 gene

Using the UCSC genome browser’s data derived from the ENCODE project, one can view the ChIP-Seq results for the acetylated histone marker (H3K27Ac), representing activated regulatory elements. The second track represents ChIP-seq results for H3K4Me3, a marker of activated proximal promoter regions. Higher peaks represent positive findings from ChIP-seq. RNA-seq and RNA PolII ChIP-seq data are also shown, representing transcription levels and RNA polII occupancy on the gene. The 5’ region of the HSPG2 gene is shown, including the 40 kilobase intron one, which appears to be associated with active regulatory elements. The region up to 7 kilobases upstream of the start site of transcription also appears to house active regulatory elements. Exons are labeled, as are the first intron and the upstream region. The numbering above identifies the position on chromosome 1. Light blue peaks – HUVEC cells. Light green – HSMM cells. Orange peaks – GM12878 cells. Purple peaks – K562 cells.
Figure B.2 TNF-α secretion by cell lines. Conditioned medium was collected after 24 h in culture with cell lines. ELISA detected TNF-α concentration in conditioned medium and is depicted in absolute concentration (A). Cells were counted and concentration was calculated on a per-cell basis (B). Bars, mean +/- SEM, n = 3. *, p<.05 vs HS-5. **, p<.01 vs HS-5. ***, p<.001 vs HS-5.
Figure B.3 Negative controls for immunofluorescent imaging of p65 nuclear localization.
Negative control The blocking peptide (A,C) was applied to cells prior to addition of primary anti-p65 antibody. Primary anti-p65 antibody was omitted in secondary antibody-only controls (B,D). Images represent HS-27a (A,B) and HS-5 (C,D) cells imaged using the 63X objectives. Blue, DAPI. Green, phalloidin-488.
### Appendix C

*T. adhaerens* gene orthologues for heparan sulfate pathway components

<table>
<thead>
<tr>
<th>Gene homologue</th>
<th>BLASTp hit</th>
<th>BLAST bit score/E</th>
<th>Scaffold:position</th>
<th>JGI transcript</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-O-sulfotransferase</td>
<td>heparan sulfate 6-O-sulfotransferase 3</td>
<td>231/6e^-70</td>
<td>1:8,112,590-8,111,520</td>
<td>19920</td>
</tr>
<tr>
<td>Glucuronic acid epimerase</td>
<td>D-glucuronol C5-epimerase</td>
<td>417/1e^-136</td>
<td>6:4,921,962-4,920,905</td>
<td>27335</td>
</tr>
<tr>
<td>N-deacetylase</td>
<td>bifunctional heparan sulfate N-deacetylase/N-sulfotransferase 3</td>
<td>626/0</td>
<td>8:1,676,314-1,673,687</td>
<td>28257</td>
</tr>
<tr>
<td>Xylosyl transferase</td>
<td>xylosyltransferase 2</td>
<td>494/6e^-163</td>
<td>2:593,710-593,091</td>
<td>21022</td>
</tr>
<tr>
<td></td>
<td>glypican-6 precursor</td>
<td>674/0</td>
<td>32:138,407-138,636</td>
<td>61971</td>
</tr>
<tr>
<td>Glypican</td>
<td>glypican-6 isoform X1</td>
<td>486/1e^-165</td>
<td>32:122,166-122,438</td>
<td>61970</td>
</tr>
<tr>
<td></td>
<td>glypican 1 precursor</td>
<td>132/2e^-37</td>
<td>32:120,222-120,296</td>
<td>61969</td>
</tr>
<tr>
<td>Agrin</td>
<td>Agrin</td>
<td>207/3e^-53</td>
<td>21:9,697-8,677</td>
<td>61048</td>
</tr>
<tr>
<td>Syndecan</td>
<td>Agrin isoform X6</td>
<td>143/3e^-38</td>
<td>3:1,000,612-1,000,665</td>
<td>7530</td>
</tr>
</tbody>
</table>

**Table C.1**

**Heparan sulfate pathway component gene homologues in *T. adhaerens***
Appendix D

Analysis of *O. carmela* and *M. leidyi* perlecan BLAST hits for perlecan

**Figure D.1 Search of *O. carmela* transcriptome for HSPG2 homologues.** BLAST hits from the *O. carmela* transcriptome were translated and analyzed by PFAM (A). BLAST hit names were assigned by the transcriptome database. Under description is listed the protein homologue encoded by each transcript, what type of folding modules are included in this protein, and how many times each folding module is repeated in that protein. One protein was identified by BLAST as an *HSPG2* homologue, albeit with low confidence (red box – A). This protein is shown in schematic form in (B). Several Ig modules and LDL-receptor-like modules are encoded, as well as two Von Willebrand A modules.
Appendix E

A circular isoform of the HSPG2 transcript

Introduction

The circularization of RNA is a long-suspected [Hsu and Coca-Prados, 1979] and recently verified phenomenon that has implications in the field of micro RNAs [Memczak et al., 2013]. The common conclusion is that in some cases where gene structure is permissive, alternative splicing of transcripts can lead to circularization of several exons to result in a stably circularized RNA isoform, entirely separate from lariats formed during normal splicing [Jeck et al., 2013]. This phenomenon is widespread, occurring in animal cells as well as in prokaryotes [Danan et al., 2012]. There is some evidence that circularized isoforms of RNA act as micro RNA sponges, serving as yet another regulator of transcription by sequestering these micro RNAs without being susceptible to degradation by the Dicer-RISC pathway [Hansen et al., 2013].

Methods

Divergent (PCR)

Circularization of the HSPG2 transcript was verified by the use of divergent PCR. Primers were designed to begin amplification 5’ from exon 2 and 3’ from exon 3. The reverse complement of these primers was used a control to demonstrate the presence of these sequences in the linear transcript. Amplimers were isolated using the Qiaquick PCR Purification kit (Qiagen, Venlo, Limburg). These PCR products were cloned using the pCR®2.1 TOPO kit (Life technologies, Carlsbad, CA). These plasmids then were sequenced.
Results

HSPG2 mRNA is circularized in stromal cell lines

The HSPG2 gene has the structural hallmarks of genes whose transcripts are circularized during splicing. To investigate this possibility, divergent PCR was performed on pooled cDNA from various treatments of stromal and cancer cell lines (figure E.1A). Primers were designed to amplify the linear transcript as in conventional PCR, and the reverse compliment of these primers was used to amplify any transcript that could be circularized involving exons 2 and 3. Divergent PCR amplified detectable products in HS-27a, HS-5 and LNCaP cell cDNA pools (figure E.1B). Upon cloning and sequencing of these products, the HS-27a cell product was unambiguously aligned with exons 2 and 3 of the HSPG2 transcript and with a heretofore unrecognized exon, found 1.5 kilobase pairs 5′ of exon 2 (figure E.1C). The sequencing data suggests that there is a circularized RNA product that comprises exons 2, 3 and the novel exon (figure E.1D).
Figure E.1 The *HSPG2* transcript is circularized in bone marrow stromal cells and contains a heretofore unrecognized exon. Divergent PCR was used to amplify circularized portions of the *HSPG2* transcript. Divergent and linear primers (A) were designed as the reverse complement of one another, and amplified out of exons 2 and 3. cDNA from various treatments (TNF-α, TGFβ1, vehicle control) of stromal and prostate cancer cell lines was pooled and used for RT-PCR (B). Controls included ACTB and *HSPG2* primers used in qPCR experiments. Divergent RT-PCR resulted in 300-1000 base pair amplimers (red boxes, B) in HS-27a, HS-5 and LNCaP cells. Sequencing reveals an additional exon in the *HSPG2* gene (green, C) included in a circularized transcript (green, D).
**Discussion**

The circularization of mRNA is a recently accepted phenomenon which has been suspected to occur for some time [Hsu and Coca-Prados, 1979; Memczak et al., 2013]. Various model systems have been used to attribute a role in micro RNA activity regulation to several circularized transcripts [Hansen et al., 2013]. The general attributes of transcripts which have circularized splicing products have been delineated as such: increased length of the intron upstream of the acceptor exon and downstream of the donor exon, relative decrease in length of the intron downstream of the acceptor exon, an increase in incorporation of intron 2 into scrambled transcript isoforms [Jeck et al., 2013]. Genes with circularized transcript isoforms are also longer than the average gene and have longer introns than the average. All of these traits describe the human *HSPG2* gene, which was the inspiration for using divergent PCR to test for circularized isoforms of the transcript including exon 2.

Several tests should be performed to fully characterize the circularized isoform of the *HSPG2* transcript. The standard means for identifying truly circularized isoforms of RNA is to use RNAse R digestion [Memczak et al., 2013]. RNAse R digests only linear RNA molecules, and leaves circularized isoforms intact. Thus after RNAse R digestion, divergent reverse transcription PCR can be performed as usual. In this experiment, the linear isoform of the cDNA should not be amplified but the circularized isoform should still be amplified by divergent PCR. If this circularized *HSPG2* isoform survives this test, then it represents a truly circularized splicing product of this transcript. The only identified role for circular RNA to date is as a sponge for micro RNAs (miRNAs) [Hansen et al., 2013]. These circular isoforms contain multiple incomplete (less than 12 base pair) binding
sites for various miRNAs and thus serve as sinks for these RNAs in the nucleus but are not degraded by the miRNA-Dicer pathway [Memczak et al., 2013]. Exons 2, 3 and the newly identified exon should be tested for matching miRNA seed sequences.