

10-8-2008

# Biophysical Characterization of the Sequence-Stranded DNA-Binding Properties of Mouse Pur : a Repressor of Smooth Muscle -Actin Gene Expression

Jon Ramsey  
*University of Vermont*

Follow this and additional works at: <http://scholarworks.uvm.edu/graddis>

---

## Recommended Citation

Ramsey, Jon, "Biophysical Characterization of the Sequence-Stranded DNA-Binding Properties of Mouse Pur : a Repressor of Smooth Muscle -Actin Gene Expression" (2008). *Graduate College Dissertations and Theses*. Paper 189.

This Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks @ UVM. It has been accepted for inclusion in Graduate College Dissertations and Theses by an authorized administrator of ScholarWorks @ UVM. For more information, please contact [donna.omalley@uvm.edu](mailto:donna.omalley@uvm.edu).

**BIOPHYSICAL CHARACTERIZATION OF THE SEQUENCE-SPECIFIC  
SINGLE-STRANDED DNA-BINDING PROPERTIES OF MOUSE PUR $\beta$ :  
A REPRESSOR OF SMOOTH MUSCLE  $\alpha$ -ACTIN GENE EXPRESSION**

A Dissertation Presented

by

Jon Eric Ramsey

to

The Faculty of the Graduate College

of

The University of Vermont

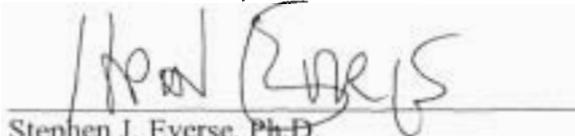
In Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy  
Specializing in Biochemistry

October, 2008

Accepted by the Faculty of the Graduate College, The University of Vermont, in partial fulfillment of the requirements for the degree of Doctor of Philosophy specializing in Biochemistry.

**Thesis Examination Committee:**

  
\_\_\_\_\_  
Robert J. Kelm Jr., Ph.D. **Advisor**

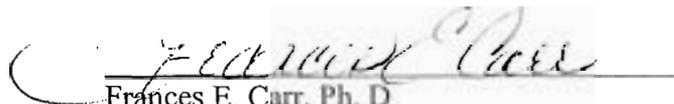
  
\_\_\_\_\_  
Stephen J. Everse, Ph.D.

  
\_\_\_\_\_  
Anne B. Mason, Ph.D.

  
\_\_\_\_\_  
Scott W. Morrill, Ph.D.

  
\_\_\_\_\_  
Mercedes Rincon, Ph.D.

  
\_\_\_\_\_  
Robert B. Low, Ph. D. **Chairperson**

  
\_\_\_\_\_  
Frances E. Carr, Ph. D. **Vice President for Research and Dean of Graduate Studies**

Date: August 29<sup>th</sup>, 2008

## ABSTRACT

Regulation of gene transcription by structural interconversions of genomic DNA is an emerging biochemical and genetic paradigm that adds to the already diverse repertoire of eukaryotic gene regulatory mechanisms. The appearance of paranemic structures coincident with changes in gene activity, as well as participation of transcription factors that recognize and bind single-stranded DNA at numerous gene promoters *in vivo* illustrates the authenticity of this concept and its importance in cellular homeostasis. Despite its acceptance, this concept has been minimally described at the biochemical and biophysical levels, as the means by which sequence-specific single-stranded DNA-binding proteins exert transcriptional influence in double-stranded genomes remains largely undefined.

Pur $\beta$  is a sequence-specific single-stranded DNA/RNA-binding protein that acts as a repressor of smooth muscle  $\alpha$ -actin (SM $\alpha$ A) gene transcription, and mRNA translation. SM $\alpha$ A is an important cytoskeletal protein that contributes contractile, antimigratory, and nonproliferative functions in smooth muscle. In concert with Pur protein family member Pur $\alpha$ , and Y-box protein MSY1, Pur $\beta$  enacts repression of SM $\alpha$ A gene expression by interacting with a cryptic *cis*-regulatory element in the 5' region of the SM $\alpha$ A promoter that has been shown to transiently adopt single-stranded conformations *in vivo*, and to confer transcriptional activation when *trans*-activator occupied while in a double-stranded conformation. Downregulation of SM $\alpha$ A gene expression has been identified to be a contributing factor to cardiovascular disease progression; therefore a thorough understanding of SM $\alpha$ A repression mechanisms is critical for clinical management of these conditions.

Although highly homologous at the primary sequence level, Pur $\beta$  and Pur $\alpha$  display significant conserved regions of sequence divergence that suggest these paralogs exert distinct cellular functions in various vertebrate classes. A goal of the studies presented herein was to delineate exhibited functional differences with respect to SM $\alpha$ A repression in pertinent mouse cell lines. Loss-of-function and chromatin immunoprecipitation studies verified that Pur $\beta$  differs from Pur $\alpha$  in that Pur $\beta$  is the dominant Pur protein repressor of SM $\alpha$ A expression in embryonic fibroblasts and vascular smooth muscle cells, although by different, cell type-specific mechanisms.

Biophysical assessment of Pur $\beta$  single-stranded DNA binding properties showed that despite the ability of Pur $\beta$  to self-dimerize in the absence of nucleic acid, Pur $\beta$  binds to the cryptic SM $\alpha$ A enhancer by a sequential and cooperative mechanism, with remarkable affinity and a terminal stoichiometry of 2 to 1. Footprinting and *in vitro* binding site characterization confirms two Pur $\beta$  binding sites exist within this element and display slight degeneracy from a proposed Pur protein-binding consensus motif. These findings delineate binding mechanisms adopted by Pur $\beta$  and provide a means to identify putative Pur $\beta$  binding sites throughout the genome.

## **DEDICATION**

***For Kendra and Cooper.....***

*I owe you both so much for your unconditional love, patience, and enthusiasm. Everyday you remind me how much I need you. Your support has made all of this possible, without it I would not have been able to pursue this goal. Your contributions to my success are beyond words, yet plentiful and essential are certainly ones that come to mind.*

## **ACKNOWLEDGMENTS**

First and foremost I wish to thank my advisor Dr. Robert Kelm, Jr. for providing me with opportunity, support, advice, and freedom to explore new and exciting avenues of scientific investigation. I would also like to thank other members of the Kelm laboratory both past and present that helped me during my graduate career: Shu-Xia Wang, Anna Knapp, and numerous rotation students who helped advance my various projects. I would like to thank the American Heart Association and Dr. Kenneth G. Mann for financial support. I would also like to express gratitude to the members of my committee, Drs. Stephen J. Everse, Anne B. Mason, Scott W. Morriscal, Mercedes Rincon, and Robert B. Low for guiding me through my thesis research. I would like to thank all members of the Department of Biochemistry, especially Stephen J. Everse, Anne B. Mason, and Scott W. Morriscal for providing me with technical resources, Paula Tracy for generous advisement, and Margaret Daugherty for employment, encouragement, and plentiful advice before and during my graduate career. I am also grateful to members of the Vascular Biology group at the Colchester Research Facility, Drs. David Schneider, Burton Sobel, Chris Holmes, and Jeffrey Spees for insightful comments regarding my research. I am deeply indebted to Ewa Folta-Stogniew of Yale University for her help with light-scattering studies, and to Keith Connaghan and David Bain of the University of Colorado for extensive and courteous help with the nuts and bolts of DNase I footprinting and analysis.

None of this would have been possible without the loving support of my wife, Kendra, who has encouraged me at every step of my pre-doctoral training, my son Cooper, who is my inspiration, and my parents Jacqueline and Richard and sister Jan for telling me I could accomplish anything and everything I put my mind to.

Thank You.

# TABLE OF CONTENTS

DEDICATION .....	ii
ACKNOWLEDGMENTS .....	iii
LIST OF TABLES .....	vii
LIST OF FIGURES .....	viii
LIST OF USED ABBREVIATIONS .....	xi
CHAPTER I. INTRODUCTION .....	1
BACKGROUND TO THE PRESENT WORK .....	1
PHENOTYPIC PLASTICITY OF VASCULAR SMOOTH MUSCLE CELLS .....	2
FUNCTIONAL ASPECTS OF SMOOTH MUSCLE $\alpha$ -ACTIN IN SMOOTH MUSCLE CELL BIOLOGY .....	13
REGULATION OF SMOOTH MUSCLE $\alpha$ -ACTIN GENE EXPRESSION .....	15
PUR PROTEIN STRUCTURE AND FUNCTION .....	32
THE BREADTH OF PUR PROTEIN FUNCTION .....	32
STRUCTURAL CONSIDERATIONS .....	39
NUCLEIC ACID-BINDING PROPERTIES .....	48
CHAPTER II. SCOPE OF THE PRESENT WORK .....	67
CHAPTER III. PUR PROTEIN LOSS-OF-FUNCTION STUDIES IDENTIFY PUR $\beta$ AS A DOMINANT REPRESSOR OF SM $\alpha$ A EXPRESSION .....	70
INTRODUCTION .....	70
MATERIALS AND METHODS .....	75
RESULTS .....	85
DISCUSSION .....	96
CHAPTER IV. SOLUTION CHARACTERIZATION REVEALS THE QUATERNARY STRUCTURE OF RECOMBINANT MOUSE PUR $\beta$ .....	102
INTRODUCTION .....	102
MATERIALS AND METHODS .....	105

RESULTS .....	114
DISCUSSION .....	130
CHAPTER V. THERMODYNAMIC ANALYSIS OF SEQUENCE-SPECIFIC SINGLE-STRANDED DNA-BINDING BY RECOMBINANT PUR $\beta$ REVEALS A COOPERATIVE MECHANISM OF NUCLEOPROTEIN ASSEMBLY.....	136
INTRODUCTION .....	136
MATERIALS AND METHODS .....	143
RESULTS .....	159
DISCUSSION .....	190
CHAPTER VI. TOWARDS THE CHARACTERIZATION OF THE MINIMAL SINGLE-STRANDED DNA BINDING SITE OF PUR $\beta$ WITHIN THE SM $\alpha$ A PROXIMAL MCAT ENHANCER ELEMENT .....	203
INTRODUCTION .....	203
MATERIALS AND METHODS .....	206
RESULTS .....	210
DISCUSSION .....	213
CHAPTER VII. CONCLUDING REMARKS .....	220
COMPREHENSIVE BIBLIOGRAPHY .....	222

## LIST OF TABLES

Table 1.1. Oligonucleotides with established interactions with Pur proteins .....	51
Table 1.2. Pur $\alpha$ binding properties of mutant <i>c-myc</i> -associated <i>PUR</i> -element oligonucleotides.....	55
Table 4.1 Sedimentation equilibrium data: Parameters from global analysis.....	129
Table 5.1. Microscopic configurations and corresponding interaction free energies for a two site system. ....	154
Table 5.2. Microscopic configurations and corresponding interaction free energies for multiple models of two site-binding.....	160
Table 6.1. Oligonucleotides used in this study as fluid-phase competitors in Pur $\beta$ ssDNA-binding functional ELISA. ....	209

## LIST OF FIGURES

Figure 1.1. Regulation of SM $\alpha$ A gene expression involves numerous <i>trans</i> -activating and <i>trans</i> -repressing factor interactions at multiple <i>cis</i> -regulatory elements. ....	17
Figure 1.2. Multiple primary sequence alignments of Pur $\alpha$ orthologs versus Pur $\beta$ orthologs. ....	41
Figure 1.3. Pur $\alpha$ nucleotide preferences within putative origin-associated PUR-elements. ....	57
Figure 1.4 Nucleotides flanking the core MCAT enhancer are critical for Pur protein association. ....	60
Figure 3.1. Target positions of Pur $\alpha$ - and Pur $\beta$ -specific shRNA. ....	84
Figure 3.2. Specific knockdown of Pur $\alpha$ or Pur $\beta$ by shRNA-mediated RNAi. ....	87
Figure 3.3. De-repression of the SM $\alpha$ A promoter by Pur $\beta$ shRNA. ....	91
Figure 3.4. De-repression of the SM $\alpha$ A promoter in response to knockdown of Pur $\alpha$ and/or Pur $\beta$ and the requirement for Pur/Pyr element integrity. ....	94
Figure 3.5. Relative Levels of Pur $\alpha$ , Pur $\beta$ , and MSY1 occupancy at the SM $\alpha$ A promoter by ChIP. ....	97
Figure 4.1. Expression and purification of N-HisPur $\beta$ . ....	115

Figure 4.2. Molecular size measurements of N-HisPur $\beta$ in solution by light scattering techniques.....	117
Figure 4.3. Sedimentation velocity analysis of N-HisPur $\beta$ . .....	121
Figure 4.4. Analysis of N-HisPur $\beta$ sedimentation velocity by direct fitting of time-resolved concentration difference curves.....	123
Figure 4.5. Analysis of concentration-dependant N-HisPur $\beta$ solution non-ideality by sedimentation equilibrium. ....	126
Figure 4.6. Sedimentation equilibrium analysis of N-HisPur $\beta$ .....	128
Figure 4.7. Species plot of N-HisPur $\beta$ .....	132
Figure 5.1. Schematic of <i>cis</i> -regulatory elements of the SM $\alpha$ A promoter.....	142
Figure 5.2. Extent of recombinant and nuclear extract-derived Pur $\beta$ binding to PE32-F. ....	162
Figure 5.3 Serial dilution EMSA determination of the N-HisPur $\beta$ :PE32-F* nucleoprotein complex stoichiometry.....	169
Figure 5.4. Quantitative electrophoretic mobility shift assay of N-HisPur $\beta$ binding to PE32-F*. ....	171
Figure 5.5. Quantitative DNase I footprinting of N-HisPur $\beta$ :SMP382-F* interactions.	176

Figure 5.6. N-HisPur $\beta$ binds to the SM $\alpha$ A MCAT enhancer element via a cooperative 3' to 5' mechanism. ....	181
Figure 5.7. Monte Carlo error simulations to assess model confidence. ....	184
Figure 5.8. Monte Carlo error simulations to assess converged parameter confidence. ....	187
Figure 5.9. N-HisPur $\alpha$ and N-HisPur $\beta$ bind supercoiled and linearized plasmids VSMP8 and pBLCAT3. ....	189
Figure 5.10. Comparison of quantitative footprinting-resolved parameters to individual quantitative electrophoretic mobility shift species isotherms. ....	194
Figure 6.1. Analysis of N-HisPur $\beta$ :PE32-F nucleotide interactions by truncated oligonucleotide competition. ....	212
Figure 6.2. Analysis of N-HisPur $\beta$ :PE32-F nucleotide interactions by scanning oligonucleotide competition. ....	214
Figure 6.3. Recognition and binding of N-HisPur $\beta$ to the purine-rich strand of the SM $\alpha$ A cryptic MCAT enhancer element. ....	217

## LIST OF USED ABBREVIATIONS

$\alpha$ MHC,  $\alpha$ -myosin heavy chain; ANF, atrial natriuretic factor; ATP, adenosine triphosphate; bHLH, basic helix-loop-helix; bp, basepair; BPV, bovine papilloma virus; BSA, bovine serum albumin; CArG, 5'-CC(A/T)<sub>6</sub>GG-3' element; AT, chloramphenicol acetyl transferase; CE, coding element; *dhfr*, dihydrofolate reductase; DLS, dynamic light-scattering; DNA, deoxyribonucleic acid; dsDNA, double-stranded deoxyribonucleic acid; dsRNA, double-stranded ribonucleic acid; DTT, dithiothreitol; EDTA, ethylenediamine tetraacetic acid; ELISA, enzyme-linked immunosorbent assay; EMSA, electrophoretic mobility shift assay; HeLa, Henrietta Lacks cervical cancer cell line; HIV, human immunodeficiency virus; HJCV, human John Cunningham virus; hnRNP, heterogeneous nuclear ribonucleoprotein; LCE, lytic control element; LLS, laser light scattering; MBP, myelin basic protein; MCAT, muscle-specific 5'-CATTCCCT-3' box; MEF, mouse embryonic fibroblast; mRNP, messenger ribonucleoprotein; MSY1, mouse Y-box protein 1; NRE, negative regulatory element; nt, nucleotides; OB fold, oligonucleotide/oligosaccharide-binding fold; PCR, polymerase chain reaction; PDGF, platelet-derived growth factor; PE, promoter element; PNR, purine-rich negative regulatory element; pRb, retinoblastoma protein; PrM, purine-rich motif; PUR, purine-rich Pur protein-binding element; RI, refractive index; RNA, ribonucleic acid; SEC, size-exclusion chromatography; SM $\alpha$ A, smooth muscle  $\alpha$ -actin; SMC, smooth muscle cell; snRNA, small nuclear ribonucleic acid; SRF, serum response factor; SRP, signal recognition particle; SSB, single-stranded deoxyribonucleic acid-binding protein; ssDNA, single-stranded deoxyribonucleic acid; ssRNA, single-stranded ribonucleic acid;

TCE, transforming growth factor  $\beta$ 1-control element; TEF-1, transcriptional enhancer factor 1; TGF- $\beta$ 1, transforming growth factor  $\beta$ 1; THR, transforming growth factor  $\beta$ 1-hypersensitive region; USF, upstream stimulatory factor; UTR, untranslated region; UV, ultraviolet; VACssBF, vascular actin single-stranded deoxyribonucleic acid-binding factor; VSMC, vascular smooth muscle cell; YB-1, Y-box protein 1.

# CHAPTER I. INTRODUCTION

## BACKGROUND TO THE PRESENT WORK

Investigations into cellular processes that contribute to tissue remodeling associated with development, response to injury, and disease progression have determined that these courses are the result of phenotypic modulation of cells resident to remodeled areas. Sequence-specific single-stranded DNA/RNA-binding transcription/translation factor Pur $\beta$  (Pur protein isoform  $\beta$ ), has been implicated in the phenotypic modulation of vascular smooth muscle cells that become activated and phenotypically reprogrammed during vessel wall remodeling associated with arteriosclerotic disease progression. A cause and effect relationship between vascular smooth muscle cell phenotype-switching and cytoskeletal protein smooth muscle  $\alpha$ -actin expression has been experimentally established. Repression of smooth muscle  $\alpha$ -actin, accomplished in part by Pur $\beta$ , has been linked to activation, migration, proliferation, and hypersynthetic properties of resident smooth muscle cells at sites of vessel wall remodeling. Contrastingly, derepression, or activation of smooth muscle  $\alpha$ -actin expression is coincident to injury-induced myofibroblast activation and adoption of contractile properties necessary for wound closure and resolution by these cells. The potential involvement of Pur $\beta$  in these pathophysiological processes have made this protein an important target of investigation for understanding dynamic smooth muscle  $\alpha$ -actin expression in phenotypically reprogrammed cells, as well as a model for understanding sequence-specific single-stranded DNA-protein interactions.

The following is a review of the existing literature regarding the various physiological and pathological processes that rely heavily on vascular smooth muscle cell phenotype plasticity, molecular mechanisms governing dynamic smooth muscle  $\alpha$ -actin expression, and Pur protein structural/functional properties as they relate to these important aspects of vascular smooth muscle biology.

### **PHENOTYPIC PLASTICITY OF VASCULAR SMOOTH MUSCLE CELLS**

The primary recognized role of vascular smooth muscle cells (vascular SMCs, VSMCs) is that of generating contractile force within blood vessels, thus providing a means to regulate vessel tone, blood pressure, and the appropriate distribution of blood to the periphery. The contractile phenotype of VSMCs in adults represents the full extent of differentiation for this cell type, and in addition to contractility, is generally regarded as being quiescent or slow to proliferate, nonmigratory, and nonsynthetic. The contractile phenotype is routinely characterized at the molecular level by positive expression of a repertoire of contractile proteins, cell-surface receptors, and ion channels that have been identified as necessary for contractility (reviewed in (210, 315)). However, unlike their skeletal and cardiac muscle cell counterparts, in whom contractile phenotypes represent terminal differentiation, VSMCs can undergo reversal of differentiation, or dedifferentiation, and revert to a broad and continuous spectrum of cell phenotypes, or various levels of differentiation that defy categorization, ranging from contractile to those reminiscent of fibroblasts. Consistent with this spectrum of cell phenotypes, combinations and levels of differentiation marker expression, which in all likelihood

dictate phenotype and by which these cell types are characterized, are vastly heterogeneous, making practical and absolute categorical determination of cell phenotypes very difficult. Nevertheless, VSMC-to-fibroblast phenotype conversion is marked by loss of contractility, hyperproliferation, increased migratory capacity, and elevated synthesis of extracellular matrix proteins and proteases (209).

The value of vascular SMC phenotypic plasticity has long been debated. This discussion arises primarily from the fact that phenotypic modulations of VSMCs and other contractile cell types are observed in various physiological and pathological processes including tissue development, wound repair, and disease progression. The roles and duties carried out by SMCs in these pathophysiological scenarios will be described here to divulge the importance of phenotypic plasticity by examining our current understanding of this phenomenon, as well as point out gaps in the existing knowledge.

The full breadth and complexity of vasculature development including the spatial and temporal participation by VSMCs in this elaborate process is beyond the scope of this review, except to say that phenotypic plasticity of cells that participate in the construction of blood vessels is critical. A common belief in the field of developmental biology is that phenotypic plasticity possessed by VSMCs provides multifunctionality and hyper-responsiveness to environmental cues that coordinate developmental events (210) Examples of this utility include the ability of VSMCs to exhibit synthetic phenotypes during investment in vessel wall construction, manufacturing significant levels of collagen, elastin, paracrine factors, and adhesion molecules while expression of

contractile apparatus components is either downregulated or not yet activated (122). Evidence supports that unifying contractile phenotypes are derived from diverse lineages during embryonic development, however the precursor origins of differentiated VSMCs and SMC-like cells are not completely known. Cells comprising the ectodermal neural crest and mesodermal proepicardium have been shown to differentially commit to the construction of distinct vessels (great vessels and epicardial vessels, respectively) and assume SMC-like properties, despite originating from different transient embryonic entities (12, 53, 193). It appears partitioning of these cells of differing origins during development is coincident with required morphological alterations characteristic of their fully-developed vessel destinations. A wide variety of environmental cues regulate cell commitment during embryogenesis and development, however direction of VSMC commitment appears to be dominated by signaling of transforming growth factor  $\beta$ 1 (TGF $\beta$ 1, (38, 254, 261)), and requires activation of several downstream coordinators, including the transcription factors Msx2 and Necdin (25), and none more important than serum response factor (SRF)(147). Transient dedifferentiation of committed VSMCs appears to also occur during vessel remodeling that transpires in later developmental stages (122, 193). Collectively, these findings suggest that forward (differentiation) and reverse (dedifferentiation) phenotypic plasticity of VSMCs is tightly regulated and crucial for proper vasculogenesis.

Wound repair represents another aspect of smooth muscle biology that relies heavily on phenotypic plasticity of contractile cell types beyond differentiated VSMCs, and similar to tissue development, is regarded as being beneficial to vertebrates. Similar to

developmental mechanisms, non-vessel wound repair requires phenotypic modulation of resident fibroblasts to gain contractile capacity, similar to SMCs, necessary for closure and resolution of the wound, while maintaining synthetic, migratory and proliferative properties necessary for populating the wound area and secretion of growth factors that aid the healing course (63). Clearly, this modulated cell type is neither VSMC-like nor fibroblast-like, but shares properties of both. Accordingly, this cell type has been termed the myofibroblast to reflect the contractile fibroblast trait (89). Biochemically, myofibroblasts within granulation tissue are characterized by positive expression of protein components of contractile stress fibers, in particular smooth muscle  $\alpha$ -actin (SM $\alpha$ A), vimentin, desmin, lamin, and tubulins (127, 241) as well as non-muscle myosin and collagen type I (60). Resident fibroblasts of diverse tissues throughout the human body exhibit transdifferentiation capacities to adopt myofibroblast phenotypes, and accordingly roles beyond wound closure have been assigned to myofibroblasts, but typically involve contraction and secretion of extracellular matrix proteins and cytokines necessary for development, repair, and maintenance of anatomical structures (reviewed in (219, 253)). Worth noting, however, is the possibility that myofibroblasts originate from discrete progenitor stem cells early in development, not necessarily resident fibroblasts, and reside in tissues as quiescent proto-myofibroblasts (20). Controversy in the literature surrounds this issue as propagation of cultured fibroblasts in media containing TGF- $\beta$ 1 has certainly shown the ability of these cells to assume myofibroblast phenotypes *ex vivo* (64), and the detection of fibroblastic cells in animal injury models that stain positive for SM $\alpha$ A further supports the notion that transdifferentiation of resident fibroblasts gives

rise to myofibroblasts (60, 240). This disparity may be a simple reflection of differences in the developmental stages of the organisms in which these observations were made, and both observations may be accurate. The detection of circulating myofibroblasts progenitors, termed fibrocytes (26), clouds this issue further, but describes the complexity of cellular reprogramming and recruitment in wound healing, as well as the importance of phenotypically-flexible cells in supporting this process. It should also be noted that epithelial cells have also shown the capacity to transdifferentiate to myofibroblast-like phenotypes during epithelial-to-mesenchymal transitions observed in metastatic processes (83).

As noted above, an additional similarity between the transdifferentiation of fibroblasts in the formation of myofibroblasts and the phenotypic commitment of VSMCs in tissue development is the involvement of platelet-derived TGF- $\beta$ 1 in signaling this progression (64), suggesting that this factor signals similar downstream events that coordinate programmed expression of smooth muscle associated genes. Constitutive overexpression of this factor in rats causes systemic fibrosis characterized by high levels of collagen deposition (45). Platelet-derived growth factor-BB (PDGF-BB) (283), interleukin 4 (IL-4) (72), heparinoids (67), thrombin (119), and the ED-A domain of fibronectin (252) have also shown pro-transdifferentiation properties on fibroblasts. Other cytokines have displayed potential to subtly modulate myofibroblast phenotypes, with particular respect to the way these cells respond and adjust to extracellular matrix dynamics (reviewed in (253)). Whereas fibronectin ED-A epitopes are generated by traumatic shear force (252), the source(s) of cytokines for propagation of the injury

response is (are) less clear. Studies indicate that infiltrate white blood cells and local endothelial cells secrete these factors (45) or that myofibroblasts secrete factors themselves in an autocrine fashion (13). Post-healing withdrawal of myofibroblast proliferation, migration and synthetic character, as well as SM $\alpha$ A expression, can be initiated by exposure to interferons  $\alpha$  and  $\gamma$  (66) secreted by natural killer lymphocytes that infiltrate the wound area shortly after injury (124). In most instances, myofibroblasts undergo apoptosis after resolution of the wound (60), however under circumstances in which myofibroblasts forgo programmed cell death, for reasons that are not clear, pathological wound healing and scarring (fibrosis) is the result (65).

The arena in which VSMC phenotype plasticity has gained the most attention is that of vascular disease progression, with the most prominent human condition being arteriosclerosis. Interest in this area has been fueled by the alarmingly high mortality rate in humans afflicted with this disease. The American Heart Association reports that in 2004, cardiovascular diseases accounted for nearly 37% of all deaths in the United States of America, whereas more than 27% of all U.S. citizens exhibit symptoms of the disease, including high blood pressure, coronary artery disease, peripheral vessel disease, stroke, and heart failure. In the period 1994 to 2004, deaths related to cardiovascular disease decreased by nearly 25%, suggesting that efforts aimed at understanding the pathological progression of this disease are beneficial (statistics obtained from [www.americanheart.org](http://www.americanheart.org)). Despite this positive progress, deaths related to coronary artery disease continue to be the leading cause of death among cardiovascular disease-associated conditions.

Atherosclerosis is a complex disease, as indicated by the historical shift in the mechanistic view of atherosclerotic progression, or atherogenesis. Early perceptions of atherogenesis were that it was precipitated by continual deposition of circulating low-density lipoproteins, particularly those rich in cholesterol, on the vessel wall. This view has changed substantially to one in which atherogenesis is now considered to be a response and overcompensation to vascular injury, and as such, is primarily an inflammatory disease (174, 232, 233). The nature of causative agents is also hotly debated and may be a variety/combination of factors including hypercholesterolemia, dyslipidemia, hyperhomocysteinemia, hypertension, diabetes, oxidative stress, infection, genetic predisposition, and/or trauma, all resulting in either denudation or dysfunction of the endothelium (232). Maturation of atheroma is the result of a vast combination of contributing events and factors including initial insult on resident endothelial and VSMCs, plasma proteins, cellular blood components, oxidized lipoproteins, and cellular inflammatory mediators such as lymphocytes and monocytes. Contrary to early belief, mature atheroma are highly cellular structures, consisting of a fatty core composed of lipidated macrophage or foam cells, layers of smooth muscle, and if progressed, a fibrous cap. The response of resident medial VSMCs and adventitial (myo)fibroblasts to promote morphological changes at sites of atherogenesis is collectively referred to as vessel remodeling, and employs phenotypic modulation capabilities of these cells. It should be noted that vascular wall remodeling exists in other pathophysiological conditions, namely post-angioplasty restenosis, and venous graft transplant vasculopathy, where endothelial disruption occurs and injury responses ensue.

To understand the role of VSMC phenotype plasticity in atherogenesis, it must first be put into the context of the sequence of events that proceed after primary insult to the vessel (the following description has been extensively reviewed in (174, 232)). In response to vascular injury, for example, disruption of vessel lumen endothelium, endothelial cell activation and resultant exposure of collagenous extracellular matrix surfaces recruits a host of circulating cells and cell particles (platelets) to the site of injury. The most important of these are monocytes which attach to the endothelium via interaction with adhesion molecules (vascular cell adhesion molecule-1, intercellular adhesion molecule-1) presented by the endothelium upon disruption and initiation of early inflammatory events, or are prone to localization in areas of turbulent and reduced blood flow. Proinflammatory, chemoattractant cytokines (monocyte chemoattractant protein-1) expressed in the subendothelium stimulate the migration of monocytes through the endothelial layer and occupation of the intima. Upon intimal residency, monocytes assume macrophage phenotypes, scavenge lipids and aid in the propagation of the inflammatory process by further secretion of inflammatory cytokines that recruit T-lymphocytes to the atheroma and activate endothelial cells and VSMCs. T-lymphocytes and activated vascular cells then amplify the response by presentation of cytokines and growth factors that cause phenotypic modulation of medial VSMCs. It is this dedifferentiation that permits vascular wall remodeling: migration through the elastic laminae to the intima, proliferation, and population of “newly formed” vessel wall, or what is termed the neointima. However, due to the popular view that neointimal vessel occlusion is a significant cardiovascular complication, phenotypic plasticity of VSMCs in

the context of atherosclerosis is often considered a contributor to disease progression; however more recent findings have implicated the phenomenon as beneficial in stabilizing vulnerable atheroma from rupture and subsequent thrombosis (2, 61).

A closer look at the phenotypic modulation or activation of VSMCs during the course of atherogenesis shows vast reprogramming of gene expression which we will now consider. Comparative analyses looking at medial VSMCs from normal vessels and neointimal VSMC or VSMC-like cells have highlighted some of the prominent molecular, morphological, and functional differences between differentiated and phenotypically modified VSMCs *in vivo* that are the result of reprogramming of over 140 genes (96).

As stated before, differentiated VSMCs have been traditionally characterized as contractile and quiescent (nonproliferative and nonsynthetic), and as expected, VSMCs undergoing dedifferentiation exhibit alternate phenotypes. By-and-large, the most widely used markers of contractile phenotypes are the expression levels of SM $\alpha$ A and SM-MHC, and as such reduced levels of these proteins are observed in migratory and proliferative VSMCs compared to controls (3, 150, 165, 199). It is the feeling of some researchers, however, that SM-MHC is a more appropriate marker of VSMC activation, as a more drastic reduction in SM-MHC expression is observed in neointimal VSMCs compared to that observed for SM $\alpha$ A (3). Loss of contractile protein expression has also been shown to correlate with loss of contractile function *in vivo* (71), validating the use of the molecular approaches to gauge the extent of VSMC activation.

Closely related to contractile function at the molecular level is migratory capacity of activated VSMCs, and cell morphology. Increased migratory (chemotactic) capacity of dedifferentiated VSMCs has been a hallmark of VSMC differentiation status both *in vivo* (248), and as a diagnostic *in vitro* (5, 171, 320). A distinctive morphological characteristic of differentiated VSMCs is their elongated spindle shape dictated by cytoskeletal filament arrangement. Not surprising is the finding that loss of cytoskeletal protein expression accompanying phenotypic modulation coincides with rounding of the cell ultrastructure (150, 199).

Extensive genetic reprogramming is also observed, and has been inferred as necessary, for proliferative properties in activated VSMCs. Specifically, increases in cyclin and proliferative cell nuclear antigen (PCNA) expression (100), and DNA synthesis has been observed in coronary atherosclerotic plaques *ex vivo* (100), and in cultured cell models *in vitro* (137, 299).

The array of cytokines, cell mediators, reactive oxygen species, lipids and lipid products that act as effectors of VSMC dedifferentiation are vast, however a few standouts that have profound effects on this process. PDGF isoforms, namely the BB homodimer, as noted above promotes transdifferentiation of fibroblasts toward a myofibroblasts phenotype (283), causes downregulation of numerous VSMC differentiation markers including SM $\alpha$ A (16, 17, 52), smooth muscle myosin heavy-chain (SM-MHC) and smooth muscle  $\alpha$ -tropomyosin (SM- $\alpha$ TM)(117). Thus, a disparity on the effects of PDGF-BB on contractile apparatus proteins in different cell types has been pointed out. However, exposure of cultured VSMCs and myofibroblasts to PDGF-

BB has shown this factor to be a chemoattractant, promoting migration of both cell types *in vitro* (5, 170, 320). Whereas TGF- $\beta$ 1 promotes SM $\alpha$ A and SM-MHC expression in fibroblasts (64), this potent growth factor displays prosynthetic effects when over expressed in medial VSMCs *in vivo*, causing extensive collagen deposition and pronounced neointimal growth (247). A protective role for TGF- $\beta$ 1 in fibrous and stable cap formation has since been verified in mice treated with TGF- $\beta$ 1-neutralizing antibodies (185). Interestingly, both PDGF (42) and TGF- $\beta$ 1 (181) cause upregulation of matrix metalloproteases (MMPs) and downregulation of tissue inhibitors of MMPs (TIMPs) in cultured SMCs. In normal tissues, matrix remodeling is held in check by a MMP:TIMP ratio less than one. A positive shift in the MMP:TIMP ratio has been linked to increases in neointimal hyperplasia and plaque vulnerability in late stages of disease progression, but is generally regarded as necessary for initial stages of vessel remodeling (78, 168). The notion of matrix remodeling by MMPs propagating further phenotypic modulation of VSMCs in lesion areas has been hypothesized (210) but not demonstrated.

The large numbers of signaling molecules present within atherosclerotic lesions have profound and varied effects on expression of genes necessary for maintenance and modulation of VSMC phenotypes. It is in this manner that the continuous spectrum of VSMC-like and myofibroblasts-like phenotypes are generated *in vivo*. Difficulties in assessing the extent of differentiation of myofibroblasts and VSMCs (and endothelial cells and fibroblasts, for that matter) and the origin of these cells in injured or diseased tissue stems from the continuous, diverse, and combinatorial spectra of marker expression. This difficulty has prevented the establishment of an index of differentiation

for VSMCs and myofibroblasts. The fact that marker expression is not exclusive to these cells compounds the difficulties encountered by researchers, however, measurement of marker expression remains the best suited tool for studying phenotypic modulation. An accepted caveat in the field is that the profile of marker expression ultimately dictates the phenotypic state of VSMCs and myofibroblasts, and probably holds true for all cell types. Thus, an understanding of the mechanisms controlling marker expression is critical.

## **FUNCTIONAL ASPECTS OF SMOOTH MUSCLE $\alpha$ -ACTIN IN SMOOTH MUSCLE CELL BIOLOGY**

Absolute verification of VSMC and myofibroblast lineage by detection of a single marker is both inappropriate and impossible (210). As mentioned above, detection of SMC-like phenotypes has been routinely carried out by analysis of marker expression fingerprints. Batteries of marker-specific antibodies and marker mRNA-specific primers have facilitated this pursuit, and have been employed literally hundreds of times in the literature (for examples see (25, 55, 76, 120, 126, 186, 189, 218, 242, 302), reviewed extensively in (210)). Among the mentioned markers for detection of SMC lineage, SM $\alpha$ A is one of six actin isoforms in mammals, and is the most abundant protein in VSMCs and is a major component of the cytoskeletal foundation on which contractile force is generated (257). SM $\alpha$ A has been noted to account for ~40% of total cellular protein in differentiated VSMCs (79), and resultant ease of detection has made it the most prominent marker for SMC and myofibroblasts differentiation despite the fact that

expression of SM $\alpha$ A is not restricted to VSMCs and myofibroblasts (83, 309). SM $\alpha$ A has been determined to be the principle contractile protein expressed by activated myofibroblasts at sites of wound closure, pathophysiological fibrosis, and neoplastic stromal response (226, 231, 271).

Gene knockout studies have substantiated the claim that SM $\alpha$ A is absolutely required for proper vessel dilation for blood pressure homeostasis, as null mice display defects of vascular contractility (243). Additionally, SM $\alpha$ A expression and contractile function is necessary in wound closure by myofibroblasts (219, 287), and neural tube formation in early embryogenesis (12, 53, 122). In addition to contractile functions, SM $\alpha$ A expression has been shown to foster migration-restrictive properties in myofibroblasts. Electroporation of monoclonal antibodies specific for the amino-terminal epitope (Ac-EEED), which has been shown to be necessary for polymerization of SM $\alpha$ A (35, 43), decreases migratory capacity of cultured myofibroblasts in *in vitro* migration assays (231). These findings are in line with the nonmigratory characteristics of differentiated VSMCs. SM $\alpha$ A expression has also been linked to inhibition of cell proliferation capacity as drastic downregulation of SM $\alpha$ A expression coincides with transformation to tumorigenicity in mouse and rat fibroblasts (166).

SM $\alpha$ A has itself been shown to modulate the phenotypic properties of VSMCs and myofibroblasts (281), and related cardiomyocytes (44, 271). Although unequivocal verification of SMC lineage by detection of SM $\alpha$ A is not possible, the phenotypic altering properties of this protein have made the regulation of SM $\alpha$ A expression a topic

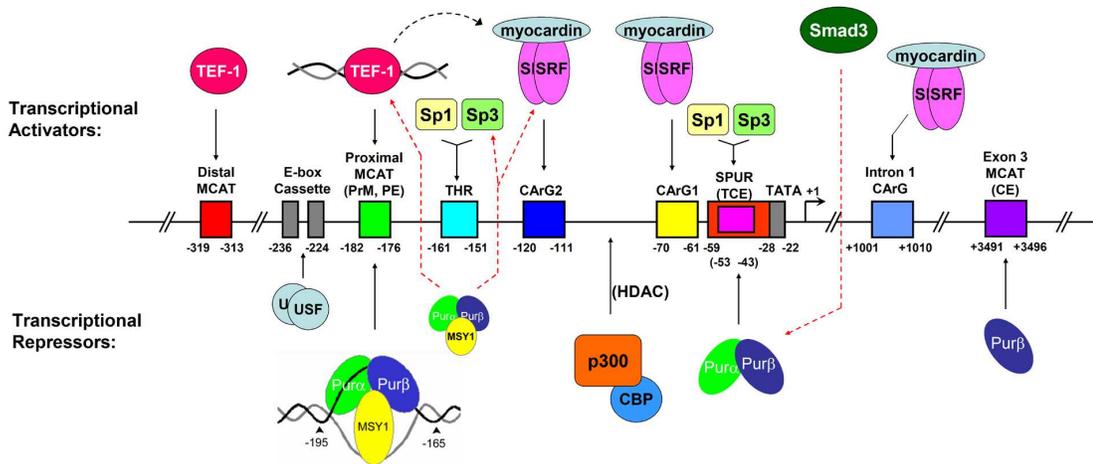
of intense study, as well as a focus of this dissertation, and shall constitute a considerable portion of this review.

## **REGULATION OF SMOOTH MUSCLE $\alpha$ -ACTIN GENE EXPRESSION**

Since the initial characterization of the SM $\alpha$ A promoter in the pursuit of understanding pathophysiologically specific scenarios and mechanism of gene expression, a vast array of *cis*- and *trans* regulatory elements have been identified that permit highly plastic and responsive modes of transcription. The following is a summary of the genomic and protein components of the SM $\alpha$ A regulatory network that is partially limited, by design, to those components which feature Pur protein participation or involvement. Extensive literature reviews have been published elsewhere that cover the broad scope of SM $\alpha$ A gene regulation (140, 158, 315). For a diagrammatic summary of composition, location, and designation of SM $\alpha$ A gene regulatory elements, as well as a pictorial synopsis of *trans*-acting factors and their reported regulatory functions, refer to Figure 1.1.

Initial investigations aimed at identifying regulatory elements of the SM $\alpha$ A gene promoter were geared towards surveying the 5' upstream region of the gene for sequences necessary for activation of expression. The first steps made in this effort were the isolations of genomic promoter sequences from chicken (29), mouse (194), and humans (225). These studies were quick to point out extensive homology of at least two (depending on species) *cis*-regulatory elements bearing high resemblance to CArG

(CC(A/T)<sub>6</sub>GG) boxes, that had been noted previously in the  $\alpha$ -cardiac actin gene 5' flanking region (195, 197). Demonstrable interaction of  $\alpha$ -cardiac actin promoter CArG boxes with *trans*-activating factors (196) underpinned the need to study the role of CArG box *cis*-regulatory elements in the SM $\alpha$ A promoter. Similar to what is observed for  $\alpha$ -cardiac actin, SM $\alpha$ A promoter CArG box elements provide for activatable transcription of both chicken and mouse promoter-reporter constructs transfected in cultured SMCs (15, 194), however this effect has been determined to be cell type-dependent (30). Serum requirements for CArG-box mediated activation of SM $\alpha$ A expression in fibroblastic cell lines suggest that these elements are downstream targets of serum growth factors (147, 270). Furthermore, interruption of serum stimulation by *c-fos* overexpression implies that serum response factor (SRF) is responsible for *trans*-activation of SM $\alpha$ A expression by interaction with CArG *cis*-elements (147), a finding that has been verified by immunological techniques (258). It has now been established that interaction of SRF with muscle-specific (as opposed to cell growth-specific) SREs is directed by modulating factors such as Mhox (111), Nkx3.1 (31), Barx2b (116), SSRP1 (264), and/or the master smooth muscle regulator, myocardin (172, 300), as well as the position, number and precise sequence of the CArG boxes present (reviewed in (158)). The tandem nature of the proximal CArG boxes in the SM $\alpha$ A promoter has been deemed a signature promoter motif of smooth muscle-specific genes, and drives timely, tissue-specific expression in smooth muscle tissue by virtue of myocardin's capacity to modulate SRF sequence specificity (300).



**Figure 1.1. Regulation of SM $\alpha$ A gene expression involves numerous *trans*-activating and *trans*-repressing factor interactions at multiple *cis*-regulatory elements.** Activation of SM $\alpha$ A expression occurs by binding of heteromeric SRF:myocardin complexes to multiple CArG elements located both upstream and downstream of the transcriptional start site. Sp-proteins activate transcription from THR and SPUR (TCE) elements. TEF-1 activates transcription from a distal MCAT element and a cryptic MCAT element that is proximal to the transcription start site via cooperative interactions with SRF:myocardin-occupied CArG2 (black dashed arrow). TEF-1:MCAT elements interactions rely on the element assuming a double-stranded configuration. Proximal MCAT-mediated enhancement can be alleviated by structural interconversion of this region and occupation by sequence-specific SSBs Pur $\alpha$ , Pur $\beta$ , and MSY1 in a single-stranded configuration. Pur $\alpha$  and Pur $\beta$  also mediate transcriptional repression by interaction with Sp1/3 at the SPUR element. Pur $\beta$ -mediated repression of expression by interaction with an exonic MCAT element has also been reported. Furthermore, Pur $\alpha$ , Pur $\beta$ , and MSY1 may mediate repression by sequestration of factors

TEF-1, SRF:myocardin, and Sp1/3 (red dashed arrows). Pur $\alpha$  and Pur $\beta$ , themselves can be sequestered by Smad3 to preserve expression (red dashed arrow). Epigenetic control of SM $\alpha$ A expression is accomplished by chromatin rearrangement created by p300/CBP histone deacetylase (HDAC) activity within the immediate upstream CArG region. Tissue specific and cell density-responsive repression can be enforced by bHLH factors (USF) binding to upstream tandem E-box elements. Numbers represent nucleotide positions relative to the transcriptional start site.

The causative extracellular signaling component for SRF upregulation in SM $\alpha$ A expressing cell-types has been determined to be TGF- $\beta$ 1, however CARG-box involvement only accounts for partial activation. This finding led to the discovery of an additional positive *cis*-regulatory element, the TGF- $\beta$ 1-control element, or TCE, located at position -53 to -43 relative to the transcriptional start site (110). Interestingly, occupation of TCE is not accomplished by SRF in response to TGF- $\beta$ 1 exposure, but instead by Kruppel-like factors, Sp1 and Sp3, and Smad signaling proteins during activation of SM $\alpha$ A expression (47, 110). Examination of the 5' flanking region of the mouse SM $\alpha$ A promoter also identified six putative E-box sequences (CANNTG) that have been shown to bind *trans*-activating basic helix-loop-helix (bHLH) proteins involved in differentiation of cells of myogenic lineage (131, 135, 194, 304). The function for these elements appears cell type and cell density restrictive (82, 131), as combinatorial control by bHLH protein USF (upstream stimulatory factor) at a tandemized E-box cassette located between positions -236 and -224 confers repression of SM $\alpha$ A expression in rat aortic SMCs (131, 135).

When it was noted that full activation of SM $\alpha$ A expression in mouse embryo fibroblastic cell line AKR-2B required the deletion of a 33 base-pair (bp) sequence corresponding to bases -224 to -191, the realization came that this region contains a negative regulatory element (82, 270). However, direct sequence examination also concluded that the region directly 3' to this newly recognized repressive element (bases -181 to -176) contains a potent muscle-specific inverted CATTCT (MCAT) activation motif which is able to bind transcriptional enhancer factor 1 (TEF-1) in a double-stranded

configuration (48, 275) . Indeed, binding of the proximal MCAT enhancer element by TEF-1 provides synergistic activation of SM $\alpha$ A expression in collaboration with the distal CArG 2 element (48), which has been shown in previous studies to be essential for high level activation of expression by SRF (270). Deletion of the core sequence of the proximal MCAT element results in diminished expression (48). Further investigative elaboration of the function of the proximal MCAT and repressive elements, by Cogan, Getz, Strauch and coworkers, verified overlapping sequences between bases -195 and -164 that constitute both a transcriptional enhancing-, and repressive element (48). Thorough analysis of the rat and mouse SM $\alpha$ A promoters detected two MCAT elements (proximal, -182 to -176; distal -320 to -314), both of which have been shown to be important for transcriptional activation of the gene in fibroblasts, myoblasts, and aortic endothelial cells (278), however cryptic enhancer activity, as possessing both activator and repressor character, has only been established for the proximal site in the SM $\alpha$ A promoter. The work performed by Cogan et al (48), and Sun, et al (275), not only verified binding of TEF-1 to the double-stranded proximal MCAT element, but also detected binding of several *trans*-repressors that exhibit affinity for sequences adjacent to the core MCAT sequence only when in single-stranded conformations. The identities of the *trans*-acting repressors were perceived unique at the time of their discovery, and were referred to as vascular actin single-stranded binding factors 1 and 2 (VACssBF1 and VACssBF2).

A closer examination of the region encompassing the proximal MCAT enhancer also showed an interesting characteristic; this stretch of sequence possesses extensive

polypurine/polypyrimidine asymmetry. The coding (forward, non-transcribed) strand of this region shows greater than 80% purine content in the mouse and rat promoters, whereas human and chicken are slightly lower (275). This is an important finding, as was the fact that VACssBF1 and VACssBF2 display specificity for binding of the pyrimidine-rich and purine-rich strand, respectively (48, 275). The significance and repercussions of polypurine/polypyrimide tracts within gene promoters will be discussed in a following section.

MCAT enhancer sequences and their corresponding *trans*-binding factors are common in numerous promoter regions of cardiac, skeletal, and smooth muscle genes (314). These elements generally drive enhanced expression by binding of TEF-family enhancer factors, although as noted before, cell type restrictions have been pointed out (278). A mechanistic basis for differential MCAT function was examined by Gan, and colleagues (94), in which they noted that MCAT-driven SM $\alpha$ A expression follows distinct mechanisms in activated myofibroblasts compared to differentiated SMCs. Briefly, MCAT elements are essential for *de novo* expression of SM $\alpha$ A in myofibroblasts and developing embryonic SMCs, but appear dispensable for expression in adult differentiated SMCs. Furthermore, this effect also coincides with TEF-family binding activity of the MCAT elements, as knockdown of RTEF-1 (TEF-3) decreases expression of SM $\alpha$ A promoter-driven reporter expression in myofibroblasts and embryonic SMCs, whereas knockdown of all TEF-family members has no effect on reporter expression in adult SMCs. These findings speak clearly to the mechanism of SM $\alpha$ A expression during vascular development and myofibroblast transdifferentiation, and provide a means for

distinguishing myofibroblasts from dedifferentiating adult SMCs, both of which express high levels of SM $\alpha$ A. These studies fail, however, to describe the role of cryptic MCAT enhancer-mediated repression in cell types with known pathophysiological functions; explicitly activated SMCs. The location of nucleotide mutations within the proximal MCAT element used by Gan, et al. were engineered to preclude (R)TEF-1 binding, but have been shown previously to not substantially effect VACssBF binding (48, 275). The implications of SM $\alpha$ A downregulation in pathophysiological situations have been outlined above; hence factors affecting SM $\alpha$ A repression are an important area of investigation.

Delineation of molecular mechanism fostering bifacial character of *cis*-regulatory elements, as observed for the proximal MCAT enhancer element of the SM $\alpha$ A promoter, represents a challenge for genetic researchers. As stated above, mutational analysis of the coding strand of the -195 to -164 region, referred to at the time as the purine-rich motif, showed that mutations abolishing TEF-1 binding do not affect VACssBF1 or VACssBF2 binding, and mutations reducing VACssBF2 binding do not affect transcriptional activation (48, 275). These results suggest that binding of these putative activation and repression factors to particular nucleotides is not mutually exclusive *per se*, however the possibility of disparate double-stranded versus single-stranded entities could not be discounted at the time, and a mechanism by which binding of VACssBF1 and VACssBF2 to their respective strands causes disruption of the double-stranded MCAT sequence was proposed (275). This latter caveat was the focus of subsequent investigations by Kelm, Strauch and Getz and colleagues, whose primary aim was to

delineate the repressive mechanism employed by the cryptic MCAT enhancer element and associated *trans*-repressor proteins, as well as the biochemical characterization of these novel proteins (28, 47, 129, 141-145, 148, 149, 222, 272, 297, 319).

Initial experiments aimed at delineating VACssBF-mediated SM $\alpha$ A repression illustrated ssDNA-binding specificity of both VACssBFs and TEF-1 towards the proximal promoter MCAT element (promoter element, PE) versus a newly identified MCAT element positioned within exon 3 (coding element, CE) (144). These studies showed that VACssBF1 and TEF-1 shows no affinity for the CE sequence (single, and/or double-stranded) despite 100% conservation of the GGAATG TEF-1 recognition motif; however the purine-rich strand of the CE is tightly bound by VACssBF2. Additionally, replacement of the PE sequence with CE does not permit transcriptional activity in reporter based assays, consistent with lack of TEF-1 binding capacity for CE. However, introduction of mutations nullifying VACssBF2 binding rescues transcription suggesting that VACssBF2 possesses enhancer disruption activity and that this activity is independent of enhancer recognition by TEF-1. These studies also highlighted the importance of downstream regulatory elements that govern SM $\alpha$ A expression. The discovery of CArG elements within the first intron fortifies this claim (184), and the importance of downstream promoter elements for SMC-dependent expression *in vivo* has been further verified (296, 321).

Preliminary electrophoretic mobility shift studies gauging purine-rich strand binding by VACssBF2 illustrated that either this factor assembles on purine-rich ssDNA sequences to varying stoichiometric degrees, forms heterogeneous nucleoprotein

complexes, or is in-fact not a single polypeptide (48, 144, 275). The indicated possibilities stem from the observation that multiple electrophoretically shifted species are generated when cell extracts are allowed to bind single-stranded purine-rich probes derived from the SM $\alpha$ A promoter. Biochemical characterization of these complexes confirmed that VACssBF2 is in fact multiple polypeptides able to form homo- and hetero-multimers (141). Cloning of these factors led to the identification of VACssBF2 as Pur $\alpha$  and Pur $\beta$  (143), which was augmented by later work identifying VACssBF1 as MSY1 (141). The identification of these proteins was of critical importance as it showed that their respective activities characterized therein were consistent with other biological systems/scenarios in which their involvement had been detected. For example, Pur $\alpha$  (HeLa Pur Factor) had been implicated in recognizing and binding stretches of single-stranded DNA generated at origins of DNA replication proximal to the *c-myc* promoter (10). In this and subsequent studies it was observed that Pur $\alpha$  recognizes purine-rich ssDNA, especially those rich in guanine nucleotides, and was assigned a consensus (*PUR*) binding sequence GGNNAGGAGARRRR (N = any nucleotide, R = A/G) based on other known origin sequences, although Pur $\alpha$  binding activity was not substantiated at these sites (10, 11). Cloning of Pur $\alpha$  from human cDNA libraries showed the presence of a distinct isoforms, Pur $\beta$ , however no function was assigned at this time (11). In this regard, the human ortholog of MSY1, YB-1 for Y-box binding protein 1, was shown to facilitate, and bind a pyrimidine-rich ssDNA sequence important for regulation of the human *DRA* promoter (major histocompatibility complex II gene), called the Y-box (69, 183).

Consistent with early models of cryptic MCAT enhancer regulation of SM $\alpha$ A expression, Pur $\alpha$  and YB-1 were shown to cooperatively regulate expression from the human JC virus (HJCV) lytic control element (36). This finding, along with studies by Sun, et al (275), suggested that Pur $\alpha$ , and YB-1/MSY1, and possibly Pur $\beta$ , constituted an ensemble of cooperative transcriptional regulators, which by interaction was able to modulate and augment individual ssDNA-binding properties (36). As stated above, studies by Kelm and colleagues (141) investigating molecular interactions between recombinant Pur $\alpha$ , Pur $\beta$ , and MSY1, established that Pur proteins are able to bind the forward strand of PE (PE-F) as homo- and/or heteromultimers, suggesting that Pur protein self-association and Pur $\alpha$ /Pur $\beta$  association may either facilitate binding to ssDNA, represent a regulatory step in ssDNA-binding, or determine repressive activities of the PE-nucleoprotein complex. Furthermore, these studies also showed direct interaction between MSY1 with Pur $\alpha$  and Pur $\beta$ .

To test the hypothesis that sequence-specific ssDNA-binding proteins (sequence-specific SSBs) are responsible for preventing MCAT occupation by TEF-1, Carlini, et al, systematically analyzed the ability of SM $\alpha$ A promoter constructs harboring mutations that selectively inhibit binding of Pur $\alpha$ , Pur $\beta$ , and MSY1, but not TEF-1, to drive expression of a reporter gene (28). These studies showed that loss of sequence-specific SSB-binding releases repressive effects of the cryptic MCAT enhancer, and that drastic reduction of Pur protein binding can be induced by deletion of nucleotides -194 to -192, highlighting the importance of these nucleotides in Pur protein-ssDNA nucleoprotein complex formation. Additionally, these studies also showed that Pur $\alpha$ , Pur $\beta$ , and MSY1

possess the capacity to interact with TEF-1, SRF, and Sp3, as well as with double-stranded PE, albeit weakly, by virtue of low-stringency DNA pull-down techniques. This data supports the concept that sequence-specific SSB-mediated repression of the MCAT enhancer element occurs by protein-protein interaction masking effects, although observable indirect association (DNA or ternary complex-mediated) could not be discounted by this approach. Nonetheless, opposing competitive and masking models for cryptic MCAT enhancer regulation by Pur $\alpha$ , Pur $\beta$ , and MSY1 were proposed (28), however, a thorough experimental attempt to discredit either of these models has not yet been performed. The implications of detectable interaction of Pur $\alpha$ , Pur $\beta$ , and MSY1 with SRF and Sp3 are that sequence-specific SSB-mediated repression of SM $\alpha$ A expression may be accomplished by disruption of SRF and Sp3 *trans*-activation properties by virtue of protein-protein interactions. On a similar note, Pur $\beta$  has been shown to competitively disrupt muscle-specific CArG box binding, and not *c-fos* CArG box binding, by SRF and gene activation in cardiac muscle gene expression (103).

A consistency in the literature surrounding Pur $\alpha$ , Pur $\beta$ , and YB-1/MSY1 function in the mechanism of cryptic MCAT enhancer element regulation of SM $\alpha$ A is that these proteins are able to function as sequence-specific SSB transcription factors in a predominantly dsDNA genome despite exhibiting low affinity for dsDNA. A general supposition for sequence-specific SSB activity at gene promoter sequences is the coincident existence of structural perturbations within the DNA duplex structure that are either created by virtue of SSB binding or facilitated by auxiliary factors to provide binding sites for SSBs. Widespread dynamic structural rearrangements have been noted

in promoter regions of numerous protein encoding genes, notably *c-Myc* (8, 75, 153, 192, 288), platelet-derived growth factor A-chain (298, 301), vascular endothelial growth factor (274), tyrosine kinase pp60<sup>c-src</sup> (229), high mobility group protein A (235), insulin receptor (285), androgen receptor (39), and epidermal growth factor receptor (130). The distribution of non-B-DNA structures *in vivo* appears to be non-random, as it is limited to specific genes, although transcriptional activity is not universally coincident with structural alterations (33, 161). Observed non-B-DNA, or paranemic structures within gene promoters include ssDNA, slippage structures, cruciforms, (left handed) Z-DNA, (triple helix) H-DNA, quadruplexes, and protein stabilized paranemic structures (313). The observation of non-B-DNA structures is most common in sequences of DNA possessing tracts of asymmetrically distributed nucleotides (A/T or G/C rich), polypurine/polypyrimidine, alternating purine/pyrimidine, or dinucleotide repeats (204, 227, 235). Often intervening or non-B-DNA/B-DNA transition nucleotides exhibit stable ssDNA character (105) detectable by use of ssDNA-specific reagents (9, 130, 192, 229, 235, 274, 285, 298, 301). Physical analysis has shown that topological stress further facilitates B-DNA to non-B-DNA structural conversions (204, 260), as well as localized duplex melting *in vitro* (153) and *in vivo* (154), and that non-B-DNA structures typically possess lower melting temperatures (227, 249). The nature of topological stress *in vivo* is believed to be negative supercoiling which is generated upstream of transcriptional machinery by associated helicase-mediated unwinding occurring downstream of transcription (154, 179, 260). The reality of localized melting of promoter sequences is exemplified not only by ssDNA-specific reagent sensitivity, as mentioned above, but also

by the involvement of sequence-specific SSB transcription factors that regulate expression at these and other promoters. Collectively, these findings support the importance of non-B-DNA structures in mechanisms of gene regulation.

The existence of the proximal MCAT enhancer of the  $SM\alpha A$  promoter embedded within a region of extensive purine/pyrimidine asymmetry coupled with the involvement of sequence-specific SSBs in the regulation of this element fueled speculation of DNA structural interconversion as being a possible component of  $SM\alpha A$  transcriptional regulation. As suspected, detectable single-stranded character is observed within the vicinity of the proximal MCAT enhancer element of the  $SM\alpha A$  promoter (9). Treatment of cultured AKR-2B mouse embryonic fibroblasts with TGF- $\beta$ 1, a serum factor well established to induce  $SM\alpha A$  expression in fibroblasts (64), causes transient changes in the sensitivity of genomic  $SM\alpha A$  promoter DNA to reagents that preferentially react with unpaired or unstacked nucleotides, including chloroacetaldehyde and potassium permanganate, as assessed by ligation-mediated polymerase chain reaction (PCR) techniques. Specifically, hypereactivity in bases immediately upstream of the core MCAT sequence is observed upon activation of  $SM\alpha A$  expression, signifying induced vacancy of this region by dissociation of sequence-specific SSBs at nucleotides deemed to be necessary for binding of these factors (28, 48, 275). As a result of TGF- $\beta$ 1-induced hyper-reactivity, the region encompassing the core MCAT element and adjacent sequences was referred to as the TGF- $\beta$ 1 hypereactive region, or THR (9). The importance of this region proximal to the cryptic MCAT enhancer element was verified by responsiveness to TGF- $\beta$ 1 treatment in AKR-2B and *trans*-activation by Sp1/3,

similar to what was observed for TCE (47). It has been noted that upon TGF- $\beta$ 1 treatment of cultured human pulmonary fibroblasts, YB-1 dissociates from the pyrimidine-rich strand of the SM $\alpha$ A MCAT enhancer element and shuttles to the cytosol as SM $\alpha$ A messenger ribonucleoprotein complexes (mRNPs) via a mitogen activated kinase pathway (319) and possible C-terminal processing mechanism (269). A similar mechanism for Pur protein shuttling has been proposed, but not published (A.R. Strauch, personal communication), as has been mRNP involvement by Pur proteins. This latter point has been hypothesized based on observations of Pur proteins binding to the CE in the 5' untranslated region (UTR) of reporter mRNAs causing reduced levels of translation (142). Thrombin treatment of human pulmonary fibroblasts induces dissociation of Pur $\alpha$ , Pur $\beta$ , and YB-1 from mRNPs and subsequent shuttling of these proteins back to the nucleus, thus permitting fast translation of SM $\alpha$ A transcripts and thin filament assembly (319). These findings have direct implications in the role of TGF- $\beta$ 1, thrombin, and sequence-specific SSB/RNA-binding proteins Pur $\alpha$ , Pur $\beta$ , and MSY1 in excessive myofibroblast differentiation and subsequent destructive tissue remodeling, and may represent a permissive mechanism for SM $\alpha$ A derepression.

Pur protein involvement in regulation of SM $\alpha$ A gene transcription and translation has been detected outside of cryptic MCAT enhancer repression and mRNP sequestration. As noted before, serum-responsive *cis*-elements beyond the non-canonical CArG elements contribute to SM $\alpha$ A gene regulation. Namely, *trans*-activation of gene expression by binding of Sp1/3 to the TCE and THR elements has been demonstrated (47, 110). Examination of sequences flanking the TCE identified an overlapping purine-

rich subdomain similar to what is observed in the THR, suggesting cryptic character of this *cis*-element as well. Examination of the ability of this sequence to interact with Pur proteins verified not only binding capacity, but also that occupation of this element by sequence-specific SSBs can occur in a double-stranded configuration in cell extracts. Furthermore, overexpression of Pur $\alpha$  or Pur $\beta$  reduces reporter expression from a SM $\alpha$ A promoter construct in which the proximal MCAT element had been deleted, suggesting that this element, designated as the SPUR element (Sp1/3 – Pur protein), confers both positive and negative regulatory functions *in vivo* (272). Additionally, TGF- $\beta$ 1 exposure is coincident with dissociation of Pur proteins from SPUR, as a detectable Pur protein:Smad2/3 complex, thus demonstrating physical interaction of Pur $\alpha$  and Pur $\beta$  with Smad proteins, and elucidating a regulatory mechanism for sequestration of repressors in SM $\alpha$ A activation during myofibroblast differentiation and tissue remodeling (272). What's more, SRF has exhibited potential to circumvent Pur $\alpha$ -mediated repression at SPUR in stressed adult cardiac myocytes undergoing SM $\alpha$ A reprogramming (271, 318). This capacity of SRF to neutralize Pur $\alpha$  repression comes from its ability to form a SRF:Pur $\alpha$  heteromeric complex. Similarly SRF-overexpression has been shown to circumvent Pur $\beta$ -mediated repression of SM $\alpha$ A (145), thus underlying the importance of protein-protein interactions and the dynamic interplay of *trans*-regulatory factors in regulating SM $\alpha$ A expression during phenotypic modulation by a variety of cell types.

The complex nature of SM $\alpha$ A promoter regulation by the involvement of numerous *cis*-regulatory elements, and diverse interactions with a multitude of interacting *trans*-

acting factors epitomizes eukaryotic gene expression. However, additional layers of epigenetic regulation have been identified at the level of chromatin modification as well. Extensive chromatin histone acetylation and concomitant SRF occupation in the -150 to -50 (CArG1 and CArG2) region of the genomic SM $\alpha$ A promoter is observed in SMC-lineage restricted cell types *in vivo* (189). Interaction of SRF with myocardin augments association of SRF to acetylated histones during gene activation, and deacetylation of histone H4 is coupled to SM $\alpha$ A repression in response to vascular injury. Adenoviral E1A cotransfection experiments in AKR-2B mouse embryonic fibroblasts and rat smooth muscle cells confirmed that specific targeting of the CBP/p300 family of histone acetyltransferases and pRb pocket proteins causes SM $\alpha$ A promoter inactivation in a *cis*-element and cell cycle-dependent fashion, thus implicating these proteins in epigenetic and cell cycle-dependent regulation of SM $\alpha$ A expression (297).

In conclusion, the regulated expression of SM $\alpha$ A expression in various cell types is the consequence of a diverse and extensive protein-protein and protein-DNA interaction ensemble that, in turn, permits highly plastic expression of this important functional filamentous protein. Adaptability of SM $\alpha$ A expression is a crucial component for cellular responses to physiological and pathological stimuli that impart either beneficial or malevolent phenotypic consequences. The documented involvement of Pur proteins in pathophysiological SM $\alpha$ A repression makes them suitable targets for intense biochemical study.

## **PUR PROTEIN STRUCTURE AND FUNCTION**

The Pur family of proteins is comprised of a group of functionally-related, highly homologous DNA/RNA-binding proteins, and consists of four members in mammals, Pur $\alpha$ , Pur $\beta$ , and the two isoforms, Pur $\gamma_a$  and Pur $\gamma_b$ , however in lower eukaryotes, namely *Drosophila melanogaster*, multiple Pur $\alpha$  isoforms have been detected (118). Despite the fact that a relatively high amount of knowledge regarding Pur $\alpha$  and Pur $\beta$  structure and function has accumulated in recent years, very little is known about corresponding properties in Pur $\gamma_{a,b}$ . The focus of the following section is a review of Pur $\alpha$  and Pur $\beta$  structure/function characteristics, as this pertains to the scope of the work presented here.

## **THE BREADTH OF PUR PROTEIN FUNCTION**

Much of what we know about structure and function of Pur proteins comes from their involvement in diverse cellular events, with binding of nucleic acids being a common aspect in these processes, whether direct or indirect. Discovery of Pur $\alpha$  was the result of a survey of proteins that were believed to be responsible for enforcement of structural perturbations that are observed in origins of DNA replication in HeLa cells (10). Investigators examining a zone of DNA replication neighboring the *c-myc* promoter in humans identified a protein that binds a purine-rich element (so called, *PUR*-element) at a site of DNA bending that, presumably, causes disruption of helix conformation and localized melting of strands. It was also noted that this protein displays specificity for purine-rich ssDNA sequences, endorsing the claim that bending induces unpairing of

complementary strands in this region of genomic DNA, or permits helix disruption by occupancy of SSBs. Methylation interference patterns showed that this protein forms specific contacts with guanines in ssDNA sequences representative of the *PUR*-element, and similar purine-rich motifs were identified within other known zones of DNA replication, suggesting the importance of these sequences and the *PUR*-binding factor in the initiation of cellular DNA replication (10, 259) as well as in the replication of viral genomes (36, 136, 155, 238). Cloning and sequencing of the so-called *PUR*-factor led to the identification of Pur $\alpha$  (11).

Numerous roles for Pur $\alpha$  have been elucidated in regards to regulation of cell cycle progression. Aside from cellular replication origin interactions described above, clues to further involvement in cell cycle control came from examination of the replication of viral genomes. Regulation of lysogeny in the human JC polyoma virus (HJCV) has been shown to require the differential and reciprocal activities of Pur $\alpha$  and YB-1, modulated by interactions with the JC tumor antigen (36). The competition between opposing activities of Pur $\alpha$  and YB-1 dictates entry into the lytic cycle, with Pur $\alpha$  maintaining lysogeny by governing expression of early genes via displacement of YB-1. Modulation of YB-1 ssDNA-binding affinity by HJCV tumor antigen association causes disruption of Pur $\alpha$ :promoter interactions, expression of late genes, and commitment to the lytic cycle (238), whereas association of Pur $\alpha$  and HJCV tumor antigen results in attenuation of T-antigen-mediated transcriptional activation of viral genes necessary for lytic entry (93). Overexpression of Pur $\alpha$  suppresses replication of HJCV genomic DNA in infected glial cells (34), but has the opposite effect on HIV-1 genome replication (41). Interestingly,

the capacity of coinfecting HIV Tat protein to bind, sequester and modulate Pura $\alpha$  sequence specificity to activate late gene expression and impose lytic entry by HJCV supports the notion that partitioning of Pura $\alpha$  activity is important for regulation of transcription of viral genes and viral DNA replication (155).

The repertoire of Pura $\alpha$ -viral protein interactions utilized in the regulation of viral genome replication suggested similar mechanisms may be employed by cells during regulation of cell cycle progression. Examination of Pura $\alpha$  primary structure identified a signature motif utilized by several cell cycle regulators for protein-protein interaction (11, 182). Namely, the presence of a C-terminal “psycho” motif in Pura $\alpha$  suggested this protein might interact with the tumor suppressor retinoblastoma protein, pRb, a factor known to cause G<sub>0</sub> cell-arrest when normally expressed or loss of growth control when mutated or deleted. Indeed, direct interaction of Pura $\alpha$  with hypophosphorylated pRb was established in a psycho motif-dependent manner, and this interaction modulates the binding affinity of Pura $\alpha$  for its ssDNA recognition element in the *c-myc* origin of replication. The implications of Pura $\alpha$  and pRb association were not realized until studies were performed that correlated Pura $\alpha$ :pRb complex level alterations with growth phase entry suggesting that this complex may aid in preventing assembly or processivity of replication machinery from origins of replication. Levels of Pura $\alpha$  significantly drop just prior to the onset of S-phase, and return prior to mitosis in CV-1 fibroblast cell lines (125). This is consistent with findings that show microinjection of NIH3T3 mouse fibroblasts with Pura $\alpha$  causes cell cycle arrest in populations in early S phase and G<sub>2</sub>

(265), and that overexpression of Pur $\alpha$  inhibits *Ras*-induced colony formation in NIH3T3 cells (7) similar to its effects on HJCVCV (34). Furthermore, the subcellular localization of Pur $\alpha$  is governed by either pRb or cyclin A. Detectable interaction of Pur $\alpha$  with cyclin A, via interaction with a cyclin A/Cdk2 ternary complex, has been shown to stimulate the phosphorylation of histone H1 by cyclin A/Cdk2 *in vitro* (177), suggesting that cyclin A-sequestration of Pur $\alpha$  may result in permitting S-phase entry and facilitation of chromatin decondensation necessary for replication. Pur $\alpha$  may also play dominant-negative roles on other regulatory proteins, as association of Pur $\alpha$  with transcription factor E2F-1 suppresses E2F-1-induced activation of S phase-specific genes necessary for cell cycle progression (58). Collectively, these studies point out that Pur $\alpha$  exerts negative regulation of cellular and viral DNA replication, although mechanistic aspects of Pur $\alpha$  protein-protein and protein:ssDNA interactions remain uncertain, as do possible parallel roles for Pur $\beta$ .

Since its discovery, Pur $\alpha$  has been implicated in numerous aspects of nucleic acid processing aside from DNA replication, none more important perhaps than its involvement in gene regulation. The verification of Pur $\alpha$  involvement in transcriptional regulation is the direct result of identification of gene regulatory *cis*-elements that bear resemblance to purine-rich *PUR*-elements. The first identification of Pur $\alpha$  as a transcriptional regulator was in the regulation of the clusterin gene in quail infected with the Rous sarcoma virus, showing that this protein not only was important for mammalian DNA replication (182), but is also conserved for transcription regulation in numerous

vertebrate classes (115). Involvement of *PUR*-elements and Pur $\alpha$  interaction is especially widespread in the expression of genes important for neuron development and function. For example *PUR*-elements and Pur $\alpha$  binding has been identified in promoter regions of protein FE65 (316), neuronal nicotinic acetylcholine receptor (74), and myelin basic protein (107), where Pur $\alpha$  activates transcription. Contrastingly, Pur $\alpha$  has also displayed negative regulatory functions as a neuronal transcriptional repressor. For example, Pur $\alpha$  represses expression of the neuronal and hematopoietic transcription factor *Gata2* in the central nervous system, which is opposed by the *trans*-activator Sp8 (214). Similarly, cAMP-response elements of the somatostatin and tyrosine hydroxylase gene promoters in opiate-exposed neuronal cell line extracts are suppressed by Pur $\alpha$  in a *PUR*-element-dependent manner (70, 207, 237). Autoregulation of Pur $\alpha$  expression in glial cells has also been reported, as *PUR*-element-dependent repression of Pur $\alpha$  promoter activity by ectopic expression of Pur $\alpha$  has been observed (200).

Tightly controlled expression of Pur $\alpha$  during neuronal development suggests that gene activation afforded by Pur $\alpha$ :promoter interactions is crucial to development of functioning neurons (107). Exemplification of this trend is provided by mouse models harboring nullifying homozygous mutations in Pur $\alpha$  alleles (*PURA*<sup>-/-</sup>) which exhibit severe postnatal neurological defects and eventual fatality due to lack of neuron population, neuron myelination, and neurofilament assembly (146), moreover heterozygous deletions of both Pur $\alpha$  and Pur $\beta$  have been noted in human patients with acute myeloid leukemia (169). In addition to temporal expression and promoter

interaction of Pur $\alpha$ , protein-protein and protein-RNA interactions appear to be a crucial component of Pur $\alpha$ -mediated gene expression. For example, temporal expression, and interaction of Sp1 with Pur $\alpha$  augments Pur $\alpha$ :promoter interactions and expression of myelin basic protein (MBP) in neuronal cells (291). Association of HJCIV tumor antigen with Pur $\alpha$  causes a downregulation of MBP expression, and subsequent hypomyelination of infected, non-lytic brain cells *in vivo* (290). Translocation of Pur $\alpha$  from the nucleus to the cytoplasm during neuronal development by virtue of interactions with three distinct, yet uncharacterized, Pur-binding proteins are key to developmental timing in mouse brains (317), suggesting that compartmentalization of Pur proteins is a mechanism utilized by cells to partition transcriptional/translational regulatory activities of these proteins in accordance to what is observed for DNA replication-governing by Pur $\alpha$  (125).

As discussed extensively in a previous section, Pur proteins have been implicated in the transcriptional regulation of numerous genes, outside the realm of neuron-specific proteins. For example, Pur $\alpha$  and YB-1, in conjunction with Pur $\beta$  and AP-1 (c-Fos and c-Jun), regulate expression of the cell death-associated surface protein, Fas, in a complex and dynamic manner. Specifically, Pur $\alpha$  and YB-1 overexpression results in repression, which can be antagonized by Pur $\beta$ , suggesting that either co-association of these proteins modulates DNA-binding properties of Pur $\alpha$ :YB-1 complexes or that Pur $\beta$  competes for promoter sites resulting in derepression (162). These studies also highlight functional differences between Pur $\alpha$  and Pur $\beta$  in mechanisms of gene regulation. Pur $\alpha$  has also been shown to repress expression of CD43 in activated leukocytes in conjunction with

heterogeneous nuclear ribonucleoprotein K (hnRNP-K) (54), and the expression of the nicotinic acetylcholine receptor  $\beta 4$  subunit in neurons (74), also in conjunction with hnRNP-K, Sp1, Sp3, and Sox10 (190). Similarly, Pur $\alpha$ -Tat heteromeric complexes repress TNF $\alpha$  expression in HIV-1-infected glial cells (59). The implication of Pur $\alpha$  and Pur $\beta$  in repression of SM $\alpha$ A expression, as discussed extensively in a previous section, has also created a focus on these proteins in regulating other muscle-specific genes. Similar to SM $\alpha$ A expression, Pur $\alpha$  and Pur $\beta$  negatively regulate expression of cardiac  $\alpha$ -myosin heavy chain ( $\alpha$ -MHC) via interactions with a purine-rich regulatory element in the first intron which is essential for cardiac-restricted expression, and by direct interaction with the transcript thus regulating translation of the mRNA in the cytoplasm (103, 104). Contrary to repressive roles exhibited by Pur proteins in cellular gene expression, Pur $\alpha$  has shown the capacity to activate transcription of the PDGF-A gene via interactions with a purine-rich, and S1 nuclease-sensitive region of the promoter (322).

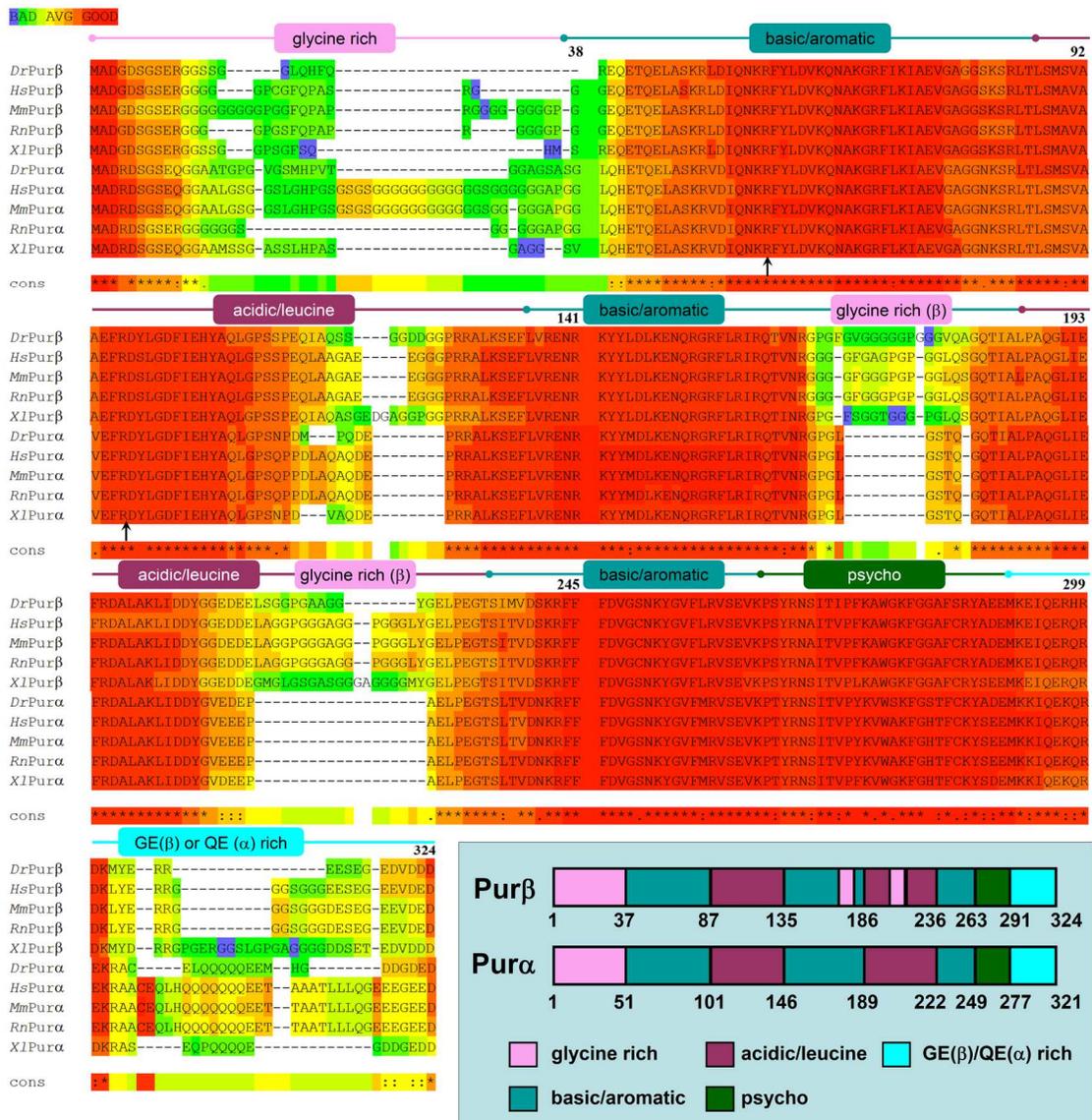
As indicated, several studies have implicated Pur proteins in transcriptional and translational regulatory mechanisms. In addition to SM $\alpha$ A and cardiac  $\alpha$ -MHC translational regulation, in which binding of mRNA by Pur $\alpha$  and Pur $\beta$  (and MSY1 in the case of SM $\alpha$ A mRNA) attenuates translation of the mature transcripts (104, 142), Pur protein involvement in the translational regulation of numerous mRNAs has been identified. Pur $\alpha$  and Pur $\beta$  were identified as two of 42 proteins involved in kinesin-associated large mRNA transport granules (138, 206). A multitude of interactions

between Pur $\alpha$  and other proteins with extensive mRNP involvement, namely YB-1/MSY1 (50, 77, 87, 112, 188) and hnRNP-K (190, 286), have fueled speculation as to whether Pur proteins are intimately involved with widespread translational regulation. It has been speculated that YB-1 at low concentrations destabilizes mRNA tertiary structures to permit ribosomal access and assembly on the template, whereas higher concentrations may restrict access (188). It is unknown whether similar mechanisms are employed by Pur proteins. Copurification of Pur $\alpha$  with 18S ribosomal RNA prompted the discovery of Pur $\alpha$ 's ability to limit translation in reconstituted assays in a dose-dependent manner (90). Validation of this theme has been noted in neuronal cells in which association of Pur $\alpha$  with ribosomes in mouse brain homogenates increases during postnatal brain development (173). RNA-mediated activity modulation of Pur $\alpha$  has also been noted. For example, association of HIV-1 Tat protein with Pur $\alpha$  is dependent upon ligand-induced reconfiguration of Pur $\alpha$  upon binding to RNA derived from the HCV *PUR*-element (311). Similarly, association of RNA homologous to the signal recognition particle (SRP) 7 SL RNA (so called Pur $\alpha$ -associated, or PU-RNA) has been shown to negatively alter Pur $\alpha$  binding affinity for the MBP promoter (289), suggesting that additional layers of regulatory activity exist within this multifunctional protein.

## **STRUCTURAL CONSIDERATIONS**

High primary sequence homology between Pur $\alpha$  and Pur $\beta$  suggests that these proteins are paralogs within given organisms and have arisen as a result of gene

duplication and divergence. This is exemplified by primary sequence alignment, as shown in Figure 1.2 (alignments performed with T-COFFEE web-based software, <http://tcoffee.vital-it.ch>, (205), using amino acid sequences deposited to the National Center for Biotechnological Information). Various regions of high sequence homology and identity can be observed between paralogs, and these similarities are present within orthologs from diverse species spanning the vertebrate branches of the phylogenetic tree, suggesting that these regions correspond to modular structural elements critical for function. Despite extensive sequence identity between paralogs (~71% in mouse), distinct regional differences can be observed, especially in the terminal regions, as well as intervening sequences between putative domains. The level of sequence conservation within regions of paralog divergence among orthologs indicates that these regions are critical for distinct paralog functions. As described above, examination of protein components that regulate and recognize origins of DNA replication led to the discovery of Pur $\alpha$ , a HeLa cell nuclear component that binds to purine-rich ssDNA sequences derived from *c-myc* and *dhfr* origins (10). Initial cloning and analysis of Pur $\alpha$  alluded to modularity and repetition of tertiary structure within the central region and led to putative domain assignments for the detected domains (11). The proposed domain architecture for Pur $\alpha$  was described, as indicated in Figure 1.2, inset. As of the date in which this review was composed, three-dimensional structure data (crystallographic, NMR, high-resolution microscopy) has not been reported for any Pur protein, nor has any low resolution analysis been presented to argue against the hypothetical domain assignments. The modular structure proposal stems from the observation of two repeating sequence



**Figure 1.2. Multiple primary sequence alignments of Pur $\alpha$  orthologs versus Pur $\beta$  orthologs.** Primary sequences alignments of Pur $\alpha$  and Pur $\beta$  orthologs from diverse vertebrate species showing sequence homology between paralogs and orthologs. Alignments were performed with TCOFFEE software (<http://tcoffee.vital-it.ch>), and alignment score color coding is explained at the top left corner. Regions of homology

follow a modular protein structural arrangement as indicated above the alignments and in the lower right panel. Positions of residues implicated in ssDNA-binding are denoted by arrows (Wortman, et al (2005) *Biochem Biophys Acta* 1743:64-78). Primary amino acid sequences were obtained from the National Center for Biotechnological Information protein database. *Dr*, *Danio rerio* (Pur $\alpha$  NP\_00101846, Pur $\beta$  Q6PHK6); *Hs*, *Homo sapiens* (Pur $\alpha$  Q00577, Pur $\beta$  AAK72642); *Mm*, *Mus musculus* (Pur $\alpha$  NP\_033015, Pur $\beta$  NP\_035351); *Rn*, *Rattus norvegicus* (Pur $\alpha$  XP\_001063244, Pur $\beta$  NP\_001017503); *Xl*, *Xenopus laevis* (Pur $\alpha$  NP\_001086909, Pur $\beta$  NP\_001079178).

motifs. The first class, referred to as class I, consists of three 23 amino acid repeats separated by the other repeating motifs, two class II repeats of 26 amino acids. Both class I and class II repeats exhibit high degrees of homology. Class I repeats are greater than 65% homologous with 17% amino acid identity, whereas class II repeats are 69% homologous with 34% of residues completely conserved. The names of these repeats have been changed in order to more completely reflect sequence character, as indicated in Figure 1.2, inset. Class I repeats are abundant in basic and aromatic residues, while class II shows high levels of leucine and acidic residues (92, 145).

The primary function that has been proposed for both Pur $\alpha$  and Pur $\beta$  is preferential binding to ssDNA that is rich in purine nucleotides, particularly guanines, and as such, has been referred to as sequence specific. This character is based on compiled findings of many groups that have examined DNA-binding properties of these proteins. Assignment of putative domain functions was carried out using Pur-deletion mutants in ssDNA/RNA-binding functional assays. In human Pur $\alpha$ , it was shown that a core domain consisting of the first two class I and class II repeats (amino acids 65-191 of the human ortholog) is necessary and sufficient for binding to short purine-rich ssDNA sequences derived from the human JC virus Mad1 control element pentanucleotide repeat (36), the *c-myc* replication origin (133, 310), MBP promoter regulatory sequence MB1 (289) and cAMP response elements (CREs) of somatostatin and tyrosine hydroxylase promoters (70). Additionally, this core region was identified to be necessary for helix-destabilization properties of Pur $\alpha$ , as assessed by the protein's ability to displace short strands of DNA complementary to regions of single-stranded M13mp18 plasmid DNA (56). Equivalent

regions in mouse Pur $\beta$  also constitute binding to ssDNA and RNA corresponding to the purine-rich strands of CE and PE (PE-F), albeit with lower affinity than the full-length protein (145). However, a difference between Pur $\alpha$  and Pur $\beta$  arises from the requirement of the third class I repeat for binding of Pur $\beta$  to ssDNA, whereas this repeat appears dispensable in Pur $\alpha$  (145). Point mutations of Arg residues within class I repeats one and two of human Pur $\alpha$  (R71E and R110E, respectively) appear to have drastic inhibitory consequences in ssDNA binding, as well as the ability of the protein to displace shorter complementary strands of linear partial duplex DNAs, suggesting that these residues may participate in ssDNA-binding via sidechain guanidinium electrostatic interactions with either the phosphate backbone or hydrogen bonding face(s) of nucleotides (presumably guanines), or by cation- $\pi$  interactions (310). Experiments evaluating equivalent point mutations in mouse Pur $\beta$  suggested, however, that these mutations may result in destabilization of tertiary and quaternary structure of the protein and do not necessarily reflect that these arginines make atomic contacts with ssDNA (149). Unequivocal identification of ssDNA-ligating amino acids in Pur proteins has not been reported. Reasons for this are clearly that the degree of sequence identity between Pur protein paralogs and orthologs has made the identification of critical residues non-obvious. Additionally, modularity may serve as a means of compensating for deleterious amino acid substitutions in regard to nucleic acid-binding functions.

The repeat region of Pur $\alpha$  has also been implicated in contributing to diverse protein-protein interactions. For example, interaction of Pur $\alpha$  to viral protein HIV-1 Tat protein, and HJCVCV and SV-40 large T-antigens is confined to amino acids 85-215 (155), and 72-

123 (93), respectively, while association of YB-1 has been localized to amino-acids 174-215 (238). The overlapping nucleic acid-binding and protein-protein interaction properties of Tat and YB-1 association regions may help explain, in part, how Tat and YB-1 enhance ssDNA/RNA-binding affinities of Pur $\alpha$  for specific sequences of the HJCV Tat-responsive element (155) and lytic control element (238), respectively. Furthermore, RNA-mediated high-affinity self-association of Pur $\alpha$  has been localized to the second class II acidic/leucine-rich repeat (91), which is interruptible by Tat association (155), suggesting that this region contributes to heterogenous protein-protein interaction and self-association in a mutually exclusive manner, and that self-association may enhance nucleic-acid binding in a way similar to Tat and YB-1 heteromeric association.

Speculation regarding the tertiary structure of Pur proteins has been limited in large-part by the fact that Pur proteins share very little sequence homology to other proteins of known structure. However close examination of the primary sequences of mouse and human Pur $\alpha$  led to the discovery that a structural motif located near the C-terminal end bears striking similarities to other cellular and viral proteins known to be involved in initiation of DNA replication. Initial analysis of this region suggested that it might adopt an amphipathic helix, based on prediction rules (11). A core motif conserved among several viral transforming T-antigen peptides is Pro-Ser-Tyr (PSY), followed downstream by a Cys (C) residue, while Pur $\alpha$  exhibits slight degeneration of this motif with PTY. Based on this character, this region was termed the psycho motif (182). Interestingly, the corresponding motif in the simian virus large T-antigen 40 protein is known to interact

with pRb. As suspected, and described above, the psycho domain of Pur $\alpha$  constitutes a pRb interaction face (133). Thus far, *in vivo* interaction between Pur $\beta$  and pRb has not been reported, despite ~92% motif homology between murine paralogs. Perhaps positions of non-conservation represent residues that contribute important atomic contacts, although this matter has not been investigated.

As mentioned above, striking differences in primary sequence are observed near the N-, and C-termini of Pur $\alpha$  and Pur $\beta$ , especially in regard to the positioning of pronounced polyglycine tracts. Pur $\beta$  contains two stretches of polyglycine of eight and nine residues separated by FQPAPR, whereas Pur $\alpha$  has a seventeen residue polyglycine stretch interrupted by a single serine. The lack of N-terminal conservation among Pur $\alpha$  and Pur $\beta$  orthologs suggests that the functions contributed by these regions are species-specific, and has been the subject of study (discussed below). Internal polyglycine stretches in Pur $\beta$  are also observed that disrupt the second basic/aromatic, and the second acidic/leucine rich repeat, in addition to a tract found near the C-terminus, all of which are absent in Pur $\alpha$  (Figure 1.2). All told, glycine content reaches 22.2% and 15.0% in murine Pur $\beta$  and Pur $\alpha$ , respectively, perhaps signifying a reason for the lack of highly ordered crystal structures for these proteins. High glycine content and polyglycine stretches have been observed in several proteins possessing nucleic acid helix-destabilization properties, including the UP1 subunit of heterogeneous ribonucleoprotein A1 (46, 86), nucleolin (99), and the *c-myc* promoter far-upstream element binding protein (FBP)(62), although the exact function attributable to polyglycine tracts is unknown. Peptide backbone flexibility imparted by these regions may aid in adaptive structural

rearrangements necessary for concerted modular domain interaction and function. Additionally, removal of the N-terminus (residues 1-84) in Pur $\alpha$  has created aggregation problems making its study difficult (133).

Multiple studies have shown that the N-terminal regions of Pur proteins are dispensable for single-stranded nucleic-acid binding (145, 310), however, appear critical for binding to long dsDNA molecules (310). Therefore, N-termini are generally regarded as regulatory motifs, necessary for seemingly distinct functions among the paralogs based on sequence divergence, although specific interaction partners remain unidentified. Despite sequence differences, similar repressive activities of Pur $\alpha$  and Pur $\beta$  have been localized to the N-terminus. Removal of amino acids 1-86 of Pur $\alpha$  results in loss of autorepression of Pur $\alpha$  promoter:reporter constructs in multiple cell types (200), whereas equivalent deletion in Pur $\beta$  results in loss of repression of a full-length SM $\alpha$ A promoter:reporter construct in rat smooth muscle cells that is otherwise repressed when cotransfected with constructs that overexpress wild-type Pur $\beta$  (145). It remains to be seen if these N-terminal-mediated mechanisms of repression are equivalent in terms of protein-protein interaction profiles.

Other structural distinctions between Pur $\alpha$  and Pur $\beta$  made obvious by sequence alignment are the divergence observed in the C-termini. While both paralogs possess high Glu (E) content, mouse Pur $\alpha$  displays a high number of Gln (Q) residues including a seven polyglutamine stretch. The corresponding motif in mouse Pur $\beta$  is replaced by a polyglycine stretch (Figure 1.2). Glutamine-rich domains have historically been ascribed transcriptional activation functions (97), however polyglutamine tracts have also been

associated with DNA helix distortion and unwinding properties in *Drosophila* GAGA factor (307). Removal of the C-terminal portion of Pur $\alpha$ , (215-322) has deleterious effects on the ability of the protein to bind linear dsDNA (310). Similar to Pur $\alpha$  N-terminal deletion, removal of amino acids 264-324 of Pur $\beta$  also results in a loss of repression of SM $\alpha$ A expression in an MCAT enhancer element-dependent manner (145) and loss of interaction capabilities with MSY1, and Pur $\alpha$  (142). Divergent means of MSY1/YB-1 association between Pur $\alpha$  and Pur $\beta$  suggests that varied stoichiometric combinations of these corepressors at the SM $\alpha$ A MCAT enhancer may serve different regulatory functions. Loss of function studies addressing this issue is the focus of Chapter III. Interestingly, despite different reported interaction interfaces utilized in YB-1/MSY1 association, Pur $\alpha$  (238) and Pur $\beta$  (145) both display modulated ssDNA-binding properties in the presence of these Y-box proteins, however a mechanism attributable to this feature has not been described.

## **NUCLEIC ACID-BINDING PROPERTIES**

Discovery of the prototypic Pur protein, Pur $\alpha$ , was the result of investigation into cellular protein components that bind and stabilize purine-rich regions of the *c-myc* associated origin of DNA replication that displays single-stranded character (10, 11). From the very beginnings of Pur protein investigation, a paradigm has emerged regarding Pur protein function, as well as those of other known sequence-specific SSBs. The challenge has been defining the mechanism by which sequence-specific SSBs recognize

and stably bind target sequences in the presence of complementary strands that possess high degrees of affinity for the same target sequence. Speculative hypotheses to this problem have predicted that inherent capabilities of sequence-SSBs or associated SSBs to destabilize double-helix conformations of DNA permit stable interaction, either by ATPase (helicase) activity or by direct thermodynamic competition with annealing. Of course, speculation has also proposed that coupling of helix destabilization by topological stress generated by negative supercoiling may aid in thermodynamic competition by reducing melting free energies of DNA duplexes. To delineate these issues, and to further identify genomic targets of sequence-specific SSBs, researchers have examined many aspects of ssDNA-binding of sequence-specific SSBs, including sequence-specificity, affinities, stoichiometries, and strand-displacement capabilities. In this regard, work that has been performed on Pur protein DNA-binding properties has suffered from inability to designate “consensus” sequences for Pur proteins, and if they differ between paralogs and/or orthologs. As stated above, high degrees of conservation within nucleic acid-binding domains of Pur $\alpha$  and Pur $\beta$  among diverse organisms suggests a preservation of nucleic acid-binding properties; however this has not been substantiated, and the lack of designation of ligating amino acids within these domains leaves this claim open to scrutiny. The following is a description of the nucleic acid-binding properties of Pur $\alpha$  and Pur $\beta$  as reported in the literature thus far.

Towards the goal of understanding shared and distinct nucleic acid-binding properties of Pur proteins, numerous and diverse ssDNA and RNA oligonucleotides have been used to gauge Pur protein involvement in promoter, replication origin, and ribonucleoprotein

interactions. Comprehensive examination of the historically utilized nucleic acid probes, as they are reported in the literature, unfortunately fails to provide speculative insight into the nature of Pur protein-nucleic acid interactions due to the vast heterogeneity of these sequences. Table 1.1 offers a wide-ranging summary, albeit non-all-inclusive, of various sequence identities, nucleotide lengths, and Pur protein interaction affinities for these molecular probes, compiled from numerous reports in the early and contemporary literature (6, 7, 10, 11, 28, 34, 48, 91, 103, 104, 106, 107, 129, 133, 136, 141-145, 149, 155, 238, 259, 272, 275, 289, 291, 322).

What can be ascertained from Table 1.1 is that Pura and Purb possess diverse targets throughout numerous genomes and transcriptomes, with most target sequences exhibiting a high degree of purine content, and a high level of affinity where it has been reported. Sequences with a relatively low purine content show reduced affinity, as demonstrated by comparison of affinities of Pur proteins for short complementary oligonucleotides. Comparing affinities of BPV1 to BPV2, for example, shows a difference in affinity close to an order of magnitude. This also, however, shows that Pur proteins exhibit low levels of affinity for pyrimidine-rich ssDNA. For this reason Pur protein ssDNA/RNA-binding is generally regarded as promiscuous based on the fact that these sequences exhibit extensive heterogeneity, and a clearly defined binding site can not be established by this simple approach. A consistency in the literature surrounding Pur protein target sequences is that these proteins bind to defined (GGN)<sub>n</sub> repeat motifs. This trend stems from early surveys of demonstrated targets, and can be seen in Table 1.1. However, issues arise from this definition of a Pur-binding site due to the fact that demonstrated

**Table 1.1. Oligonucleotides with established interactions with Pur proteins**

Oligonucleotide	Description	Sequence (5'-3')	$K_d^{APP}$ (M)
MF0677	<i>c-myc</i> origin of replication (ssDNA, 24mer)	GGAGGTGGTGGAGGAGAGAAAAG	$7.0 \times 10^{-10}$ (Pur $\alpha$ ); $8.0 \times 10^{-10}$ (Pur $\beta$ )
PUR1 (MF0677)	<i>c-myc</i> origin of replication (ssDNA, 24mer)	GGAGGTGGTGGAGGAGAGAAAAG	$3.3 \times 10^{-9b}$
MB1a	MBP promoter (ssDNA, 24mer)	TCAGAGGGCCGTGCTTTGAAGGTC	
MBP Pura consensus	MBP promoter (ssDNA, 12mer)	GGAGCGGAGGC	
HICV Pura Ori	Hypothetical Pura binding site (ssDNA, 24mer)	GGAGCGGAGGCGGAGCGGAGGC	
HICV Pura Ori 1c	HICV origin of replication (ssDNA, 18mer)	AAGGGAAAGGATGGCTG	
VG-1	HICV origin of replication (ssDNA, 27mer)	AGCTTGGAGGCGGAGGCGGCTCGGCG	
ss-G	Human telomeric repeat (ss/dsDNA, 24mer) <sup>a</sup>	TTAGGGTTAGGGTTAGGGTTAGGG	
Rgk2	<i>c-myc</i> origin of replication (ssDNA, 24mer)	TTTTTTTTGGGAGAGTTTTTTTT	
PE30-F (Pm-F, PE-F)	SM $\alpha$ A MCAT enhancer (ssDNA, 30mer)	GGAGCGAGAACAGAGGAATGCAGTGGAAAGAG	$1.7 \times 10^{-9}$ (Pur $\alpha$ ); $1.7 \times 10^{-9}$ (Pur $\beta$ )
CE-F	SM $\alpha$ A Coding sequence (ssDNA, 30mer)	GGGAGTAATGGTGGAAATGGCCAAAAAGA	
CE-RNA	SM $\alpha$ A Coding sequence (ssRNA, 30mer)	GGGAGUAUUGGUUGGAUUGGCCAAAAAGA	
PE32-F	SM $\alpha$ A MCAT enhancer (ssDNA, 32mer)	GGGAGCGAGAACAGAGGAATGCAGTGGAAAGAGA	$1.4 \times 10^{-9}$ (Pur $\alpha$ ); $2.3 \times 10^{-9}$ (Pur $\beta$ )
SPUR32-F	SM $\alpha$ A SPUR element (ssDNA, 32mer)	GAAGCGAGTGGGAGGGGATCAGAGCAAGGGGC	$1.1 \times 10^{-9}$ (Pur $\alpha$ ); $7.0 \times 10^{-10}$ (Pur $\beta$ )
THR32-R	SM $\alpha$ A THR element (ssDNA, 32mer)	GCAGTGGAAAGAGACCAGCCAGGCTTGCCACCC	$6.1 \times 10^{-9}$ (Pur $\alpha$ ); $1.6 \times 10^{-9}$ (Pur $\beta$ )
SHS	PDGF-A promoter (ssDNA, 33mer)	CTAGAGACGTGGGAGGGGCTGCAGGTGTGT	
NHE	PDGF-A promoter (ssDNA, 33mer)	CTAGAGGGGGCGGGGGCGGGGGCGGGGGAGGGG	
ssPu	Rat aldolase B replication origin (ssDNA, 22mer)	GGCAGGAGGAGGAGGCAAAAG	
TAR	HIV-1 TAR snRNA (RNA, 61mer)	GGGUCUCUGGUAGACCAGAUUCUGAGCCUGGGAGCUCUCUGGUCAACUA GGGAACCA	
23L	HICV late gene promoter (ssDNA, 23mer)	TAGGGAGGAGCTGGCTAAAAGT	
ORI-1	BPV-1 origin lower strand (ssDNA, 100mer)	TTTTTCCCGCTGAAAAAAGCGGTGATGGTGTGATTATGTTAACAACAATT ATTCACCTGGGAAAAAATACATAGCTTTACTTACCGGTTTCGGTGAGCT	$6.7 \times 10^{-10b}$
P-l	BPV-1, plasmid maintenance sequence domain 1, lower strand (ssDNA, 96mer)	AATTCGTCTCTAATATCGATGAGGTAGGAGGCTGCACACTATTTCCCA ATTTCAAGCACAGGGGATAAGTCTTGCAGATGTGACACAGG	$1.1 \times 10^{-10b}$
ORI-u	BPV-1 origin upper strand (ssDNA, 100mer)	AGCTCACCGAAACCGGTAAGTAAAGACTATGATTTTTCCAGTGAATAA TGTGTGTTAACAATAATCACACCATACCGTTTTTCAAGCGGAAAAA	$3.1 \times 10^{-10b}$
P-u	BPV-1, plasmid maintenance sequence domain 1, upper strand (ssDNA, 96mer)	AATTCCTGTGCACATCTGCAAGGACTATGCCCTCTGTGTTGAAAAATTGG GAAATAGGTGTGCAGCTCTCTACCTATCGATATTAGAGGACAG	$6.2 \times 10^{-10b}$
WW1	Polylinker-oligonucleotide (ssDNA, 78mer)	AATTACATATGGCAGCTAAGCTTGAAGACCGCCGCCATGGTCTCG CGAATTCGGAACCGTACGTCTAG	$2.5 \times 10^{-9b}$
PGK1+1	<i>S. cerevisiae</i> phosphoglycerate kinase, upper strand dimer (ssDNA, 78mer)	GATCTAATCTGCATAAATGGTCAATGCAAGAAAGATAGGATCTAACTGC ATAAATGGTCAATGCAAGAAAGATAG	$2.7 \times 10^{-9b}$
PGK2+2	Dimer of PGK2 (ssDNA, 78mer)	GATCCTATCTTCTTGCATTGACCAATTAAGCAAGTTAGATCTATCTTTC TTGCATTGACCAATTTAAGCAAGTTA	$9.1 \times 10^{-9b}$
BPV2	BPV-1 origin, truncated lower strand (ssDNA, 43mer)	GATCTTATGGTGTGATTATTGTTAACAACAATTATCACTGG	$1 \times 10^{-8b}$
PGK2	<i>S. cerevisiae</i> phosphoglycerate kinase, upper strand (ssDNA, 39mer)	GATCCTATCTTCTTGCATTGACCAATTAAGCAAGTTA	$5.6 \times 10^{-8b}$
BPV1	BPV-1 origin, truncated upper strand (ssDNA, 43mer)	GATCCCAAGTGAATAATTGTTGTTAACAATAATCACACCATCAA	$7.1 \times 10^{-8b}$
Soma-ss-CRE	Somatostatin promoter CRE (ssDNA, 36mer)	CTGGGGCGCCCTCTGGCTGACGTCAGAGAGAGAG	
Th-ss-CRE	Tyrosine hydroxylase promoter CRE (ssDNA, 36mer)	GGGGGACCCAGAGGGGCTTTGACGTCAGCCTGGCCT	
$\alpha$ -MHC, CarG1	$\alpha$ -MHC promoter (ssDNA, 33mer)	GGCTAAGCAGACCTTTCATGGGAAACCTCAGG	
$\alpha$ -MHC, CarG2	$\alpha$ -MHC promoter (ssDNA, 35mer)	GTCCAGCAGATGACTCCAATTAGCAGCAGGCA	
$\alpha$ -MHC, PNR	$\alpha$ -MHC promoter (ssDNA, 35mer)	ACCTAGAGGGAAAGTGTCTCCCTGGAAGTGGGCT	
$\alpha$ -Skeletal actin, CarG	$\alpha$ -Skeletal actin promoter (ssDNA, 75mer)	CGCCGACACCCAAATATGGCGACGA	
$\alpha$ -Cardiac actin, CarG	$\alpha$ -Cardiac actin promoter (ssDNA, 26mer)	GGGACCAATAAGGCAAGGTGGCAG	
ANF, CarG1	Atrial natriuretic factor promoter (ssDNA, 40mer)	TCGCTGGACTGATAACTTTAAAAGGGCATCTTCTCTGGC	
ANF, CarG2	Atrial natriuretic factor promoter (ssDNA, 37mer)	TGCTCTCTGCCCTAATTTGGAGCCCTGACAGCTG	
$\beta$ PNRE-S	B-Myosin heavy chain promoter, NRE (ssDNA, 22mer)	GTGGTCTGGTGGTGGTGTCA	

<sup>a</sup> This (TTAGGG)<sub>4</sub> repeat sequence was shown to bind Pur $\alpha$  in a ssDNA configuration, and in a dsDNA configuration with a complementary (CCCTAA)<sub>3</sub>, 18mer, thus possessing 5' and 3' overhangs (Wortmann, et al (2005) *Biochem Biophys Acta* 1743:64-78).

<sup>b</sup> Affinity determinations were made using preparations of copurified Pur $\alpha$  and Pur $\beta$ , with no indications of the relative amount of each protein in these preparations (Jurk, et al (1996) *Nucleic Acids Res* 24:2799-2806).

targets display extensive heterogeneity regarding the number, position, spacing and exact sequence of these GGN repeats. For example, It has been stipulated that high affinity binding requires multiple GGN repeats, and that N not be G (92, 148), however binding of Pur $\alpha$  to probe RgR2 (310) which harbors a single purine-rich motif, (GGGAGAG) and only one GGN repeat, refutes this claim, although no indication of the relative affinity of Pur $\alpha$  for RgR2 compared to other, more elaborate purine-rich probes was given. The heterogeneity of GGN repeat spacing among established probes also confounds attempts to unambiguously define a Pur protein binding site. An interesting trend pointed out by Jurk, et al (136) is that apparent binding affinities increase with increasing probe length. This effect could be due to increases in binding site concentrations that accompany the use of probes with multiple binding sites, or that Pur proteins bind to multi-site lattices via cooperative mechanisms. This possibility has not been thoroughly examined through dissection of intrinsic interaction energetics between Pur proteins and individual binding sites or intersite communications. Furthermore, the stoichiometric species of Pur $\alpha$ , Pur $\beta$ , or heteromeric complexes thereof that are competent to bind nucleic acids has not been clearly established. Many groups have shown that particular Pur protein nucleoprotein complexes are heterogeneous with respect to Pur protein content (136, 141), whereas others have shown that homomultimers of Pur $\alpha$  only form in the presence of RNA (91), although the exact stoichiometry of self-association in these cases is unknown.

Most of what we know today regarding Pur protein nucleic acid binding properties is the result of early work performed by Bergemann, et al (10, 11). This work represents the initial discovery of Pur $\alpha$  which was preceded by identification of a purine-rich motif

flanking the *c-myc* promoter shown to adopt bent conformations and serves as an origin of DNA replication in HeLa cells. Competition-based analyses identified a candidate 24mer ssDNA oligonucleotide sequence for Pur $\alpha$  binding (termed MF0677): 5'-GGAGGTGGTGGAGGGAGAGAAAAG-3'. The length of this sequence in this study was largely arbitrary, as it represents a portion of the entire bend-region of the replication origin under investigation, and thus it was unknown at the time whether this sequence represented a partial binding site, or an array of multiple binding sites. Competition experiments resulted in the disappearance of a single electrophoretic species suggesting that this sequence represented an entirety of binding sites in the context of the *c-myc*-associated origin of replication. These studies also demonstrated a clear preference of Pur $\alpha$  for ssDNA versus dsDNA sequences. Furthermore, methylation interference footprinting of the Pur $\alpha$ :MF0677 complex detected several guanine-specific contact points, suggesting that Pur $\alpha$  has a guanine preference over adenines, and substitution of several guanines with adenines results in loss of affinity. The ability of ssDNA oligonucleotides derived from other replication origin purine-rich elements to compete for binding to the *c-myc*-associated purine-rich motif suggested that this motif may be common to replication origins, thus implicating Pur $\alpha$  in cell cycle regulation. This discovery also prompted an evaluation of sequence similarities among putative Pur $\alpha$  binding sites throughout known origins of replication, using the candidate MF0677 24mer sequence as a reference (10). In doing so, a core 16mer Pur $\alpha$  consensus sequence was developed from 11 different *PUR*-elements identified in six different organisms: 5'-GGNNGAGGGAGARRRR-3' (N = any nucleotide, R = A/G).

A quasi-systematic validation of this consensus sequence by direct demonstration of Pur $\alpha$  binding capacity to parent *PUR*-elements was performed by measuring the ability of Pur $\alpha$ , in the context of a library-candidate  $\lambda$  phage-infected *E.coli* cell extract, to alter the electrophoretic mobility of oligonucleotides harboring mutations of parent MF0677 (10, 11). A compiled view of these analyses is displayed in Table 1.2. The results of this analysis suggest that substitution of consensus-defined guanines within the core-consensus element disrupt binding (MM0677), consistent with methylation interference data suggesting these residues represent essential contact points within the nucleoprotein complex. Transversional mutation (G to T) of these nucleotides also shows the importance of these conserved guanines in high affinity nucleoprotein assembly (ME0677 and MG0677), but also highlights the dispensability of nucleotides at the 3' end of the core consensus sequence (MC0677). Disruption of GNGG or GGNGG motifs in MF0677 within the core consensus sequence (MA0677 and MB0677) also diminishes binding of Pur $\alpha$ , validating the notion of the core consensus sequence. This is further supported by results of competitions with oligonucleotides in which the G at position 1 of the core consensus sequence is restored (compare MI0677 to MH0677). However, guanine nucleotides flanking this putative core sequence were identified to be important for nucleoprotein stability (MJ0677), suggesting that the core consensus sequence notion is not resolute. Interestingly, an oligonucleotide representation of the hamster dihydrofolate reductase (*dhfr*)-associated replication origin *PUR*-element (DR3529), which completely obeys consensus rules, binds poorly to Pur $\alpha$ , again suggesting that the core consensus sequence generated by sequence data mining and mutagenesis may be

**Table 1.2. Pur $\alpha$  binding properties of mutant *c-myc*-associated *PUR*-element oligonucleotides**

Oligonucleotide <sup>b</sup>	Sequence (5' to 3') <sup>d</sup>	Pur $\alpha$ binding affinity <sup>e</sup>
MF0677	GGAGGTGGTGGAGGGAGAGAAAAG	+++
MM0677	GGAGGTGGTGGAAAAAGAGAAAAG	+
MA0677	GGAGATAGTGGAGGGAAAGAAAAG	+
MB0677	GGAGATAGTAGAGGGAGAGAAAAG	+
MC0677	GGAGGTGGTGGAGGGTTTTTTTTTT	+++
ME0677	GGAGGTGGTGGAGGTTTTTTTTTTT	++
MG0677	GGAGGTGGTGGAGTTTTTTTTTTTT	+
MH0677	TTTTTTTGTGGAGGGTTTTTTTTTT	+
MI0677	TTTTTTGGTGGAGGGTTTTTTTTTT	++
MJ0677	TTTGGTGGTGGAGGGTTTTTTTTTT	+++
DR3529 <sup>c</sup>	TGATGAGGGAGAGGGAGAAGGGAT	+
dG <sub>24</sub>	GGGGGGGGGGGGGGGGGGGGGGGG	-
dA <sub>24</sub>	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	-
Consensus	NNNNNNGGNNGAGGGAGARRRRNN	?

<sup>a</sup> Tabulated data are results from experiments described in Bergemann, et al (1992) *Mol Cell Biol* 12:1257-1265 and Bergemann, et al (1992) *Mol Cell Biol* 12:5673-5682.

<sup>b</sup> Oligonucleotide MF0677 is derived from human *c-myc* upstream *PUR*-element. Other oligonucleotides denoted M\_0677 indicate mutants derived from MF0677.

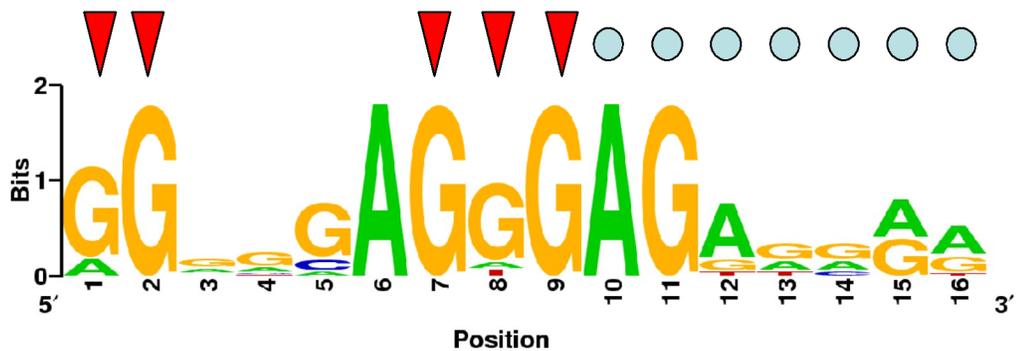
<sup>c</sup> Oligonucleotide DR3529 is derived from hamster dihydrofolate reductase (*dhfr*)-associated *PUR*-element

<sup>d</sup> Nucleotides in red correspond to “core” consensus sequence. Underlined nucleotides are positions of substitution with respect to MF0677

<sup>e</sup> Interaction affinities are based on electrophoretic mobility shift competition experiments as described in references noted in <sup>a</sup>.

inaccurate. The inability of Pur $\alpha$  to bind dG<sub>24</sub> shows that guanine content alone cannot establish a high affinity interaction. A sequence Logo of the core Pur $\alpha$  binding site, based on compilation of putative PUR-elements flanking origins of DNA replication as determined by Bergemann, et al (10), is shown in Figure 1.3 (performed using web-based software: <http://weblogo.berkeley.edu/> (246)). This shows that regions of high conservation are both essential and dispensable for binding to Pur $\alpha$ , as indicated by arrows and circles, respectively. The Logo does not however indicate the positions of critical nucleotides outside this core region, as detected with MJ0677. It should also be noted that conserved residues in oligonucleotide MJ0677, as compared to MF0677, may represent a more accurate core sequence based on similar binding affinities between these sequences.

The appropriateness of consensus sequence usage has been questioned by several researchers that point out potential genomic binding sites for proteins with defined consensus sequences are often missed due to the relatively low amount of information portrayed in consensus sequence representations and the inherent sequence promiscuity of DNA-binding proteins (245). Usage of sequence Logos has sidestepped this limitation to some degree; however their construction relies on accurate footprints of protein binding sites and that they be constructed from numerous sequences to accurately depict nucleotide preference patterns. This problem is likely enhanced when defining consensus sequences for sequence-specific SSBs, especially those which have no defined footprint. Specifically, errors in consensus sequence and sequence logo accuracy may arise from the fact that ssDNA is generally more flexible than its dsDNA counterparts. Binding site



**Figure 1.3. Pur $\alpha$  nucleotide preferences within putative origin-associated PUR-elements.** Sequence comparisons of origin of replication-associated purine-rich elements identified twelve putative *PUR*-elements (Bergemann, et al. (1992) *Mol Cell Biol* 12:1257-1265). Shown is a sequence Logo (Schneider, et al (1990) *Nucleic Acids Res* 18:6097-6100) depicting relative nucleotide usage within these aligned *PUR*-elements. *Triangles* depict positions of nucleotides contributing to stable binding of Pur $\alpha$ , whereas *circles* denote nucleotides that are dispensable to nucleoprotein assembly based on *in vitro* binding studies described in the text and Table 1.2 (Bergemann, et al. (1992) *Mol Cell Biol* 12:1257-1265, Bergemann, et al (1992) *Mol Cell Biol* 12:5673-5682). Compilation of these findings shows that highly conserved positions in putative origin-associated *PUR*-elements are both essential and dispensable for Pur $\alpha$  association.

recognition in dsDNA usually proceeds via scanning of the major or minor groove of dsDNA until a chemical group signature is encountered. The orientation of hydrogen-bond donors and acceptors, as well as methyl groups is recognized by complementary surfaces on the protein, which permits reversible interaction. Sequence promiscuity arises from slight variations in the chemical landscapes of the major or minor groove that are tolerated by the surface of the protein, and this degree of tolerability, in the eyes of this author, likely dictates binding affinity. In ssDNA, however, the major and minor groove faces of nucleotides are supplemented with the available Watson-Crick base-pairing face to provide additional ligating chemical groups. Seemingly this characteristic would make ssDNA less permissive to binding promiscuity; however near-limitless rotational freedom of the glycosidic bond (syn- versus anti- rotamers) in single-stranded configurations permits a vast array of chemical-group orientations, presenting chemical groups of all three faces to the binding surface of the protein. Hypothetically, reverse directional binding (5' to 3' versus 3' to 5') should be possible as well if ligating nucleotides are rotated 180° around their glycosidic bonds, and assuming that specific sugar contacts are not crucial for interaction. Additional backbone flexibility may compound this promiscuity further by permitting bulges, loops, and other three-dimensional structures.

To circumvent issues of binding promiscuity, several researchers have used oligonucleotide MF0677 as a “standard” for Pur $\alpha$  ssDNA-binding (10, 11, 133, 136, 310), despite the lack of extensive and rigorous evaluation of the Pur $\alpha$ :MF0677 nucleoprotein complex stoichiometry or interaction energetics. Recent studies by

Wortmann, et al (310) have investigated this issue using a Scatchard analysis approach, and found that truncated recombinant FLAG epitope (DYKDDDK)-fusion forms of both Pur $\alpha$  and Pur $\beta$  (amino acids 1-314 and 1-316 of Pur $\alpha$  and Pur $\beta$ , respectively) bind MF0677 with 1:1 (Pur:MF0677) nucleoprotein complex stoichiometries exhibiting similar subnanomolar affinities (listed in Table 1.1). However, the graphical linearization procedure used by these researchers may have caused a misestimation of complex stoichiometry, as multiple binding sites (low and high affinity) were observed, as was considerable curvature of the linear Scatchard plots. Also, the use of truncated recombinant proteins may alter binding properties, although this contention has not been experimentally tested. Nonetheless, the results of this study suggest that the ssDNA-binding-competent species of Pur proteins is a monomer.

Similar mutation analyses have been performed in regards to Pur protein binding to the proximal MCAT enhancer element of SM $\alpha$ A, as described previously. Figure 1.4 depicts a summary of nucleotides identified to be critical for Pur protein interactions within the SM $\alpha$ A MCAT enhancer region based on results from various studies investigating Pur-responsiveness (repression) of promoter constructs used in transfection assays (48, 275) as well as direct binding assays (28, 148). As stated previously, transient transfection analysis of nucleotide sequences flanking the core MCAT sequence suggested that repressor binding occurs at nucleotides 5' to the TEF-1 binding site, as transversional mutation of nucleotides in this region results in loss of repression of SM $\alpha$ A promoter:reporter constructs TV191, TV189, and TV187. The prevalence of guanines in this area also supports this hypothesis. Similarly, constructs containing



binding sites with slight degeneracy of a core GGGAGA element. Data represent a compilation of findings published previously (Cogan, et al (1995) *J Biol Chem* 270:11310-11321; Sun, et al (1995) *Mol Cell Biol* 15:2429-2436; Carlini, et al (2002) *J Biol Chem* 277:8682-8692; Knapp, et al, (2006) *J Biol Chem* 281:7907-7918).

transversions of nucleotides 3' to the core MCAT motif (TV173, TV167) also display loss of repression, suggesting that these nucleotides are also responsible for binding Pur proteins (48). Furthermore, addition of nucleotides spanning -195 to -192 augments repression ( $\Delta$ 195), also suggesting that this construct contributes an additional Pur protein binding site, or that it represents a more complete binding site with higher repressor affinity than the truncated construct VSMP4 that possesses nucleotides -191 to +46 of the SM $\alpha$ A promoter (28). Direct ssDNA-binding experiments showed a similar trend, that being the importance of nucleotides flanking the core MCAT sequence (28, 275). Oligonucleotides that accurately represent the region encompassing nucleotides -196 to -164 (PE32-F) exhibit the highest affinity for Pur proteins (148), where as deletion or mutation of these nucleotides results in loss of affinity. Interesting is the observation that transversion of nucleotides -175 to -172 (TV175) results in enhanced binding affinity (275). Reasons for this are unclear but may be due to addition of a guanine at position -172 which is proximal to other recognizably important nucleotides. Collectively, the data represented by Figure 1.4 show the importance of two regions of high guanine content (red arrows), consistent with the possibility that this region contains two Pur protein binding sites, as suggested previously by experiments detecting multimeric Pur:PE complexes (141). These putative Pur protein binding sites also show resemblance to core *PUR*-element sequences as identified by researchers investigating the *c-myc*-associated *PUR*-element (10, 11). These studies proposed a core consensus sequence of 5'-GGGAGA-3', for which oligonucleotide MF0677 has only one, and has been shown to bind only one mole of Pur $\alpha$  or Pur $\beta$  (310). The two sites in PE32-F show slight

degeneracy; the site on the 5'-flank of the core MCAT element is GGGAGC, whereas the 3' site is GGAAGA (degenerate nucleotides are shown in *italics*). Experimental evidence testing this hypothesis has not been reported, but its acquisition is a goal described in Chapter VI of this dissertation. Interestingly, mutations in either putative site cause drastic loss of repression or observed ssDNA-binding consistent with the notion that cooperative interactions between sites might dictate affinity.

A remaining aspect of Pur protein nucleic acid-binding that needs consideration is that involving dsDNA. Numerous reports have surfaced reporting binding of both Pur $\alpha$  and Pur $\beta$  to dsDNA probes *in vitro*, including the *c-myc*-associated replication of origin (10, 11, 310), the rat aldolase B-associated origin of replication (259), the mouse  $\alpha$ -skeletal actin promoter CArG element (103), and mouse SM $\alpha$ A promoter elements PE (148) and SPUR (148, 272). In each case however, a clear preference has been established of Pur proteins for single-stranded configurations of *PUR*-elements. So questions remain as to whether observed binding of Pur proteins to dsDNA represents a distinct binding mode, that it is equivalent to ssDNA-binding in the sense that major contact points between Pur proteins and dsDNA exist on the purine-rich strand while the complementary pyrimidine rich strand is completely displaced (dissociates) or partially displaced (incomplete helix destabilization, or "bubble" formation), or is mediated by an unknown dsDNA-binding protein via protein-protein interactions. Work with purified proteins certainly refutes this latter claim, as investigations have shown the ability of recombinant Pur $\alpha$  to completely displace short pyrimidine-rich strands from duplexes containing protruding purine-rich strands in the absence of ATP (56, 310). The ability of

Pur proteins to displace strands from blunt-ended dsDNA probes has not been established. Both Pur $\alpha$  and Pur $\beta$  have exhibited the ability to bind dsDNA in the context of plasmid DNA, with both showing higher affinities for supercoiled configurations (310). This finding supports the notion that dsDNA binding is not a distinct functional mode of Pur proteins. This claim is based on the fact that both supercoiled and relaxed topoisomers of plasmid DNA are double-stranded *per se*, but torsional stress existing in supercoiled plasmids imparts destabilization of the forces maintaining helical structure, namely hydrogen-bonds as indicated by differences in melting temperatures (249, 293) and the increased prevalence of regions assuming non-B-DNA conformations (204, 220, 227, 260, 294). Disruption of annealing free energy associated with supercoiling may permit localized strand displacement and binding of sequence-specific SSBs in a manner that is observed as dsDNA-binding by conventional methods, like electrophoretic mobility shift assays (310). Interestingly, observed binding of Pur $\alpha$  to supercoiled plasmids pUC19, and a pUC19 construct harboring the *c-myc*-associated replication origin produces discrete, quantized bands of DNA as judged by agarose gel electrophoretic mobility shift assays (310) in a manner similar to banding patterns of topoisomers. This characteristic could suggest that binding of Pur $\alpha$  to these supercoiled plasmids increases the twisting number or writhe in neighboring regions. However, similar electrophoretic patterns observed with linearized plasmid suggests that these bands represent successive stoichiometric complexes of Pur $\alpha$ :plasmid DNA (310).

A veritable aspect of dsDNA binding by Pur proteins illustrated by studies reported by Wortman, et al (310) is that the length of the DNA molecule (number of base-pairs)

plays a significant role. For example, binding of Pur $\alpha$  to a 343 bp PCR-product containing the same *PUR*-element as MF0677 is negatively influenced by competition with MF0677 but not double-stranded MF0677 (24 bp). This shows that longer molecules of dsDNA bind to Pur $\alpha$  with greater affinity than shorter ones. Reasons for this are unclear. Nearest-neighbor algorithms for estimating two state annealing free energies of short oligonucleotides have proven to be extraordinarily accurate, comparable to experimental determinations. However, their usage with long DNA molecules breaks down considerably as they predict annealing free energies approaching infinity and extremely high melting temperatures (239). Thermal denaturation of long, linear DNA molecules has shown that melting occurs at temperatures as low as 64°C, lower than predicted by nearest-neighbor methods, and lower than that of purine-rich oligonucleotides approximately 30 nt in length. Furthermore, melting is likely not a two-step process as indicated by the presence of locally melted subdomains (234, 262). Evidently, long dsDNA molecules have an added destabilizing component which is currently unidentified; although this component may be diffusion-related as relaxed, long dsDNA molecules display lower rotational and translational diffusion coefficients than supercoiled molecules of similar molecular weight (230). Whether or not entropic limitations on rotational and translational diffusion associated with long dsDNA molecules potentiate enhancement of intramolecular vibration and localized melting/unstacking is unknown, and its involvement in a possible mechanism for helix destabilization and SSB-binding is purely speculative.

In conclusion, mechanisms describing Pur protein single stranded nucleic-acid recognition and binding *in vivo* remain poorly defined. Unsuccessful attempt to accurately determine Pur protein target sequences and to quantitatively detail binding events have undermined this effort. The detectable multi-faceted involvement of Pur proteins in gene regulation coupled with their clear preference for binding single-stranded targets *in vitro* indicates that single-stranded nucleic acid binding is also accomplished *in vivo*. To understand how this process occurs, details of Pur protein structure, function, and chemistry must be methodically and experimentally uncovered.

## CHAPTER II. SCOPE OF THE PRESENT WORK

As details of Pur protein function have emerged from various laboratories investigating aspects of transcriptional, translational, and DNA replication regulation associated with these proteins, so have numerous questions. A major goal of the laboratory of Dr. Robert J. Kelm, Jr. is the delineation of functional distinctions between mouse Pur-paralogs Pur $\alpha$  and Pur $\beta$ . As described previously, subtle yet orthologously conserved differences in Pur $\alpha$  and Pur $\beta$  primary sequences suggest that functional differences exist between these proteins *in vivo*. Towards the testing of the hypothesis that states regions of sequence heterogeneity confer distinct functions between Pur $\alpha$  and Pur $\beta$ , I and members of the Kelm laboratory undertook a parallel gain and loss-of-function approach to delineate functions of Pur $\alpha$  and Pur $\beta$  in regards to regulation of SM $\alpha$ A expression in appropriate cell culture models. Chapter III is an excerpt from two papers published by our laboratory (148, 149) which collectively describes the contributions I made to this body of work, and focuses on the loss-of-function angle. This work details the differing contributions Pur $\alpha$  and Pur $\beta$  make to SM $\alpha$ A repression in both cell-type and ortholog-dependent fashions, and establishes the dominant repressor activity of Pur $\beta$  in this context.

Identification of Pur $\beta$  as a dominant repressor of SM $\alpha$ A expression in mouse embryo fibroblasts and vascular smooth muscle cells fueled a rigorous thermodynamic analysis of the ssDNA-binding properties of this protein which I performed. Extensive studies

focusing on the solution behavior and quaternary structure of recombinant Pur $\beta$  are described in detail in Chapter IV, with the results of this study showing the capability of recombinant Pur $\beta$  to self-dimerize in the absence of nucleic acid and that monomeric and dimeric species adopt elongated, asymmetric geometries. These findings were also published (222). Establishment of Pur $\beta$  dimer-formation at protein concentrations that are thought to be attained in cells that exhibit SM $\alpha$ A repression led to the hypothesis which states Pur $\beta$  self-association represents a prerequisite step for binding to the MCAT-associated Pur protein binding sites. Testing of this hypothesis involved a rigorous thermodynamic assessment of the mechanism used by recombinant Pur $\beta$  in binding to ssDNA sequences representative of the proximal MCAT enhancer element of the SM $\alpha$ A promoter, and is described in Chapter V. These studies corroborate findings outlined in the background that suggest that two Pur $\beta$  binding sites exist in this element and that successive binding of two Pur $\beta$  monomers proceeds with high affinity stabilized by cooperative interactions, thus refuting the original hypothesis of obligate dimer formation. To substantiate the nucleotide sequence identities of putative Pur $\beta$  binding sites within the SM $\alpha$ A MCAT enhancer region, I performed a systematic analysis of the stabilities of nucleoprotein complexes composed of Pur $\beta$  and truncated mutant PE oligonucleotides. The employed methodologies and results of this approach are described in Chapter VI.

Collectively this body of work describes the ssDNA-recognition and binding mechanisms used by Pur $\beta$  in the repression of SM $\alpha$ A expression, as it pertains to

molecular mechanism governing phenotypic plasticity of cell types implicated in physiological responses to injury and pathological progressions of numerous disease states.

## **CHAPTER III. PUR PROTEIN LOSS-OF-FUNCTION STUDIES**

### **IDENTIFY PUR $\beta$ AS A DOMINANT REPRESSOR OF SM $\alpha$ A**

#### **EXPRESSION**

*The work described herein has been published as part of greater bodies of work authored by the Kelm laboratory and can be found in its original form in:*

Knapp AM, Ramsey JE, Wang SX, Godburn KE, Strauch AR, Kelm RJ, Jr. (2006) Nucleoprotein interactions governing cell type-dependent repression of the mouse smooth muscle alpha-actin promoter by single-stranded DNA-binding proteins Pur alpha and Pur beta. *J Biol Chem.*281(12):7907-18.

*and*

Knapp AM, Ramsey JE, Wang SX, Strauch AR, Kelm RJ, Jr. (2007) Structure-function analysis of mouse Pur beta II. Conformation altering mutations disrupt single-stranded DNA and protein interactions crucial to smooth muscle alpha-actin gene repression. *J Biol Chem.*282(49):35899-909.

#### **INTRODUCTION**

Phenotypic plasticity of VSMCs and adventitial fibroblasts is a feature critical to adaptive pathophysiological tissue functions supported by these cell types *in vivo*, including vessel wall remodeling associated with atherogenesis, and restenosis (120, 174,

175, 232, 242). Dynamic phenotypic modulation is the result of widespread genetic reprogramming of genes essential to cell function, and requires extensive cooperative networking of factors that regulate gene expression. Indices of differentiation status are difficult to establish due to the continuous spectrum of biomarker expression profiles that are exhibited by cell types undergoing transdifferentiation, however profiles of marker expression remain the best way to categorically measure phenotypes (209, 210). Absence or reduction of staining for contractile proteins is a characteristic feature of cells comprising the neointima of vulnerable atheroma (210). Among the battery of contractility markers commonly used for assessing phenotypic status of VSMCs and adventitial fibroblasts, SM $\alpha$ A is the most widely utilized marker due to its dynamic range of expression in the aforementioned cell types (210). Furthermore, SM $\alpha$ A is the principle contractile protein upregulated in myofibroblasts responding to signals promoting normal and pathological wound healing (226, 231, 271).

SM $\alpha$ A is the most abundant protein expressed in VSMCs, accounting for approximately 10-40% of the total protein (79, 257). Genetic inactivation of SM $\alpha$ A in mice germ cells indicate that SM $\alpha$ A plays an essential role in regulating vascular contractility and blood pressure homeostasis (98). Down-regulation of SM $\alpha$ A has been implicated in microfilament rearrangement and changes in cell shape consistent with a transformed fibroblast phenotype, suggesting a role in non-myogenic cell types (166). Loss-of-function studies conducted in cultured cells also suggest that another essential function of SM $\alpha$ A is to inhibit cell motility (231). It follows that repression of SM $\alpha$ A expression in activated VSMCs may be a key molecular event leading to enhanced cell

migration and proliferation, as well as contributing to the vulnerability of atheroma *in vivo* (88, 151). Therefore, elucidation of repressive mechanisms governing SM $\alpha$ A expression is paramount to understanding and managing conditions associated with pathophysiological vessel remodeling.

Spatial and temporal regulation of SM $\alpha$ A expression is mediated by an array of *trans*- and *cis*-acting components that rely on dynamic functional interplay to provide highly plastic expression (158, 315). Potent activation and repression of full length SM $\alpha$ A promoter constructs has been confined to a cryptic MCAT enhancer element located -195 to -164 nucleotides relative to the start of transcription (48, 275) that has been shown to possess extensive purine/pyrimidine asymmetry and to transiently adopt single-stranded conformations *in vivo* (9). Bifacial enhancer activity has been shown to be regulated by structure-specific occupation by dsDNA-dependent transcriptional activator TEF-1, or sequence-specific SSB repressors Pur $\alpha$ , Pur $\beta$ , and MSY1 (28, 48, 141, 143-145, 275). Repression has been shown to require binding of Pur $\alpha$  and/or Pur $\beta$  to the purine-rich strand of the MCAT enhancer and occupation of the pyrimidine-rich strand by MSY1. It has not been established whether structural interconversions are cause or consequence of differential transcription factor occupancy.

Pur $\alpha$  and Pur $\beta$  are members of the Pur family of proteins and were so named based on their preference for binding purine-rich ssDNA and RNA sequences (92). Pur $\alpha$  and Pur $\beta$  are ~70% identical at the amino acid level (143), with each protein possessing a minimal ssDNA/RNA-binding domain composed of highly homologous sequence repeats unique among other known nucleic acid-binding proteins (132). However, significant

sequence differences between Pur $\alpha$  and Pur $\beta$  exist near the N- and C-termini suggesting each protein may have divergently evolved to perform distinct functions. A report has suggested that Pur $\alpha$  and Pur $\beta$  bind to ssDNA as either homo- or heterodimers and can each interact with MSY1 (28). Despite these biochemical similarities, gain-of-function (over-expression) studies clearly illustrate that Pur $\alpha$  and Pur $\beta$  are not redundant in terms of their transcriptional repressor activity toward the full-length mouse SM $\alpha$ A promoter in transfected VSMCs (145).

The potential capacity of Pur $\alpha$ , Pur $\beta$ , and MSY1 to collaboratively regulate expression of SM $\alpha$ A has been extended to the post-transcriptional level as well. The sequence of the MCAT enhancer in the 5'-flanking region bears striking resemblance to a region of exon 3 in the SM $\alpha$ A open reading frame. It has been reported that Pur $\alpha$ , Pur $\beta$ , and MSY1 can form a stable messenger ribonucleoprotein complex (mRNP) with the exon 3 RNA sequence in a manner which blocks the translational efficiency of a reporter mRNA (142). These results raise the possibility that these proteins may also participate in post-transcriptional control mechanisms for SM $\alpha$ A protein levels by either directly regulating SM $\alpha$ A translation or by reducing nuclear repressor concentrations that normally restrict transcription. Pur $\alpha$ - and Pur $\beta$ -dependent repression of  $\alpha$ -myosin heavy chain gene expression in cardiomyocytes has been shown to occur at both the transcriptional and translational levels, thus supporting this notion (104).

It has been reported that profiles of interactions of Pur $\alpha$  and Pur $\beta$  with certain transcriptional activators are altered during transforming growth factor  $\beta$ 1 (TGF $\beta$ 1)-

induced myofibroblast differentiation *in vitro* (272). These activators include Sp1, Sp3, and Smad proteins (Smad2, Smad3). DNA-binding studies have revealed distinct *cis*-element binding sites for these factors that are downstream of the MCAT enhancer suggesting that Pur $\alpha$  and Pur $\beta$  may facilitate cell type-specific repression by mechanisms involving protein-protein interactions in addition to competitive DNA-binding. The importance of functional interplay between activators and Pur-repressors was also revealed in over-expression experiments in which serum response factor (SRF) was found to reverse Pur $\beta$ -mediated repression of the SM $\alpha$ A promoter (145). Given that SRF and Smad proteins are essential downstream targets of signaling pathways that promote smooth muscle differentiation (28, 47, 145, 157, 210), it is likely that the gene regulatory effects of Pur repressors are not limited to SM $\alpha$ A but likely include other markers of VSMC differentiation. Furthermore, in light of the fact that Pur $\alpha$  has also been implicated in regulation of cell cycling due to its interaction with retinoblastoma tumor suppressor protein (pRB) (133), E2F-1 (58), and cyclin A (125), the possibility exists that Pur proteins may be directly involved in mechanisms modulating cell cycle progression *in vivo* (146).

In the present study, we have employed a RNA-interference (RNAi)-mediated loss-of-function approach to study functional similarities and differences between Pur $\alpha$  and Pur $\beta$  with respect to regulation of SM $\alpha$ A expression. While stable and transient epigenetic knockdown of both Pur $\alpha$  and Pur $\beta$  results in derepression of endogenous SM $\alpha$ A expression in AKR-2B MEFs, Pur $\beta$  appears to be the dominant Pur protein repressor in this cell type. This finding is corroborated by chromatin

immunoprecipitation experiments showing elevated levels of Pur $\beta$  promoter occupancy. This effect is restricted to serum-deprived AKR-2B MEFs, as transient transfection in C57BL/6J mouse aortic VSMC outgrowth cultures showed Pur $\alpha$  to be similarly repressive. Knockdown of Pur $\alpha$  and Pur $\beta$  results in synergistic derepression of SM $\alpha$ A promoter-driven reporter expression in both cell types in a MCAT enhancer element-dependent manner. These studies, as previously published (148) point to cell type-restricted collaborative and distinct functions for Pur $\alpha$  and Pur $\beta$  in regards to repression of SM $\alpha$ A expression, and complement parallel gain-of-function studies that show similar results.

## **MATERIALS AND METHODS**

*Cell culture, transient transfection, and reporter gene assay* – Aortic segments from C57BL/6J mice were obtained following protocols approved by the University of Vermont Institutional Animal Care and Use Committee. VSMCs were isolated by cell outgrowth from aortic tissue explants and characterized as previously described (40, 244). Primary VSMCs were cultivated in a 90% air/10% CO<sub>2</sub> incubator at 37°C in growth medium consisting of Dulbecco's modified Eagle's medium, 1 $\times$  insulin-transferrin-selenium supplement (Invitrogen), and 20% v/v heat-inactivated fetal bovine serum (FBS) (Hyclone). For transient transfection studies, primary VSMCs were seeded in six-well plates at a density of 10,000 cells/well and transfected with the use of jetPEI™ reagent at a ratio of 2  $\mu$ l/ $\mu$ g plasmid DNA as directed by the manufacturer (Qbiogene).

AKR-2B mouse embryonic fibroblasts (MEFs) or rat A7r5 VSMCs were cultured and transiently transfected as previously described (297). Briefly, subconfluent AKR-2B or A7r5 cells seeded in 60 mm dishes were transfected with the use of GenePORTER™ reagent (Gene Therapy Systems) at a ratio of 3  $\mu$ l/ $\mu$ g plasmid DNA. After 48 h incubation in growth medium, cells were washed with phosphate-buffered saline and extracted using 1 $\times$  reporter lysis buffer (Roche Applied Science) supplemented with protease inhibitors. Total protein in transfected cell lysates was determined by bicinchoninic acid assay (Sigma) using BSA as a protein standard. Commercial immunoassays were used to measure chloramphenicol acetyltransferase (CAT) or  $\beta$ -galactosidase ( $\beta$ -gal) reporter proteins as directed by the manufacturer (Roche Applied Science). Reporter values were corrected for total protein content. Transfections were typically performed in triplicate and repeated two to three times to ensure reproducibility. Data sets were subjected to one-way analysis of variance to identify differences among group means at the  $p < 0.05$  significance level. Post-hoc comparisons were performed using the Bonferroni adjustment for multiple comparisons.

***Construction of shRNA expression plasmids targeting mouse Pur $\alpha$  or Pur $\beta$***  – The design of mouse Pur $\alpha$ - and Pur $\beta$ -specific shRNA sequences was facilitated by free software available through Invitrogen Corporation (<https://rnaidesigner.invitrogen.com/rnaiexpress/>). Nucleotide sequences encoding full-length mouse Pur $\alpha$  and Pur $\beta$  were previously reported (143). Parameters set to the mouse genome database for BLAST, sense-loop-antisense configuration for shRNAs, and a loop sequence of CGAA yielded two shRNA-encoding sequences predicted to be Pur $\beta$ -

specific and four predicted to be Pur $\alpha$ -specific: Pur $\beta$ -I) 5'-GTCGGTATGCAGATGAAATG**ACGAAT**CATTTTCATCTGCATACCGAC-3' (888), Pur $\beta$ -II) 5'-GATGAAATGAAAGAGATCCAG**CGAACT**GGATCTCTTTCATTTTCATC-3' (899), Pur $\alpha$ -I) 5'-GCAAGTACTACATGGATCTC**ACGAAT**GAGATCCATGTAGTACTTGC-3' (973), Pur $\alpha$ -II) 5'-GTGGACAACAAGCGCTTCTT**CCGAAGA**AAGAAGCGCTTGTTGTCCAC-3' (1191), Pur $\alpha$ -III) 5'-GACAACAAGCGCTTCTTCTT**CCGAAGA**AAGAAGCGCTTGTTGTC-3' (1194), Pur $\alpha$ -IV): 5'-GCGTGTTTATGCGAGTGAGTG**CGAAC**ACTCACTCGCATAAACACGC-3' (1237), Pur $\beta$ -Scm) 5'-GATCCTAAGTCTGACTTGCAAC**GAA**GGTCAATCCTATAGTGCTAAG-3', Pur $\alpha$ -Scm1) 5'-GTCATCGAATGCCATGTCAGT**CGAA**ACTGACATGGCATTTCGATGAC-3', Pur $\alpha$ -Scm2) 5'-GGTATGCGTTAGTGCTGAGTG**CGAAC**ACTCAGCACTAACGCATACC-3'. Numbers in parentheses to the right of each sequence indicate the first nucleotide in the open reading frame of either Pur $\alpha$  or Pur $\beta$  mRNA predicted to be targeted by that particular shRNA. Bold type indicates the position of the loop sequence. Sequence 5' of the loop corresponds to the sense strand of the transcribed shRNA. Scrambled sequences (Scm) that contained similar base content as the Pur $\alpha$ - or Pur $\beta$ -specific shRNAs were

designed for use as negative controls in knockdown experiments. Complementary oligonucleotides sequences encoding the shRNAs were generated by chemical synthesis and obtained from a commercial vendor (Sigma-Genosys). To facilitate unidirectional cloning into pENTR™/U6 shRNA expression vector (Invitrogen), CACC was included on the 5' end of sense strand oligonucleotides while AAAA was included on the 5' end of antisense strand oligonucleotides. A double-stranded DNA insert encoding a LacZ-specific shRNA was supplied by the manufacturer. Annealing of complementary strands, ligation into pENTR™/U6, and subsequent transformation into *E. coli* strain TOP10 were performed as recommended by the manufacturer. Colonies were selected in the presence of 25 µg/ml kanamycin. *E. coli* were propagated in 5 ml cultures in the presence of antibiotic. Plasmids were isolated with the use of a mini-preparation kit (Qiagen). Insertional and sequence integrity was verified by automated DNA sequencing performed in the DNA Analysis Core Facility of the Vermont Cancer Center.

***Construction of lentiviral Pur shRNA expression constructs and transduction of AKR-2B cells*** – Selected U6 RNAi expression cassettes (Purβ-I or Purβ-Scm) were transferred from pENTR™/U6 vector into pLenti6/BLOCK-iT™-DEST vector via a site-specific LR recombination reaction as directed by the manufacturer (Invitrogen). Recombination reactions were used to transform One Shot® Stbl3™ competent *E. coli*. Recombinant subclones were selected based on resistance to blasticidin (50 µg/ml) and ampicillin (100 µg/ml) and sensitivity to chloramphenicol (30 µg/ml). Successful transfer of each U6 RNAi cassette into pLenti6/BLOCK-iT™-DEST vector was confirmed by restriction enzyme digestion of purified plasmids with *NdeI*. Cloned expression plasmids

were maintained in One Shot® *ccdB* Survival™ *E. coli* cells and purified by double cesium chloride gradient ultracentrifugation. To generate lentiviral stocks, pLenti6/BLOCK-iT™-DEST/U6 RNAi expression constructs (3 µg) were co-transfected along with the ViraPower™ packaging mix (9 µg) into the 293FT producer cell line (6 × 10<sup>6</sup> cells per transfection) using Lipofectamine™ 2000 reagent in serum-free Opti-MEM® I medium as directed by the manufacturer (Invitrogen). Cells were maintained for 16 h at 37° C in a humidified 5% CO<sub>2</sub> incubator. Transfection medium was replaced by complete growth medium supplemented with 1 mM sodium pyruvate and cells were cultured for an additional 72 h. Cell supernatants containing lentivirus were collected and titers of each lentiviral construct were determined based on transduction efficiency of cultured AKR-2B MEFs under blasticidin-resistant selection conditions.

To isolate stable lentiviral-transduced AKR-2B MEF clones, subconfluent AKR-2B cells (4 × 10<sup>5</sup> in a 25 cm<sup>2</sup> flask) were infected with 250 transducing units of lentivirus in McCoy's 5A medium (Gibco BRL/Invitrogen) with 5% v/v FBS for 24 h. After an additional 24 h incubation in growth medium without virus, cells were switched to selection medium (McCoy's 5A, 5% v/v FBS with 50 µg/ml blasticidin) and cultured for ten days with periodic replacement of selection medium in order to remove dead (non-transduced) cells. Once a suitable number of blasticidin-resistant cells were obtained (~25-50% coverage of the flask with densely colonized cells observed by microscopy), diluted cell suspensions were prepared by trypsinization and resuspension in selection medium. Cells were counted and then seeded by serial 2-fold dilution in a 96-well plate starting at 500 cells/ml (50 cells/well) to allow for colony growth and expansion from

single cell clones. Cell colonies arising in wells seeded with the lowest cell density ( $\leq 1$ ) and exhibiting circular morphology (suggestive of originating from a single cell) were once again subcloned by limiting dilution in selection medium (20 days post-infection) to ensure clonality. Lentiviral-transduced MEF clones were then expanded to prepare freezer stocks. Clonal cell lines were maintained in growth medium consisting of McCoy's 5A, 5% v/v FBS with 10  $\mu\text{g/ml}$  blasticidin at 37°C in a humidified 5%  $\text{CO}_2$  incubator. Confluent cultures of lentiviral-transduced MEF cell clones were washed with cold PBS then extracted with 1 $\times$  CAT lysis buffer (Roche Applied Science) plus protease inhibitors. After centrifugation of whole cell extracts for 10 min at 14,000 rpm, protein content in cleared lysates was measured by BCA assay (Sigma). Relative expression of Pur $\alpha$ , Pur $\beta$ , SM $\alpha$ A, and GAPDH was assessed by Western blotting.

***Construction of promoter-reporter plasmids and expression vectors*** – Mouse SM $\alpha$ A promoter-reporters and mammalian expression plasmids encoding His-epitope tagged versions of mouse Pur $\alpha$  and Pur $\beta$  were described previously (28, 296). All plasmids used for transfection were purified from *E. coli* cultures by double cesium chloride gradient centrifugation.

***Western blotting of transgene-expressed and endogenous proteins*** – Ectopic expression of His-tagged Pur $\alpha$  or Pur $\beta$  was monitored by Western blotting of lysed cell protein with an RGS(H)<sub>4</sub> monoclonal antibody (Qiagen) as previously described (145). Expression of endogenous Pur proteins was similarly assessed with a rabbit polyclonal antibody that recognizes a conserved sequence present in both mouse Pur $\alpha$  and Pur $\beta$  (141). Commercial monoclonal antibodies were used for detection of SM $\alpha$ A (clone 1A4,

Sigma) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (clone 6C5, Research Diagnostics Inc.).

***Chromatin immunoprecipitation (ChIP) assay*** – AKR-2B MEFs were seeded in 10 cm dishes at a density of  $5 \times 10^5$  cells in growth medium consisting of McCoy's 5A with 5% heat-inactivated FBS. After reaching 70-80% confluency (~36 h), cells were washed twice and incubated in 10 ml serum-free MCDB-402 medium (JRH Biosciences) for 48 h in a humidified 37°C, 5% CO<sub>2</sub> incubator. A separate group of cells was left untreated and allowed to grow to near confluence in complete growth medium prior to fixation. Formaldehyde was added to culture medium of growth-arrested (serum-free) or asynchronously growing cells to a final concentration of 1% v/v. After 10 min at 37°C, medium was aspirated and the cells were washed with phosphate-buffered saline and harvested by scraping. Cells were counted and collected by centrifugation at 1500 rpm for 10 min. Cell pellets were extracted using a ratio of 0.2 ml lysis buffer (1% w/v SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8.1) per  $10^6$  cells. Cell lysates were sonicated (4 × 10 s bursts) using a Branson model 150 sonifier at maximum power and then centrifuged at 14,000 rpm for 10 min at 4°C. Supernatant was removed and diluted 10-fold in ChIP dilution buffer consisting of 0.01% w/v SDS, 1.1% v/v Triton X-100, 1.2 mM EDTA, 167 mM NaCl, 16.7 mM Tris-HCl pH 8.1 plus 1 mM PMSF and 1 µg/ml each aprotinin, pepstatin A and leupeptin. A sample of the diluted cell lysate was set aside at this step for later reverse crosslinking and isolation of input DNA for use as a positive control in PCR amplification reactions.

For immunoprecipitation, each 2 ml sample of diluted cell lysate was pre-incubated at 4°C for 1 h with 75 µl protein A agarose blocked with salmon sperm DNA (Upstate Cell Signaling Solutions). After centrifugation, the pre-cleared lysate was combined with 4 µg of primary rabbit polyclonal antibody against Puro $\alpha$  (A291-313), Puro $\beta$  (B302-324), or MSY1 (M242-267) (141). No primary antibody and non-immune rabbit IgG controls were included as well. After overnight incubation at 4°C, 60 µl blocked protein A agarose was added to each sample and incubated for 1 h at 4°C. Samples were centrifuged at 1000 rpm at room temperature for 1 min. After removing the supernatant, agarose pellets were washed sequentially with 1 ml of the following buffers: low salt wash buffer (0.1% w/v SDS, 1% v/v Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl pH 8.1), high salt wash buffer (0.1% w/v SDS, 1% v/v Triton X-100, 2 mM EDTA, 500 mM NaCl, 20 mM Tris-HCl pH 8.1), LiCl wash buffer (0.25 M LiCl, 1% v/v IGEPAL CA-630, 1% w/v sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl pH 8.1). Pellets were then washed twice with 1 ml TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) and immune complexes were eluted by resuspending the agarose pellets twice in 250 µl elution buffer (1% w/v SDS, 0.1 M NaHCO<sub>3</sub>). Eluates were combined, supplemented with 20 µl 5 M NaCl, and incubated at 65°C overnight to reverse the formaldehyde crosslinks. Samples set aside for isolation of input DNA were processed similarly here and at subsequent steps. Following reverse crosslinking, proteins were digested by addition of 10 µl 0.5 M EDTA, 20 µl 1 M Tris-HCl pH 6.5, and 1 µl 20 mg/ml proteinase K and incubating at 45°C for 1 h. Samples were then sequentially extracted with an equal volume of buffered phenol, phenol:chloroform:isoamyl alcohol (25:24:1), and chloroform. Genomic DNA

fragments in the aqueous phase were precipitated by addition of 1/10 volume 3 M sodium acetate pH 5.2 plus 20 µg each of inert carriers glycogen (Sigma) and yeast tRNA (Sigma) and two volumes of ice cold 100% ethanol. After an overnight incubation at – 20°C, precipitated DNA was collected by centrifugation at 14,000 rpm for 20 min at 4°C. Pellets were washed with 500 µl 70% (v/v) ethanol and centrifuged again. Each pellet was dissolved in 50 µl 10 mM Tris pH 8.0 for use in subsequent PCR reactions.

The 5' flanking region of the mouse SM $\alpha$ A gene (382 bp product) was amplified by PCR using the following primers (5'-TTCTGAGGAATGTGCAAACCGTG-3' and 5'-GGCTACTTACCCTGACAGCGACT-3'). PCR cycling conditions were optimized based on the calculated melting temperatures of each primer duplex and by assessing the efficiency of product formation using “input” DNA samples equivalent to 1/50 or 1/100 of diluted cell lysate. For immunoprecipitated DNA samples, PCR mixtures contained 8.5 µl nuclease-free water, 1 µl 5 µM forward primer, 1 µl 5 µM reverse primer, 2 µl template DNA, and 12.5 µl AccuPrime™ SuperMix I (Invitrogen). No template control reactions included 2 µl of water in place of the DNA. Initial denaturation was conducted at 95°C for 5 min followed by 36 cycles of amplification. Each cycle entailed denaturation for 1 min at 95°C, annealing for 1 min at 63°C, and extension for 1.5 min at 72°C. PCR products were resolved by electrophoresis on 2% w/v SeaKem® LE agarose gels and illuminated by exposure to UV light. Images of ethidium bromide-stained bands were digitally captured with the Gel Doc™ XR system (Bio-Rad).

```

NM_008989|mPurA          CCACGCGGGCGCTCAAGAGCGAGTTCCTGGTGCGCGAAAACCGCAAGTACTACATGGAT 471
NM_011221|mPurB          CCGCGCCGCGCGCTCAAGAGCGAGTTCCTGGTGCGCGAGAACCGCAAGTACTACCTGGAC 438
** ***** *****

NM_008989|mPurA          CTCAAGGAGAACCAGCGCGGCCGCTTCCTGCGCATCCGCCAGACAGTCAACCG----- 524
NM_011221|mPurB          CTCAAGGAGAACCAGCGCGGCCGCTTCCTGCGCATCCGCCAGACAGTGAACCGCGGCGGC 498
*****

NM_008989|mPurA          -----AGCTGCCCGAGGGCACCTCCTTGACTGTGGACAACAAGCGCTTCTTCTTC 696
NM_011221|mPurB          CTGTACGGAGAGCTCCCGGAGGGCACCTCCATCACCGTGGACTCGAAGCGCTTCTTCTTC 738
*****

NM_008989|mPurA          GATGTGGGTTCACAACAAGTACGGGCGTGTATTATGCGAGTGAGTGAAGCCACCTAC 756
NM_011221|mPurB          GACGTCGGCTGACAACAAGTACGGGCGTGTTCCTGCGAGTGAGCGAGGTGAAGCCGCTCCTAC 798
** ** * * *****

NM_008989|mPurA          CGCAACTCCATCACCGTGCCCTACAAGGTGTGGCCAAGTTCGGACACACCTTCTGCAAG 816
NM_011221|mPurB          CGCAATGCCATCACCGTGCCCTTCAAGGCCTGGGGCAAGTTTGGAGGCGCCTTCTGTCCG 858
*****

NM_008989|mPurA          TACTCCGAGGAGATGAAGAAGATTCAAGAGAAACAGAGGGAGAAGCGGGCCGCTTGTGAG 876
NM_011221|mPurB          TATGCAGATGAAATGAAAGAGATCCAGGAGCGACAGAGGGATAAGC-----TTTACGAG 912
** * ** ** *****

NM_008989|mPurA          CAGCTCCACCAGCAGCAACAGCAGCAGCAAGAGGAGACCACCGCTGCCACCCTGCTACTG 936
NM_011221|mPurB          CGACGCGGC-GGGGGCAGCGGTGGCGGCACGAG-----TCG 948
* * * * * *****

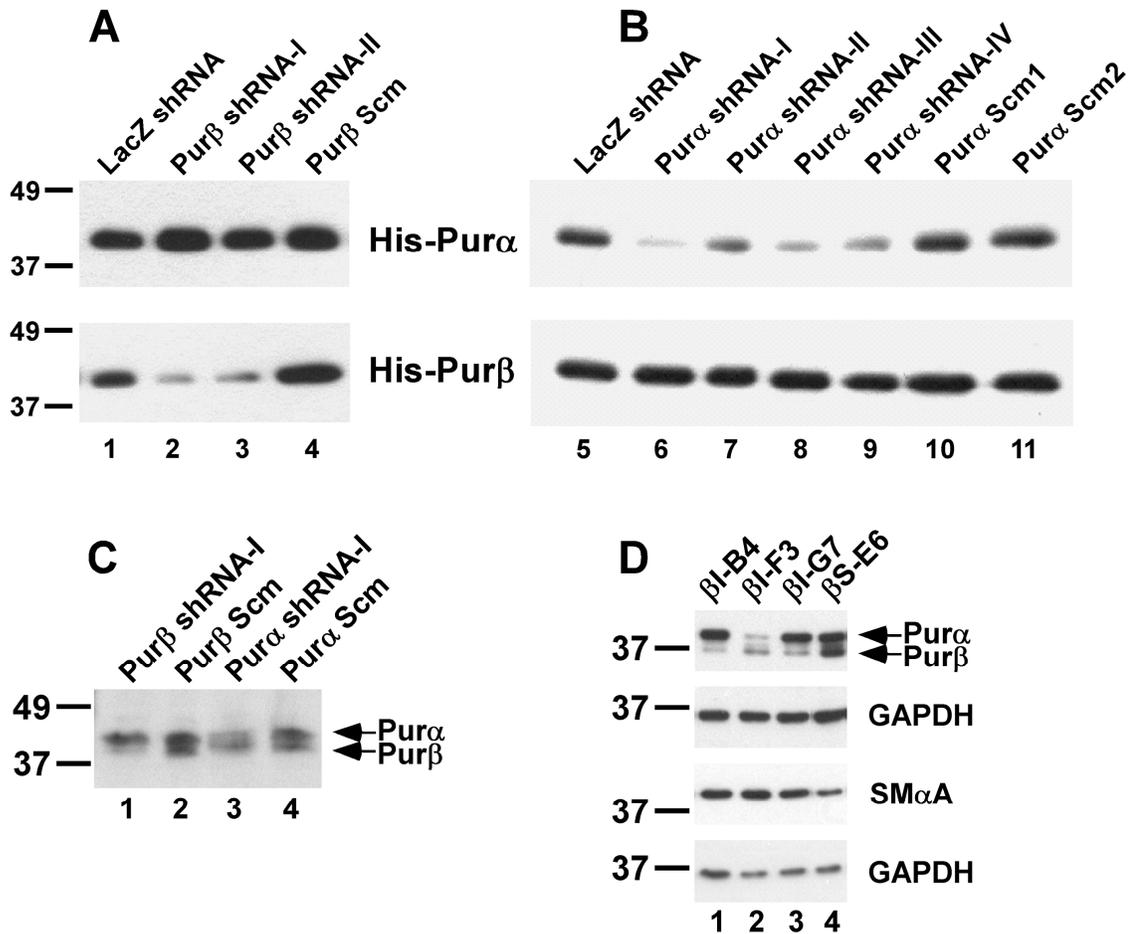
```

**Figure 3.1. Target positions of Pur $\alpha$ - and Pur $\beta$ -specific shRNA.** Open reading frame sequences of mouse Pur $\alpha$  (NM\_008989) and mouse Pur $\beta$  (NM\_011221) were aligned using ClustalW (European Bioinformatics Institute). Positions with base identity are denoted by an asterisk. Mouse Pur $\alpha$  and Pur $\beta$  show ~67% base identity in the open reading frames of cDNAs, thus making specific shRNA design difficult. Pur $\alpha$ -specific shRNA targets (blue) and Pur $\beta$ -specific shRNA targets (red) are indicated. Positions of these targets lie largely in regions of non-homology and are indicated in parentheses. Numbering assignments are based on translational start positions. Note these alignments are not complete.

## RESULTS

*Validation of shRNA-mediated knockdown of mouse Pur $\alpha$  or Pur $\beta$  in transfected fibroblasts* – To investigate the functional properties of endogenous Pur $\alpha$  and Pur $\beta$  in cultured fibroblasts and VSMCs, a loss-of-function approach utilizing shRNA-mediated RNA interference (RNAi) was undertaken. Mouse Pur $\alpha$  and Pur $\beta$  show ~67% base identity in the open reading frames of cDNAs, thus making specific shRNA design somewhat challenging. Nonetheless, taking advantage of several regions of non-homology, multiple shRNA sequences were identified by computer analysis with the theoretical potential to render the transcript encoding either mouse Pur $\alpha$  or Pur $\beta$  susceptible to destruction by the RNAi pathway (134). The sites of Pur $\alpha$ -, and Pur $\beta$ -specific transcript targeting by shRNA-mediated RNAi are shown in Figure 3.1. To test whether constructed U6 promoter-driven shRNA expression plasmids would generate functional shRNAs for knockdown of Pur $\alpha$  or Pur $\beta$  in cultured cells, immunoblotting of whole cell extracts from transiently transfected AKR-2B MEFs co-expressing His-tagged Pur $\alpha$  or Pur $\beta$  along with selected shRNAs was performed. His-tagged Pur transgenes were used because of low shRNA plasmid transfection efficiency and since endogenous Pur $\alpha$  and Pur $\beta$  migrate as a closely spaced doublet on Western blots making interpretation of specific knockdown somewhat ambiguous. Results shown in Figure 3.2 verify that transfection of AKR-2B cells with wild type Pur $\alpha$  or Pur $\beta$  shRNA constructs dramatically reduced the expression of either His-Pur $\alpha$  (Figure 3.2, panel B) or His-Pur $\beta$  (Figure 3.2, panel A) in a highly specific manner. Importantly, the Pur $\beta$ -specific shRNAs

had no discernable effect on His-Pur $\alpha$  expression while Pur $\alpha$ -specific shRNAs had no effect on His-Pur $\beta$  expression. LacZ shRNA or constructs encoding scrambled Pur shRNA sequences (Scm) served as negative controls for Pur $\alpha$  or Pur $\beta$ -specific knockdown (Figure 3.2). Similar results were obtained in transfected A7r5 cells indicating that the RNAi pathway is functional in VSMCs as well. These findings also reveal the necessity of empirically testing multiple shRNA sequences in cell types of interest as the efficiency of RNAi (i.e. relative level of knockdown) can vary significantly (Figure 3.2, panel B). To ensure that selected shRNAs would reduce endogenous Pur $\alpha$  or Pur $\beta$  expression in fibroblasts or VSMCs, several additional transfection/infection and immunoblotting experiments were performed. As shown in panel C of Figure 3.2, shRNA sequences identified as promoting specific knockdown of His-tagged Pur proteins, also facilitated knockdown of endogenous Pur $\alpha$  or Pur $\beta$  relative to scrambled control sequences when assayed in plasmid-transfected primary mouse VSMCs. Moreover, evaluation of multiple lentiviral-transduced AKR-2B MEF clones stably expressing Pur $\beta$  shRNA indicated definitive Pur $\beta$  knockdown relative to a control clone expressing scrambled sequence (Figure 3.2, panel D). As anticipated, there was some variability in the absolute level of Pur $\beta$  deficiency among different clones. One serendipitous clone was isolated which was also lacking in Pur $\alpha$  suggesting perhaps the ability of Pur $\beta$  to affect Pur $\alpha$  gene expression (200). Importantly, deficiency of one or both Pur proteins in these MEF clones correlated with an increase in smooth muscle  $\alpha$ -actin (SM $\alpha$ A) expression consistent with loss of Pur repressor activity. Relative differences in the levels



**Figure 3.2. Specific knockdown of Purα or Purβ by shRNA-mediated RNAi.** *A* and *B*, AKR-2B MEFs were transiently co-transfected with equal amounts of expression vector encoding His-Purα (*top panels*) or His-Purβ (*bottom panels*) plus the indicated shRNA expression plasmids. Cell extracts were prepared 48 h later. Total cellular protein (50 μg/lane) was separated by SDS-PAGE and transferred to Immobilon-PT<sup>TM</sup> membrane. Blots were probed with a monoclonal antibody recognizing the N-terminal RGS(H)<sub>4</sub> epitope tag (Qiagen). *A*, Blots show designed shRNAs specifically targeting knockdown of Purβ (*lanes 2 and 3*). *B*, Blots show designed shRNAs specifically targeting

knockdown of Pur $\alpha$  (*lanes 6–9*). Scm designates control constructs with scrambled sequences (*lanes 4, 10, and 11*). LacZ designates cells transfected with a LacZ shRNA construct (*lanes 1 and 5*). **C**, Primary C57BL/6J VSMCs were transiently transfected with 2  $\mu$ g of the indicated shRNA expression plasmids. After 48 h incubation in growth medium, whole cell extracts were prepared for analysis by immunoblotting (10  $\mu$ g protein/lane) with a rabbit polyclonal pan Pur antibody. **D**, Whole cell extracts of the lentiviral-transduced AKR-2B MEFs clones expressing Pur $\beta$  shRNA (*lanes 1–3*) or scrambled control sequence (*lane 4*) were assayed by immunoblotting to detect Pur $\alpha$  and Pur $\beta$  (10  $\mu$ g/lane) or SM $\alpha$ A (1  $\mu$ g /lane). Each blot was reprobbed with a GAPDH antibody to confirm equivalent protein loading. Numbers on the left denote the size of prestained protein markers in kilodaltons.

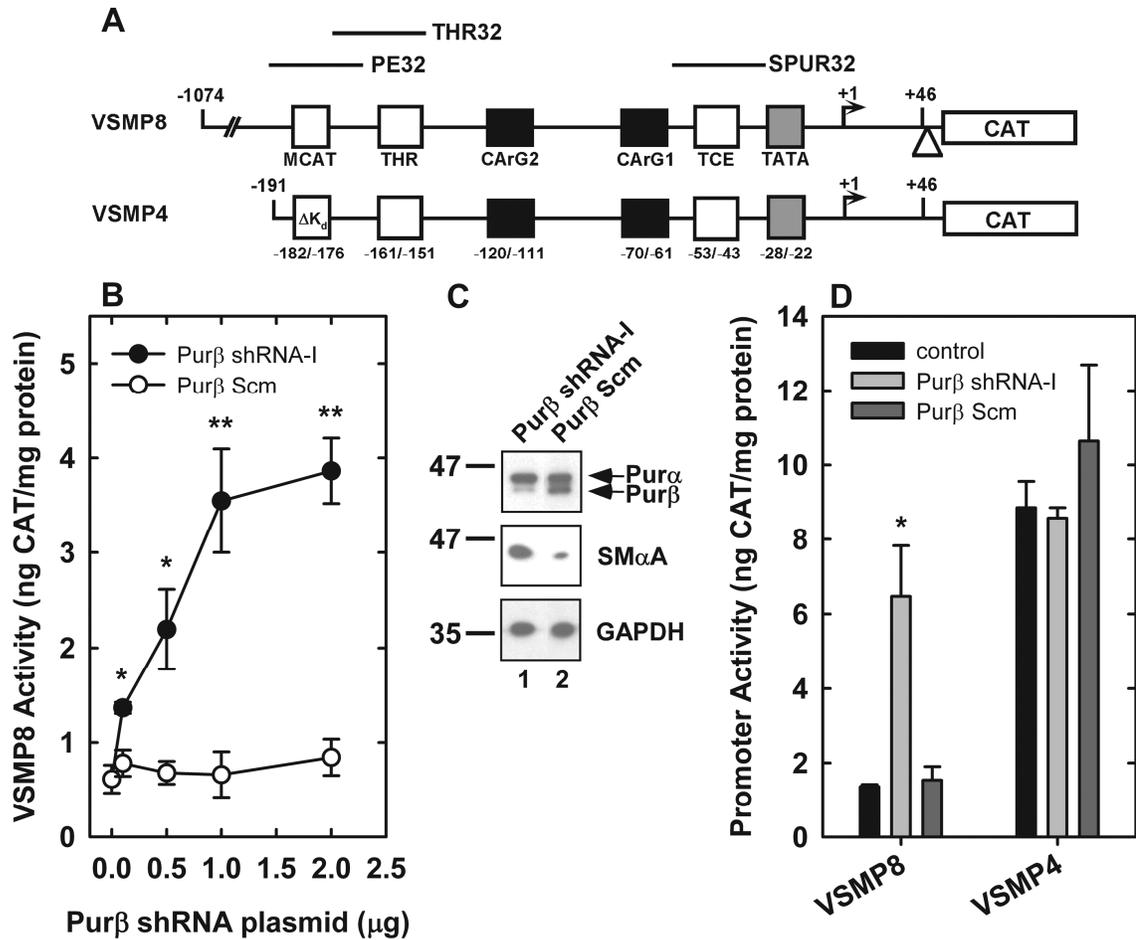
of Pur and SM $\alpha$ A expression detected in lentiviral-transduced clones were validated by reprobing immunoblots with anti-GAPDH as a loading control.

***Functional consequences of Pur $\alpha$  or Pur $\beta$  knockdown in fibroblasts and VSMCs –***

Given previous studies indicating that Pur $\alpha$  and Pur $\beta$  function as co-repressors of the SM $\alpha$ A promoter (28), we hypothesized that knockdown of Pur $\alpha$  and/or Pur $\beta$  in non-differentiated fibroblasts would result in promoter activation. A series of co-transfection experiments were conducted in AKR-2B MEFs with the use of a full-length SM $\alpha$ A promoter-CAT reporter known as VSMP8 (Figure 3.3, panel A) and selected Pur $\alpha$ - or Pur $\beta$ -specific shRNA expression plasmids. The VSMP8 promoter construct was chosen because it contains mouse SM $\alpha$ A sequence (–1074 through the first intron) required for smooth muscle-specific transgene expression *in vivo* (296) and, as such, exhibits very weak transcriptional activity in transfected AKR-2B cells due, in part, to negative regulation by endogenous Pur repressor proteins. As shown in Figure 3.3, panel B, co-transfection of Pur $\beta$  shRNA-I resulted in a dose-dependent increase in VSMP8 promoter activity while a scrambled control construct (Pur $\beta$  Scm) had no effect on this reporter. Corroborative results were obtained in lentiviral-transduced AKR-2B MEFs where stable shRNA-mediated knockdown of Pur $\beta$  augmented SM $\alpha$ A protein expression (Figure 3.3, panel C and Figure 3.2 panel D).

To evaluate whether the effect of Pur $\beta$  knockdown was promoter context-dependent, a truncated SM $\alpha$ A reporter known as VSMP4 was also tested for responsiveness to Pur $\beta$  deficiency. As previously documented (28), this mutant reporter exhibits unrestricted

MCAT enhancer activity due to the absence of 5'-nucleotides required for strand-specific Pur/Pyr element recognition by endogenous Pur repressors (see Figure 3.3, panel A for schematic). Hence it was not surprising to find that expression of Pur $\beta$  shRNA-I had no discernable effect on the transcriptional activity of VSMP4 in transfected MEFs (Figure 3.3, panel D). To assess whether knockdown of Pur $\alpha$  would yield analogous results, shRNA constructs demonstrating the maximum efficiency of Pur $\alpha$  or Pur $\beta$  knockdown (Figure 3.2, Pur $\alpha$  shRNA-I and Pur $\beta$  shRNA-I) were transfected either individually or in combination into AKR-2B MEFs together with VSMP8 or VSMP4 reporters (Figure 3.4, panel A). Knockdown of Pur $\alpha$  alone resulted in a modest ~2-fold enhancement in VSMP8 activity while knockdown of Pur $\beta$  alone induced VSMP8 by ~4-fold. Importantly, a synergistic response (~12-fold activation over reporter only control) was observed in AKR-2B MEFs expressing both Pur $\alpha$  and Pur $\beta$  shRNAs implying that endogenous Pur repressors likely function in a collaborative manner to regulate the transcriptional activity of the full-length SM $\alpha$ A promoter in this cell type. Transient transfection of the same Pur $\alpha$  or Pur $\beta$  shRNA constructs in primary mouse VSMCs yielded similar results although the extent of de-repression was less pronounced than seen in AKR-2B MEFs (Figure 3.4, panel B). Relative to VSMP8, VSMP4 exhibited little or no responsiveness to combined Pur $\alpha$  and Pur $\beta$  knockdown in both cell types. Evaluation of other SM $\alpha$ A reporter constructs containing differing lengths of 5'-flanking region indicated that significant promoter induction in response to co-knockdown of Pur $\alpha$  and Pur $\beta$  in AKR-2B cells minimally required 5' sequence extending to -195 (Figure 3.4,

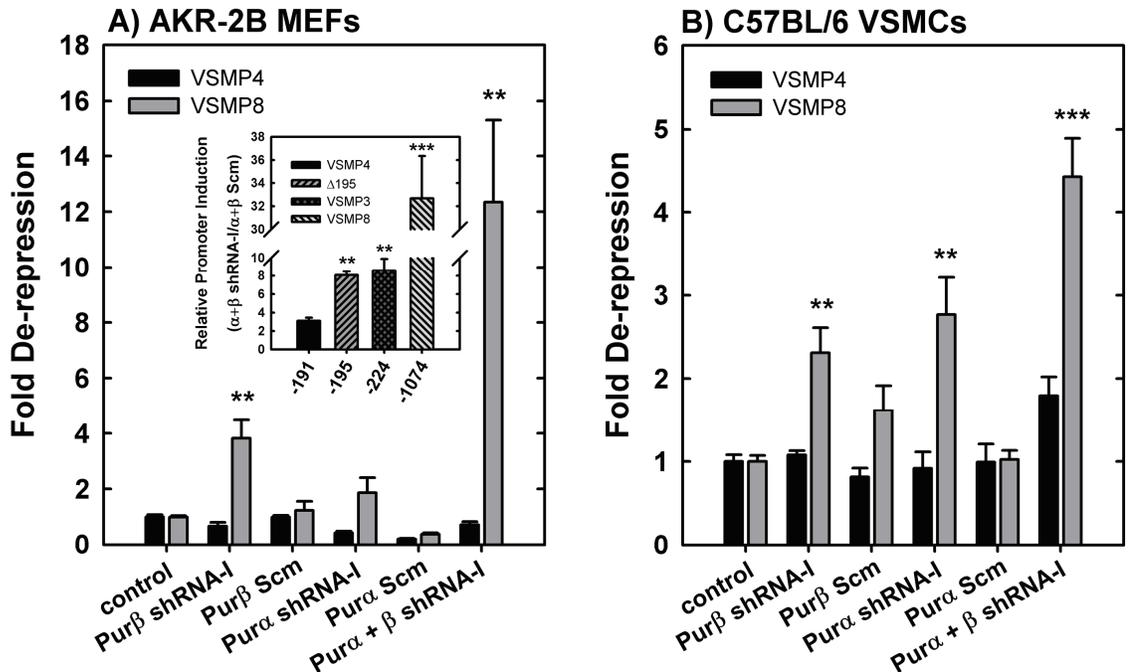


**Figure 3.3. De-repression of the SM $\alpha$ A promoter by Pur $\beta$  shRNA.** *A*, Schematic of mouse SM $\alpha$ A promoter-CAT reporter constructs shows core *cis*-elements mediating transcriptional activity (*boxes*) and binding sites (*PE*, *THR*, *SPUR*) for Pur $\alpha$ /Pur $\beta$  (18,26). VSMP4 lacks the ~2.5 kbp intron 1 (*triangle*), 5'-flanking region from -1074 to -192, and is transcriptionally activated due to exposure of a cryptic MCAT enhancer ( $\Delta K_d$ ) (18). *B*, Specific and dose-dependent activation of VSMP8 by Pur $\beta$  shRNA-I. AKR-2B cells were co-transfected with 2.8  $\mu$ g VSMP8, 0.2  $\mu$ g pCMV $\beta$ , selected amounts of plasmid encoding Pur $\beta$  shRNA-I or scrambled control sequence, and filler

DNA to 5.0  $\mu\text{g}/\text{dish}$ . Whole cell extracts were prepared 48 h after transfection and assayed for total protein and CAT reporter. **C**, Extracts of AKR-2B MEFs transduced with lentiviral vectors encoding Pur $\beta$  shRNA-I (*lane 1*) or scrambled sequence (*lane 2*) were analyzed by Western blotting for detection of Pur $\alpha$  and Pur $\beta$  (10  $\mu\text{g}$  protein/lane) or SM $\alpha$ A (1  $\mu\text{g}$  protein/lane). Note that the faster migrating Pur $\beta$  band is specifically reduced while SM $\alpha$ A is increased in shRNA-expressing cells. The SM $\alpha$ A blot was reprobed for GAPDH as a loading control. Numbers on the left denote the size of prestained protein markers in kilodaltons. **D**, De-repression of the SM $\alpha$ A promoter by Pur $\beta$  shRNA-I is promoter context-dependent. AKR-2B cells were co-transfected with VSMP8 or VSMP4  $\pm$  1  $\mu\text{g}$  of the indicated shRNA plasmids then assayed as described above. **B** and **D**, Promoter activity is expressed as CAT reporter divided by total cell protein (mean  $\pm$  SD). \*,  $p < 0.025$ ; \*\*,  $p < 0.01$  compared to reporter only control.

panel A inset, compare  $\Delta 195$  to VSMP4). This result coincides with previous biochemical studies indicating that high affinity binding by purified Pur $\alpha$  or Pur $\beta$  to the purine-rich strand of a Pur/Pyr element is impaired by transversional mutation of the GGA motif from -194 to -192 (28), as well as binding analyses showing that deletion of the GGA motif spanning positions -195 to -192 results in drastic impairment of Pur $\alpha$  and Pur $\beta$  binding (148).

***Relative levels of Pur $\alpha$ , Pur $\beta$ , and MSY1 SM $\alpha$ A promoter occupancy.*** To validate our findings that suggest Pur $\beta$  is the dominant Pur protein repressor of SM $\alpha$ A expression in MEFs we used a chromatin immunoprecipitation approach to assess relative levels of SM $\alpha$ A promoter occupancy under growth conditions shown previously to silence SM $\alpha$ A expression (9). Consistent with the notion of Pur $\beta$  repressor dominance shown in previous overexpression experiments (145) and in the present RNAi study, sequence-specific SSB occupancy of the SM $\alpha$ A promoter in the region spanning -322 to +58 is apparently dominated by Pur $\beta$  (Figure 3.5). Near identical antigen affinities of the antibodies employed validates these results (141), although effects of formaldehyde treatment, and relative immunoprecipitating antibody affinities have not been thoroughly investigated. Although highly qualitative, these results reinforce the idea that Pur $\beta$  is the dominant Pur protein repressor in the context of SM $\alpha$ A expression.



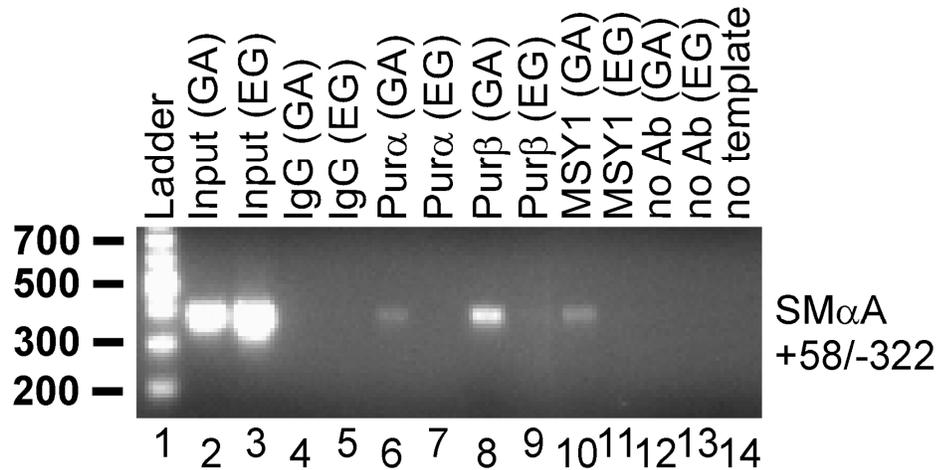
**Figure 3.4. De-repression of the SM $\alpha$ A promoter in response to knockdown of Pur $\alpha$  and/or Pur $\beta$  and the requirement for Pur/Pyr element integrity.** *A* and *B*, AKR-2B MEFs or primary C57BL/6J VSMCs were transiently transfected with 2.8  $\mu$ g VSMP8 or VSMP4 reporters, 0.2  $\mu$ g pCMV $\beta$ , and 1.0  $\mu$ g of expression plasmid encoding the indicated shRNA or scrambled control sequence. pBLCAT3 was used as filler to equalize the amount of DNA transfected at 5.0  $\mu$ g/dish. Whole cell extracts were prepared 48 h after transfection and assayed for total protein and CAT reporter. To compare the relative effect of shRNAs or scrambled controls on each promoter, corrected CAT values were normalized to values obtained in transfectants in which VSMP8 or VSMP4 were co-transfected with pBLCAT3 filler DNA only (control activity defined as 1 for each reporter). Results are expressed as fold de-repression (mean  $\pm$  SE). \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$  compared to reporter only control. *Inset*, AKR-2B MEFs were co-

transfected with the indicated SM $\alpha$ A promoter-reporter constructs and a combination of Pur $\alpha$  plus Pur $\beta$  shRNA-I plasmids or scrambled control plasmids then assayed as described above. To assess the relative level of shRNA-mediated induction of each reporter, CAT values obtained in Pur $\alpha$  plus Pur $\beta$  shRNA co-transfectants were divided by CAT values measured in scrambled control co-transfectants. Results are expressed as relative promoter induction (mean  $\pm$  SE). \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$  compared to VSMP4.

## DISCUSSION

Pur $\alpha$  and Pur $\beta$  are highly homologous proteins exhibiting ~70% primary sequence identity (143). The highest levels of conservation are restricted to the central modular repeat region of each molecule (10, 11) where the minimal nucleic acid-binding domain resides (11, 143, 212, 310). Not surprising then is the finding that these proteins display near-identical affinity for ssDNA oligonucleotides harboring PUR-elements (310). Regions of primary sequence divergence within the amino- and carboxy-termini have been implicated in directing specific interactions of Pur $\alpha$  or Pur $\beta$  with auxiliary protein and/or nucleic acid-binding partners (92, 145). This finding has fueled speculation that these regions of non-homology also confer paralog-specific function, as indicated by the distinct transcriptional properties ascribed to Pur $\alpha$  and/or Pur $\beta$  in different cell types and promoter contexts (54, 74, 162, 176, 214, 237, 238, 255, 256, 291, 303, 322). To test this speculative view, we have utilized parallel loss-of-function approach to assess Pur-paralog functional differences in the context of SM $\alpha$ A repression. These studies were published in a report describing parallel gain-of-function analyses, as well as biochemical dissection of protein/DNA and protein/protein interaction profiles comparing Pur $\alpha$  and Pur $\beta$  (148). Collectively these results confirm that paralog and cell type-specific functional differences exist between Pur $\alpha$  and Pur $\beta$  in the context of SM $\alpha$ A repression, *cis*-element recognition, and transcription factor interaction profiles.

Mechanisms of ssDNA generation/binding by Pur proteins have consistently shown that Pur $\alpha$  and Pur $\beta$  can associate with purine-rich ssDNA sequences in a sequence-



**Figure 3.5. Relative Levels of Pur $\alpha$ , Pur $\beta$ , and MSY1 occupancy at the SM $\alpha$ A promoter by ChIP.** Growth-arrested (GA) or exponentially growing (EG) AKR-2B MEFs were fixed with 1% formaldehyde. Nucleoprotein complexes were isolated from sonicated whole cell extracts equivalent to  $10^6$  cells by immunoprecipitation with rabbit polyclonal antibodies against mouse Pur $\alpha$  (lanes 6 and 7), Pur $\beta$  (lanes 8 and 9), or MSY1 (lanes 10 and 11). Non-immune rabbit IgG (IgG, lanes 4 and 5) or no primary antibody (no Ab, lanes 12 and 13) were included as negative controls. Complexes eluted from protein A agarose were processed to isolate genomic DNA for use as a template in PCR amplification of the indicated SM $\alpha$ A promoter region. Genomic DNA equivalent to  $\sim 1/100$  of the amount used for immunoprecipitation (Input, lanes 2 and 3) or buffer only (No template, lane 14) were used as positive and negative controls, respectively, for the PCR reaction.

specific fashion as either homo- or heteromultimers of varying degrees depending upon sequence and length of ssDNA (28, 141, 143, 144, 201, 310). RNA-mediated self-association of Pur $\alpha$  (91) and interaction between Pur $\alpha$  and Pur $\beta$  in the absence of nucleic acid has also been reported (141, 142) although the stoichiometric extent of Pur oligomerization and impact on nucleic acid binding remains unknown. The observation of these seemingly heterogeneous complexes from nuclear extracts suggests that collaborative activities of Pur $\alpha$  and Pur $\beta$  are necessary for *in vivo* repression of SM $\alpha$ A expression. Results from chromatin immunoprecipitation (Figure 3.5) and transient knockdown experiments in MEFs indicate that while Pur $\beta$  appears to be the dominant Pur protein repressor, Pur $\alpha$  also contributes in a synergistic and *cis*-element-dependent manner, as suggested by simultaneous knockdown in the presence of the full length promoter (Figure 3.3 and Figure 3.4). These findings substantiate the idea that Pur $\alpha$  and Pur $\beta$  are distinct and authentic repressors of the SM $\alpha$ A promoter, whereas the observed derepression of SM $\alpha$ A expression in MEFs exhibiting stable Pur $\alpha$  and/or Pur $\beta$  knockdown also validates their authenticity (Figure 3.2, panel D).

Consistent with previous models of repression, our data support a functional cooperation between endogenous Pur $\alpha$  and Pur $\beta$  for cryptic MCAT enhancer regulation (28). However marked synergistic activation of the full-length SM $\alpha$ A promoter in response to simultaneous knockdown of both Pur $\alpha$  and Pur $\beta$  MEFs suggests that transcriptional activation in AKR-2B cells is not only the result of cryptic MCAT enhancer vacancy, but also a result of derepression at additional *cis*-elements and

liberation of other *trans*-activators influenced by Pur proteins in fibroblasts (47, 272). This explanation may explain the less significant response observed during simultaneous knockdown of Pur proteins in primary VSMCs (Figure 3.4, panel B). An alternate explanation is that the existing levels of endogenous Pur $\alpha$  and Pur $\beta$  in VSMCs are sufficiently low such that SM $\alpha$ A expression is constitutively active and knockdown effects are difficult to detect.

Delineation of this issue was goal of gain-of-function studies that complemented the loss-of-function studies presented here (148). The results of these investigations showed that forced overexpression of Pur $\beta$ , but not Pur $\alpha$ , confers repression in cultured clonal and primary VSMCs from both full-length promoter constructs (VSMP8) and truncated constructs lacking Pur-binding sites flanking the core MCAT enhancer (VSMP4). These findings contrast Pur $\alpha$  overexpression studies in AKR-2B MEFs which showed Pur $\alpha$ -mediated repression is restricted to the core enhancer element. These studies confirm that Pur proteins repress expression of SM $\alpha$ A by distinct mechanisms, dependent upon cell type. The preference for Pur $\beta$ , but not Pur $\alpha$ , to repress expression of SM $\alpha$ A at *cis*-elements proximal to the MCAT enhancer element also suggests that differences exist either in the transcription factor interaction profiles between Pur $\alpha$  and Pur $\beta$  or their inherent binding affinities to these elements.

Interaction profiles of Pur $\alpha$  and Pur $\beta$  versus transcription factors known to interact and activate transcription of SM $\alpha$ A from other numerous *cis*-elements were generated by an ELISA-based method (elements and factors detailed in Figure 1.1)(28, 47, 145, 272, 319). The results of this analysis showed that indeed Pur $\alpha$  and Pur $\beta$  display differential

affinity for other known transcriptional activators. In both AKR-2B and A7r5 clonal rat VSMCs, Pur $\beta$  shows greater interaction potential with MSY1, TEF-1, and Sp1/3 than does Pur $\alpha$ , although slight cell type and subcellular compartmental differences were noted. The elevated interaction of Pur $\beta$  with Sp1/3 is particularly interesting due to recent findings which showed the capacity of Pur proteins to restrict enhancer activity by displacing Sp1/3 and Smad proteins from THR and SPUR elements (47, 272). *In vitro* binding studies examining the apparent affinities of recombinant Pur $\alpha$  and Pur $\beta$  for ss/dsDNA oligonucleotides representative of the MCAT enhancer element (PE32), THR, and SPUR show that both Pur $\alpha$  and Pur $\beta$  display high, near identical affinities for the purine-rich strands of the PE and SPUR elements. Both proteins show reduced affinity for either strand of THR, a predictable finding based on observed purine content. The results of ss/dsDNA-binding studies indicate that differences in repressive capacities observed between Pur $\alpha$  and Pur $\beta$  are not due to differences in the abilities of these proteins to recognize and bind cis-elements, but likely due to distinct capacities to interact with or displace transcriptional activators.

Collectively these findings support a model that suggests that reprogramming of constitutive SM $\alpha$ A in activated SMCs towards a repressive state likely involves deactivation of Sp1/3:SPUR enhancer complex by displacement caused by increased nuclear concentrations and binding of Pur $\beta$  to this element. Furthermore, compartmentalization of Pur $\alpha$  to the nucleus may permit this mechanism. Maintenance of the repressed state, as it exists in fibroblast-like cell types, likely proceeds through restriction of the MCAT enhancer element by both Pur $\alpha$  and Pur $\beta$ . This model also

suggests that distinct functions of Pur $\alpha$  and Pur $\beta$  aid in the plasticity of SM $\alpha$ A expression in response to environmental cues. The identity of environmental factors that upregulate, compartmentalize, or 'activate' Pur proteins for genetic and phenotypic reprogramming of VSMCs remains to be determined.

## **CHAPTER IV. SOLUTION CHARACTERIZATION REVEALS THE QUATERNARY STRUCTURE OF RECOMBINANT MOUSE PUR $\beta$**

*The experimental methods, results and interpretations described herein have been published previously in a revised form:*

Ramsey JE, Daugherty MA, Kelm RJ, Jr. (2007) Hydrodynamic studies on the quaternary structure of recombinant mouse Pur $\beta$ . *J Biol Chem.*282(3):1552-60.

### **INTRODUCTION**

Pur $\alpha$  and Pur $\beta$  are members of a highly conserved family of nucleic acid-binding proteins related by primary structure and a propensity to interact with single-stranded DNA (ssDNA)<sup>1</sup> or RNA sequences rich in purine nucleotides (reviewed in (92)). Diverse regulatory functions attributed to these proteins include gene transcription, mRNA translation, cell growth, and cell cycle progression (7, 57, 58, 125, 133, 146, 155, 177, 265). Pur $\alpha$  and Pur $\beta$  are ~70% identical at the amino acid level (143). Biochemical investigation of deletion mutants has shown that each protein possesses a minimal ssDNA/RNA-binding domain composed of a unique set of highly homologous sequence repeats (70, 133, 145). However, significant sequence differences between Pur $\alpha$  and Pur $\beta$  existing near the N- and C-termini suggest that each protein may have evolved to perform distinct functions (143). Pur $\alpha$  and Pur $\beta$  have been reported to bind to a *PUR*-element in a highly asymmetric polypurine/polypyrimidine tract located in the 5' flanking region of

the mouse SM $\alpha$ A gene (28, 141). It has been hypothesized that strand-specific binding by Pur $\alpha$ /Pur $\beta$  to this element disrupts a core MCAT enhancer motif thereby repressing SM $\alpha$ A promoter activity in cultured fibroblasts and vascular smooth muscle cells (28). Despite biochemical similarities, gain-of-function studies suggest that Pur $\alpha$  and Pur $\beta$  are not redundant in terms of their transcriptional repressor activity toward the full-length mouse SM $\alpha$ A promoter in transfected vascular smooth muscle cells (145).

In light of the specific protein-DNA, protein-RNA, and protein-protein interactions attributed to Pur $\alpha$  and Pur $\beta$ , and their potential relevance in modulating cell growth and differentiation, a need has emerged for the elucidation of the biophysical factors governing nucleic acid recognition. The mechanism of DNA-binding is of particular interest since DNA is thought to exist primarily in a double-stranded B-form helical configuration *in vivo*. Since Pur $\alpha$  and Pur $\beta$  preferentially bind to ssDNA or, in some cases, non-B-form structures (6), this aspect of Pur protein function has been investigated to a greater degree. Several reports have indicated that Pur $\alpha$  is capable of helix destabilization despite showing no intrinsic helicase or ATPase activity (56, 310). This has led to speculation that Pur proteins may promote strand displacement by opportunistic binding to transiently-formed ssDNA structures. This notion is supported by the finding that the 5'-MCAT enhancer region of the SM $\alpha$ A promoter has the propensity to adopt a partially unpaired configuration in response to transforming growth factor  $\beta$ 1 signaling (9). Another potentially important aspect of DNA-binding that was hinted at by previous studies involves the oligomeric state of Pur proteins in the presence and absence of ssDNA. Experiments reported by Kelm and coworkers (104), suggested

that Pur $\alpha$  and Pur $\beta$  bound to a SM $\alpha$ A promoter-derived PUR element as either hetero- or homodimeric nucleoprotein complexes. However, the conditions used in those experiments did not permit determination of whether or not dimerization was a prerequisite for, or a consequence of, ssDNA-binding. In this regard, a report by Gallia and colleagues has pointed to a critical role for RNA in mediating self-association of Pur $\alpha$  (91).

In the present study, we elucidate the quaternary structure of purified recombinant Pur $\beta$  by employing hydrodynamic and thermodynamic approaches to examine the macromolecular character of nucleic-acid free Pur $\beta$  in solution. Size exclusion chromatography coupled to static and dynamic light scattering-based detection systems revealed Pur $\beta$  to be an asymmetric protein capable of homodimeric self-association. This principal finding was confirmed by analytical ultracentrifugation which established that mouse Pur $\beta$  does indeed exist in a reversible monomer-dimer equilibrium characterized by a dissociation constant of  $\sim 1 \mu\text{M}$  in the absence of nucleic acid. Hydrodynamic analyses further suggested that homodimeric Pur $\beta$  assumes a non-spherical conformation in solution. We propose a model in which dimerization may affect ssDNA-binding in a manner regulated by a mass-action governed self-association of Pur $\beta$ . This type of a mechanism may be particularly relevant to pathophysiological states of the heart and vasculature where elevated Pur $\beta$  levels have been noted (104, 271).

## MATERIALS AND METHODS

*Expression and purification of nucleic acid-free N-HisPur $\beta$  – E. coli strain JM109* was transformed with the plasmid pQE30-N-HisPur $\beta$  which contains cDNA encoding an amino-terminal hexahistidine tagged mouse Pur $\beta$  fusion protein (N-HisPur $\beta$ ) (141). Transformants were selected after growth on LB-agar medium containing 40  $\mu$ g/ml carbenicillin (Sigma) for 16 h at 37°C. A single colony was used to inoculate 50 ml LB medium containing 40  $\mu$ g/ml carbenicillin. These starter cultures were allowed to grow for 8 h at 37°C to an optical density of  $>1.2$  at 600 nm, at which time 10 ml of starter was used to inoculate 1 liter of pre-warmed Terrific Broth II (QBiogene) with 40  $\mu$ g/ml carbenicillin. Cultures were allowed to grow at 37°C until an optical density of 0.6 at 600 nm was reached. Cultures were supplemented with ampicillin (Sigma) to 0.1 mg/ml and isopropylthiogalactopyranoside (Sigma) was then added to a final concentration of 1 mM to induce N-HisPur $\beta$  expression. After 4 h of additional growth, cells were harvested by centrifugation and cell pellets were stored at -80°C. Cell pellets equivalent to 4 liters were allowed to thaw in 20 ml of buffer A (50 mM sodium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, 10 mM  $\beta$ -mercaptoethanol) and completely resuspended on ice. Protease inhibitors leupeptin, aprotinin, and pepstatin A were each added to a final concentration of 1  $\mu$ g/ml and phenylmethanesulfonyl fluoride was added to 0.1 mM. Lysis was facilitated by the addition of egg white lysozyme (Sigma) to a final concentration of 1 mg/ml and incubated on ice for 20 min with occasional stirring. The cell suspension

was sonicated with a Branson Sonifier model 150 (setting 10) for a total of six 10 s bursts with 1 min incubations on ice between bursts to avoid overheating. Lysates were cleared by centrifugation at  $14,000 \times g$  for 30 min at  $4^{\circ}\text{C}$ . A total of 5 ml of 50% Ni-NTA agarose slurry (Qiagen) was added to the cleared lysate followed by 5M NaCl to a final concentration of 1 M. Bovine pancreatic DNase I and RNase A (Sigma) were each added to 59 and 92 units/ml, respectively (based on manufacturer specified activities) and the lysate-Ni-NTA agarose mixture was slowly rocked for 2 h at room temperature. The lysate-resin mixture was then gently centrifuged at  $1000 \times g$  for 2 min at  $4^{\circ}\text{C}$  to pellet the Ni-NTA agarose. Supernatant was removed and 25 ml of buffer B (50 mM sodium phosphate pH 8.0, 2 M NaCl, 10 mM imidazole, 10 mM  $\beta$ -mercaptoethanol plus protease inhibitors) was added to the resin. The mixture was then incubated approximately 14 h at  $4^{\circ}\text{C}$  with slow rocking. Gentle centrifugation at  $1000 \times g$  for 2 min at  $4^{\circ}\text{C}$  was used to pellet the resin which was subsequently resuspended in buffer A and loaded into a 1.5 cm diameter column. The rest of the purification procedure was carried out at room temperature. The resin was washed with buffer A until the absorbance of the flow through at 280 nm reached a baseline level ( $A_{280} \leq 0.02$  absorbance units). N-HisPur $\beta$  was eluted by application of buffer C (50 mM sodium phosphate pH 8.0, 300 mM NaCl, 500 mM imidazole, 10 mM  $\beta$ -mercaptoethanol). Eluted protein was concentrated using a centrifugal device (Centriprep YM-10, Millipore). Size exclusion chromatography was carried out on a 1.5 cm  $\times$  98 cm column packed with Sephacryl<sup>®</sup> 200 HR resin (Sigma) equilibrated in buffer E (50 mM sodium phosphate pH 7.5, 200 mM NaCl, 0.5 mM EDTA, 2 mM dithiothreitol) and run at a flow rate of 0.5 ml/min. The optical density of

the eluate was monitored with a Pharmacia model UV-10 UV/Vis detector. The column was calibrated using bovine serum albumin, ovalbumin, carbonic anhydrase, and cytochrome c (low molecular weight size-exclusion protein standards, Sigma). Fractions corresponding to the major dimeric peak of Pur $\beta$  were pooled and concentrated as described above. Protein purity was assessed by Coomassie Blue staining of N-HisPur $\beta$  reduced with 300 mM  $\beta$ -mercaptoethanol and resolved by SDS-PAGE on 12% (40:1 acrylamide:bisacrylamide) mini-gels. Preparations used for hydrodynamic studies were judged to be >95% homogenous N-HisPur $\beta$  under reducing conditions. To assess the level of nucleic acid-contamination, baseline-corrected absorbance spectra of purified N-HisPur $\beta$  were obtained using a Cary Bio100 dual beam spectrophotometer (Varian). A theoretical molar extinction spectrum of N-HisPur $\beta$  was calculated using SEDNTERP software (164) based on the method of Pace and coworkers (211). Protein concentration was determined spectrophotometrically assuming an extinction coefficient of 18,610 M<sup>-1</sup>cm<sup>-1</sup> at a wavelength of 280 nm and a monomeric relative molecular weight of 35,168.6 (calculated using SEDNTERP). To further test for the presence of nucleic acid, purified N-HisPur $\beta$  (1.4 mg in 500  $\mu$ l) was extracted twice with an equal volume of buffered-phenol, followed by extraction with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1), and an equal volume of chloroform. The aqueous phase of the extract was lyophilized and redissolved in ultrapure water three times and then dissolved one final time in 200  $\mu$ l of ultra pure water. Baseline corrected absorbance spectrum of the lyophilized extract was obtained to ensure the absence of a peak at 260 nm.

*Hydrodynamic analysis of N-HisPur $\beta$  by size-exclusion chromatography-laser light scattering and dynamic light scattering* – Molecular mass and hydrodynamic radius measurements of N-HisPur $\beta$  in solution were made by size-exclusion chromatography (SEC) coupled light scattering techniques. These experiments were performed at the W. M. Keck Foundation Biotechnology Resource Laboratory, Biophysics Facility, Yale University, New Haven, CT as described (81, 306). The analytical system consisted of a Superose 6 (Pharmacia-Amersham) size-exclusion column coupled to four detection systems configured in series including a UV absorbance detector (Kratos model 773 variable wavelength, Applied Biosystems), a laser (static) light scattering (LLS) detector (DAWN DSP, Wyatt Technologies), a refractive index detector (OPTILAB, Wyatt Technologies), and a dynamic light scattering (DLS) detection system (DYNAPRO TITAN, Wyatt Technologies). The DLS light source used was a DAWN EOS 633 nm laser (Wyatt Technologies). Solvent delivery was carried out by a Waters 510 HPLC pump (Waters Corp.) equipped with pulse-dampening transducers and a Rheodyne 7125 sample injection valve. Buffer E was used as the solvent in this analysis. Multiple volumes of pre-filtered (0.22  $\mu$ m Durapore, Millipore) N-HisPur $\beta$  (2.0 mg/ml) were injected and eluted at controlled flow rate of 0.3 ml/min.

LLS data was collected at an angle of 90° (the highest scattering signal) and analyzed by ASTRA software package (Wyatt Technologies) fitting implementations of the Zimm formalism of the Rayleigh-Debye-Gans light scattering model for dilute solutions of polymers, which relates the amount of scattered light to the concentration and weight average molecular weight of solute (81, 279, 306):

$$K^*c/R(\theta) = 1/M_w P(\theta) + 2A_2c \quad (\text{Equation 4.1})$$

In this relationship  $R(\theta)$  is the intensity of excess scattered light at angle  $\theta$ ,  $c$  is the concentration of the solute,  $M_w$  is the weight average molecular weight of the solute,  $A_2$  is the second virial coefficient,  $K^*$  is an optical parameter equal to  $4\pi^2 n^2 (dn/dc)^2 / (\lambda^4 N_A)$ ,  $n$  is the refractive index,  $N_A$  is Avogadro's number, and  $\lambda$  is the wavelength of the scattered light. The angular dependence of the scattered light is described by the function  $1/P(\theta)$ , whose first order expansion gives:

$$1/P(\theta) = 1 + (16\pi^2/3\lambda^2) \langle r_g^2 \rangle \sin^2(\theta/2) + \dots \quad (\text{Equation 4.2})$$

where  $\langle r_g^2 \rangle$  is the root mean square radius of gyration. The Zimm fitting formalism (81) was used for the determination of  $M_w$  of N-HisPur $\beta$ .

DLS measurements were made at an angle of  $111^\circ$  with a 2 second collection interval. Time resolved homodyne scatter intensity fluctuations were analyzed using Dynamics Software (Protein Solutions) which implements the cumulants method (152) to determine the time dependence of diffusive motion also referred to as the intensity autocorrelation function,  $G(\tau)$  (18, 213, 308):

$$G(\tau) = B[1 + \alpha \exp(-2D_T q^2 \tau)] \quad (\text{Equation 4.3})$$

where  $B$  is the average baseline intensity,  $\alpha$  is an instrument-specific correction factor,  $D_T$  is the concentration-dependent translational diffusion constant of the solute,  $\tau$  is a delay time, and  $q$  is the scattering vector equal to  $(4\pi n/\lambda)\sin(\theta/2)$ , in which  $n$  is the refractive index of the solvent,  $\lambda$  is the wavelength of the scattered light and  $\theta$  is the scattering angle. Equation 4.3 describes the relationship between the time dependency of fluctuation in scatter intensity and the translational diffusion coefficient. Generally, large molecules diffuse slowly and generate scatter signals that fluctuate slowly. The opposite is true for small molecules. The value of  $D_T$  was used to estimate the apparent hydrodynamic radius of an equivalent sphere by way of the Stokes-Einstein relationship:

$$R_h = kT/6\pi\eta D_T \quad (\text{Equation 4.4})$$

where  $k$  is the Boltzmann constant,  $T$  is the absolute temperature, and  $\eta$  is the temperature corrected viscosity of the solvent.

***Sedimentation velocity analytical ultracentrifugation of N-HisPur $\beta$***  – Sedimentation velocity experiments were carried out in a Beckman/Coulter Optima XL-I/XL-A equipped with an An50Ti rotor. Radial concentration distributions were measured at 50,000 rpm and 4°C using Rayleigh interference optics. Protein samples were gel filtered over a 1.5 cm  $\times$  98 cm Sephacryl® 200 HR size-exclusion column pre-equilibrated and eluted with buffer E (described above) as a final step prior to sedimentation. The fraction corresponding to the absorbance maximum of the presumed dimeric peak was used to prepare a dilution series of N-HisPur $\beta$  over a ten-fold

concentration range starting at 43.5  $\mu\text{M}$ . The reference buffer consisted of a buffer E eluate from the size exclusion column. Blank subtracted Rayleigh interference scans were recorded at 1 min intervals. Sedimentation velocity data were analyzed by the  $dc/dt$  method to generate apparent sedimentation coefficient distributions,  $g(s^*)$  (266), with the use of DCDT+ software (217). Direct fitting of time-resolved concentration difference ( $\Delta c$  versus radius) curves to numerical solutions to the Lamm equation describing multiple, interacting species models and kinetic models was performed using SEDANAL v4.3 software (268). Goodness of fits were judged by visualization of residuals and fitting statistics. Temperature corrected values for the partial specific volume of N-HisPur $\beta$  ( $\bar{v}$ ), as well as density ( $\rho$ ), and viscosity ( $\eta$ ), of buffer E were calculated using the program SEDNTERP. Resulting values are as follows:  $\bar{v} = 0.7109 \text{ ml g}^{-1}$ ,  $\rho = 1.0149 \text{ g ml}^{-1}$ , and  $\eta = 0.1635 \text{ g cm}^{-1} \text{ s}^{-1}$ . Molecular shape modeling was also carried out with SEDNTERP. Protein integrity was assessed after sedimentation by SDS-PAGE to ensure that the samples were intact.

*Sedimentation equilibrium analytical ultracentrifugation of N-HisPur $\beta$*  – Sedimentation equilibrium experiments were carried out in cells fitted with six-sector charcoal-Epon centerpieces (1.2-cm path). Protein and reference buffer samples were prepared as described above. Sample dilution series were made from the dimeric peak fraction over a ten-fold concentration range. Sedimentation was carried out at rotor speeds of 22,000, 28,000, and 35,000 rpm at 4°C. Five scans were averaged to remove noise. Equilibrium was judged to be achieved by the superposition of scans taken 6 h

apart and by analysis with the MATCH package implemented by HETEROANALYSIS software (49).

Blank corrected sedimentation equilibrium data were fit using HETEROANALYSIS software to numerous mathematical models describing radial concentration distributions including single ideal species (equation 4.5), single non-ideal species (equation 4.6), monomer-Nmer equilibria (equation 4.7), monomer-dimer with incompetent monomer (equation 4.8), monomer-dimer with incompetent dimer (equation 4.9) and monomer-Nmer-Qmer equilibria (equation 4.10):

$$A(r, \lambda) = \delta_\lambda + \varepsilon_\lambda \ell c_0 \exp[M^* \phi(r^2 - r_0^2)] \quad (\text{Equation 4.5})$$

$$A(r, \lambda) = \delta_\lambda + \varepsilon_\lambda \ell c_0 \exp[M^* \phi(r^2 - r_0^2) - 2A_2 M \{w(r) - w(r_0)\}] \quad (\text{Equation 4.6})$$

$$A(r, \lambda) = \delta_\lambda + \varepsilon_\lambda \ell c_0 \exp[M^* \phi(r^2 - r_0^2)] + N \varepsilon_\lambda \ell c_0^N \exp[\ln K_a + NM^* \phi(r^2 - r_0^2)] \quad (\text{Equation 4.7})$$

$$A(r, \lambda) = \delta_\lambda + \varepsilon_\lambda \ell c_0 \exp[M^* \phi(r^2 - r_0^2)] + \varepsilon_\lambda \ell c_0' \exp[M^* \phi(r^2 - r_0^2)] + 2 \varepsilon_\lambda \ell c_0^2 \exp[\ln K_a + 2M^* \phi(r^2 - r_0^2)] \quad (\text{Equation 4.8})$$

$$\begin{aligned}
A(r, \lambda) = & \delta_\lambda + \varepsilon_\lambda \ell c_0 \exp[M^* \phi(r^2 - r_0^2)] + \varepsilon_\lambda \ell d'_0 \exp[2M^* \phi(r^2 - r_0^2)] \\
& + 2\varepsilon_\lambda \ell c_0^2 \exp[\ln K_a + 2M^* \phi(r^2 - r_0^2)]
\end{aligned}$$

(Equation 4.9)

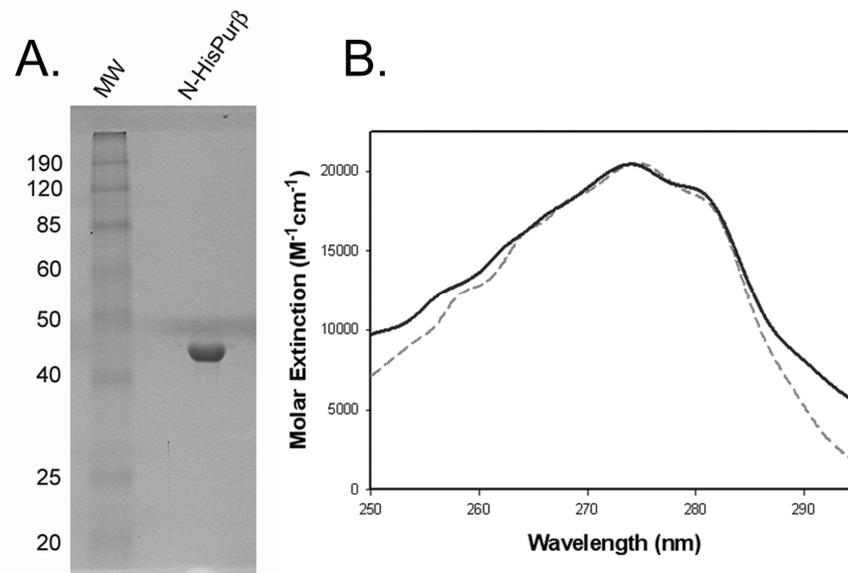
$$\begin{aligned}
A(r, \lambda) = & \delta_\lambda + \varepsilon_\lambda \ell c_0 \exp[M^* \phi(r^2 - r_0^2)] + N\varepsilon_\lambda \ell c_0^N \exp[\ln K_{a1} + NM^* \phi(r^2 - r_0^2)] + \\
& Q\varepsilon_\lambda \ell c_0^Q \exp[\ln K_{a2} + QM^* \phi(r^2 - r_0^2)]
\end{aligned}$$

(Equation 4.10)

where  $A(r, \lambda)$  is the radially-dependent absorbance at radial position,  $r$ , and wavelength,  $\lambda$ ,  $\delta_\lambda$  is the baseline offset,  $\varepsilon_\lambda$  is the molar extinction coefficient,  $\ell$  is the path length,  $c_0$  is the molar concentration of the monomer at the arbitrary reference radial position  $r_0$ ,  $c'_0$  is the molar concentration of the incompetent monomer at the arbitrary reference radial position  $r_0$ ,  $d'_0$  is the molar concentration of the incompetent dimer at the arbitrary reference radial position  $r_0$ ,  $A_2$  is the second virial coefficient,  $M$  is the monomer molecular weight,  $w(r)$  refers to the concentration of the monomer on a weight/volume scale at distance  $r$ ,  $w(r_0)$  is the weight/volume concentration at reference position  $r_0$ ,  $N$  is the stoichiometry of association,  $K_a$  is the association constant,  $Q$  is the stoichiometry of the higher order oligomeric association.  $M^*$  refers to the buoyant molecular weight, equal to  $M(1 - \bar{v}\rho)$ .  $\phi$  is equal to  $\omega^2/RT$ , in which  $\omega$  is the angular velocity,  $R$  is the gas constant, and  $T$  is the absolute temperature. Goodness of fit was judged by fit statistics and visual inspection of residuals for systematic deviations.

## RESULTS

*Purification of nucleic acid-free recombinant Pur $\beta$*  – In order to investigate the self-association of N-HisPur $\beta$  it was essential to ensure that preparations of the recombinant protein be devoid of co-purifying DNA and/or RNA. To accomplish this task, we developed a protocol that included nucleases during *E. coli* lysis and performed Ni-NTA agarose affinity chromatography under high ionic strength conditions to promote removal of weakly-associated nucleic acids. A final size exclusion chromatography step was also included to eliminate high molecular weight aggregates and low molecular weight fragments. Using the method described herein, N-HisPur $\beta$  was purified to homogeneity as judged by SDS-PAGE (Figure 4.1A). Although the molecular weight of N-HisPur $\beta$  calculated on the basis of its amino acid sequence is 35,168.6, it appears to migrate by SDS-PAGE as a ~43 kDa peptide under reducing conditions. The unusual electrophoretic mobility of the recombinant protein is consistent with the reported mobility of native Pur $\beta$  expressed in fibroblasts and vascular smooth muscle cells (28, 143). This suggests that the His tag is not the major contributing factor to the non-ideal electrophoretic behavior of N-HisPur $\beta$ . In order to assess the extent of nucleic acid contamination, we compared the absorbance spectrum (normalized to the molar extinction at 280 nm of 18,610 M<sup>-1</sup>cm<sup>-1</sup>) to a hypothetical molar extinction spectrum of N-HisPur $\beta$ , generated from amino acid content (Figure 4.1B). This comparison showed only minimal deviations between the calculated and experimental spectra in the region around 260 nm. It would be predicted that stoichiometric quantities of co-purifying

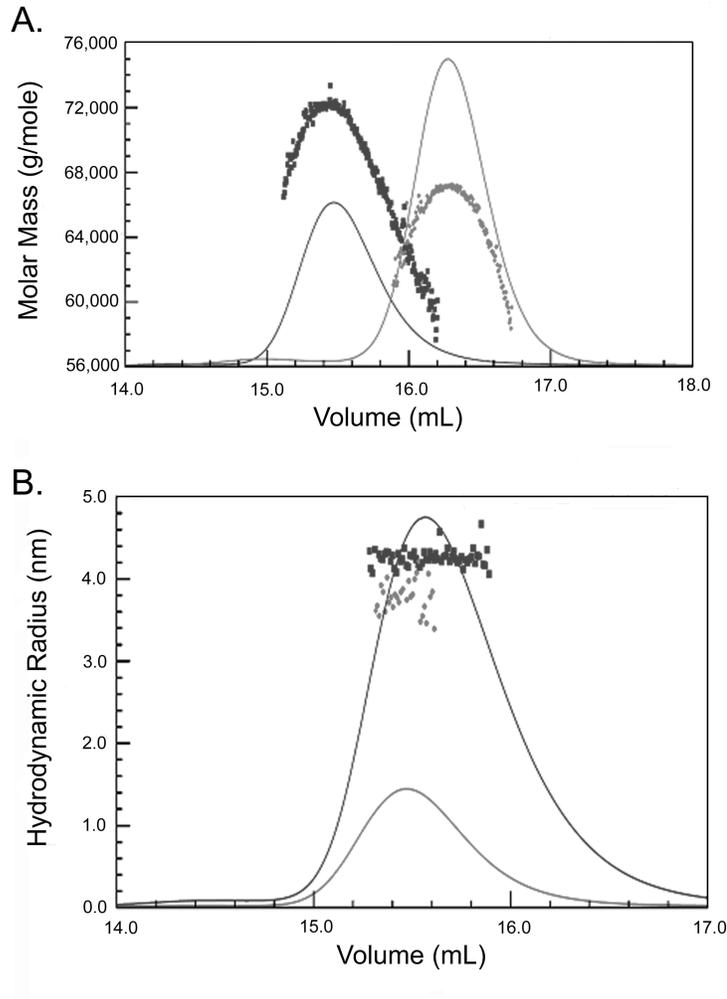


**Figure 4.1. Expression and purification of N-HisPur $\beta$ .** *A*, Reducing SDS-PAGE to analyze recombinant mouse Pur $\beta$ . 3.8  $\mu$ g of purified N-HisPur $\beta$  was loaded. Despite the fact that N-HisPur $\beta$  has a sequence predicated molecular mass of  $\sim$ 35 kDa, it shows an electrophoretic mobility of  $\sim$ 43 kDa relative to molecular weight standards (MW). *B*, Comparison of a calculated molar extinction spectra of N-HisPur $\beta$  (solid line) with an experimentally measured absorbance spectra of purified N-HisPur $\beta$  (normalized to calculated  $\epsilon_{280 \text{ nm}} = 18,610 \text{ M}^{-1} \text{ cm}^{-1}$ , dashed line). The correspondence of the curves confirms the absence of nucleic acid in preparations of N-HisPur $\beta$ .

nucleic acids would result in a large spectral difference in this range since nucleoside-5'-monophosphates have a molar extinction coefficient on the average of  $10^4 \text{ M}^{-1}\text{cm}^{-1}$  at or near 260 nm (32). Furthermore, absorbance spectra of phenol-chloroform extracts of purified N-HisPur $\beta$  preparations showed no species with a  $\lambda_{\text{max}}$  of 260 nm (data not shown). Collectively, these data indicate that the preparations of N-HisPur $\beta$  used in this study were free of co-purifying nucleic acids.

Buffer E was chosen as the buffer condition for all sedimentation and light-scattering experiments. We have determined that the solubility of N-HisPur $\beta$  relies heavily on ionic strength and reducing agent concentration (data not shown). Dialysis of protein at moderate concentrations (~ 1 mg/ml) leads to loss of protein (likely deposition on vessel surfaces, as sample can be recovered by addition of high salt buffers). The salt concentration of buffer E (200 mM NaCl) was determined to be the minimum quantity to limit this effect and to maximize stability.

***Hydrodynamic analysis of N-HisPur $\beta$  by size-exclusion coupled light scattering techniques*** – A series of physical techniques that make use of macromolecular light scattering phenomena were used to investigate the hydrodynamic properties of N-HisPur $\beta$  in solution. SEC-LLS-DLS is a well-suited means of investigating the hydrodynamic character of proteins. It is non-destructive and each individual light scattering detection technique can be performed in series after a size fractionation step. The use of SEC as a preliminary step to light scattering can eliminate some of the ambiguity created in performing weight-average measurements (81, 279, 306).



**Figure 4.2. Molecular size measurements of N-HisPur $\beta$  in solution by light scattering techniques.** N-HisPur $\beta$  was applied to a Superose 6 column and eluted with buffer E at high pressure. Elution was monitored by refractive index (RI) changes and analyzed by LLS (**A**) and DLS (**B**) to determine the solution characteristics of the recombinant protein. **A**, The RI traces (normalized to molar mass values, solid lines) and LLS molar mass measurements (points) of N-HisPur $\beta$  (dark gray) show an asymmetric elution pattern suggesting that the protein elutes as a mixture of self-associating and monomeric species, as compared to a bovine serum albumin standard (light gray). The

weight average molar mass of the elution peak determined by LLS was  $67.7 \pm 4.12$  kDa. **B**, As determined by DLS, a dependence of the weight average  $R_H$  on loading quantity was also noted. Solid lines represent RI traces whereas individual data points represent the DLS determined weight average hydrodynamic radius measurements of eluting N-HisPur $\beta$  collected at different loading concentration (19.9 pmoles in 350  $\mu$ l, dark gray; 5.1 pmoles in 100  $\mu$ l, light gray). Again, RI trace values were normalized to  $R_H$  values.

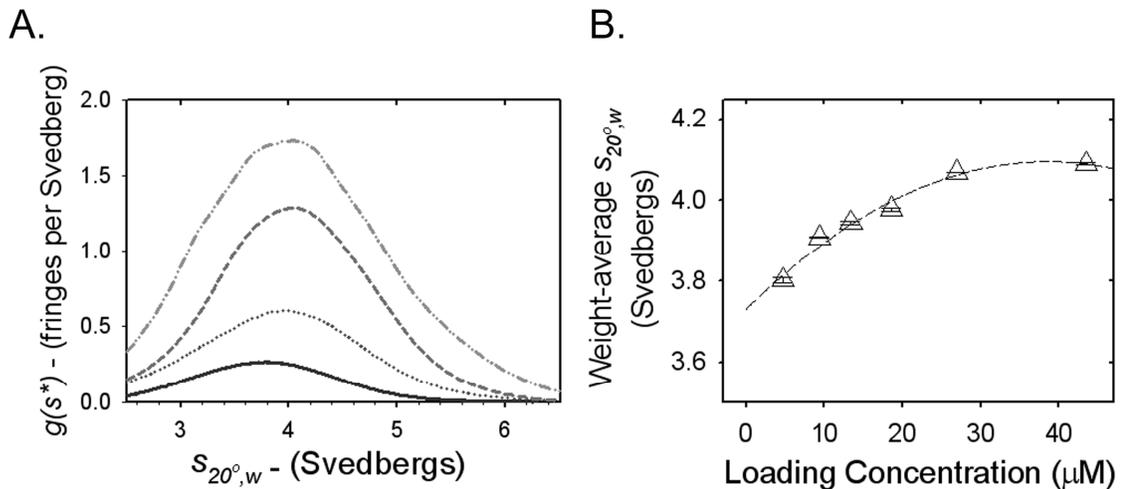
The results of applying the SEC-LLS to solutions of N-HisPur $\beta$  reveal an asymmetric distribution of refractive index (RI) and molar mass measurements in the sole eluting peak of N-HisPur $\beta$ , as compared to a bovine serum albumin standard (Figure 4.2A). These results are consistent with a polydisperse mixture that elutes as a single, albeit asymmetric peak. On the sole basis of RI signal, one might surmise that the asymmetric shape of the elution peak could arise from interaction of N-HisPur $\beta$  with the column matrix. Although this could conceivably cause a tailing effect on elution, it would not cause a broadening in the distribution of weight-average molar mass measurements as is seen for N-HisPur $\beta$  (Figure 4.2A). Instead, this effect is likely due to polydispersity. The number average of the weight average molar mass measurements across the elution peak is  $67.7 \pm 4.12$  kDa corresponding to a 6.08% degree of polydispersity. Based on the number average molar mass, this suggests that N-HisPur $\beta$  exists as an interacting mixture of monomers and dimers (expected dimeric  $M_r$  of 70,337.2).

Weight average hydrodynamic radii ( $R_H$ ) determinations on SEC fractionated N-HisPur $\beta$  by DLS were also consistent with a self-associating system. It was found that the number average  $R_H$  across the top 10% of the eluting peak (region of peak where scatter signal is strongest and  $R_H$  values are approximately constant) is dependent upon the loading quantity of N-HisPur $\beta$  (Figure 4.2B). The resultant  $R_H$  values for loading quantities of 19.9 pmoles and 5.1 pmoles were 4.3 nm and 3.8 nm, respectively. From these data, it is apparent that the weight average  $R_H$  of the eluting solution is dependent upon a mass-action governed self-association of N-HisPur $\beta$ . These  $R_H$  values are very different from the calculated  $R_H = 2.72$  nm for an equivalent non-compressible sphere

with a molecular mass of 70,337.3 Da. The  $R_H$  value of 4.3 nm is expected to result in a frictional coefficient ratio ( $f/f_0$ ) of 1.56, as determined by analysis with the Dynamics Software (Protein Solutions). Collectively, findings based on light scattering suggest the N-HisPur $\beta$  participates in a monomer-dimer equilibrium and that the dimeric form assumes an asymmetric shape in solution.

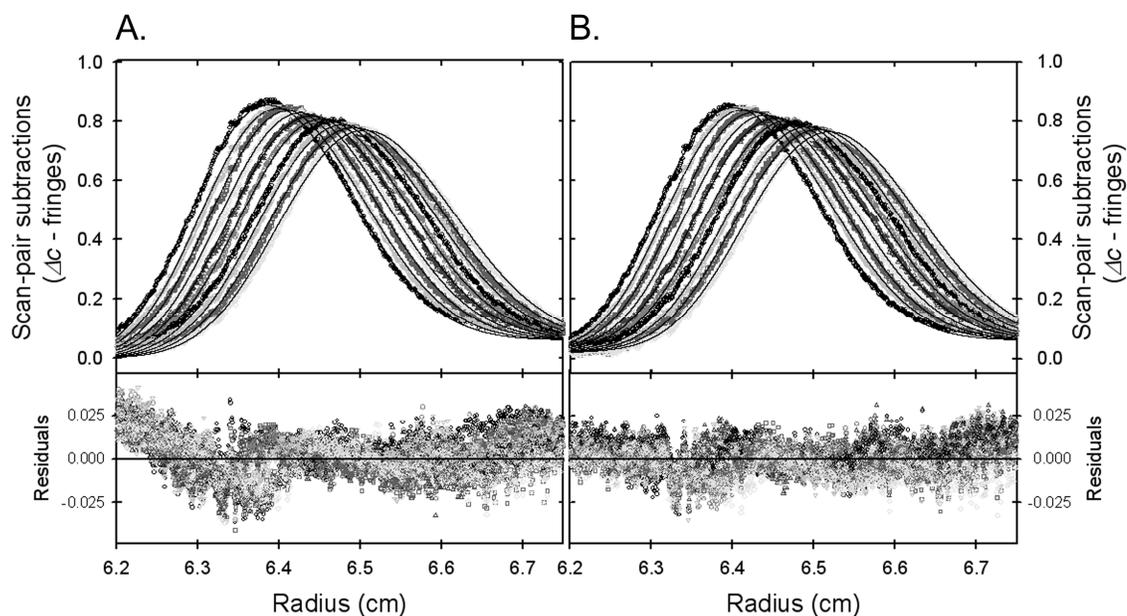
***Sedimentation velocity analysis of recombinant Pur $\beta$***  – Quantitative hydrodynamic and thermodynamic analyses aimed at investigating the oligomeric structure of N-HisPur $\beta$  in solution were performed by analytical ultracentrifugation. Sedimentation velocity experiments were carried out to investigate the hydrodynamic properties of recombinant Pur $\beta$  and to validate the observations made by SEC-LLS-DLS studies. The apparent sedimentation coefficient distribution function,  $g(s^*)$ , for solutions of N-HisPur $\beta$  at four loading concentrations ranging from 4.7  $\mu\text{M}$  to 43.5  $\mu\text{M}$  is shown in Figure 4.3A. Analysis of the normalized weight average sedimentation coefficient ( $s_{20,w}$ ) from  $g(s^*)$  distributions shows an increase as a function of loading concentration (Figure 4.3B).

The presence of a single peak in the  $g(s^*)$  data shown in Figure 4.3A may not imply a single sedimenting species. This is because the effects of diffusion in these experiments may hide heterogeneity, especially in the case where the sedimentation coefficients of all sedimenting species vary only slightly (216). However, the ensemble of data suggests a diffuse mixture of species whose sedimentation is dynamic and dependent upon the changing radial concentration distribution over time (267).



**Figure 4.3. Sedimentation velocity analysis of N-HisPur $\beta$ .** **A**,  $g(s^*)$  plots of a dilution series of N-HisPur $\beta$  made by collecting the major elution peak from a Sephacryl $^{\text{®}}$  200 HR size exclusion column. Data are shown for a 10-fold concentration range of protein from 4.7  $\mu\text{M}$  (—), 13.4  $\mu\text{M}$  (·····), 28.6  $\mu\text{M}$  (----) and 43.5  $\mu\text{M}$  (-••-). Data were collected at 50,000 rpm and 4°C. **B**, Weight average  $s_{20,w}$  as a function of loading concentration. These data show an increase in  $s_{20,w}$ , a further indication of a reversible self-associating system. Error bars represent uncertainty in the determination of the weight average value of  $s_{20,w}$ . The dashed line in **B** is intended to guide the eye.

In order to elucidate the number of sedimenting species in solutions of N-HisPur $\beta$ , direct fitting of sedimentation velocity data was employed. Despite only observing a single Gaussian peak in the  $g(s^*)$  distributions, fitting of radial  $\Delta c$  (subtraction of scan-pairs) data to a single sedimenting species model was poor relative to that of a monomer-dimer equilibrium model (Figure 4.4; as judged by an increase in randomness of residuals and by fitting statistics). Fitting to the single-species model revealed an apparent 3.97 Svedberg species with an apparent  $M_r = 66,890.1$ , which is lower than the expected molecular weight for dimeric N-HisPur $\beta$ . This is suggestive of enhanced diffusion due to self-association ( $M \sim s/D_T$ ). Further, fitting to an associating model revealed the sedimentation coefficients and molecular weights of both the monomeric and dimeric species. At a loading concentration of 13.4  $\mu\text{M}$ , it was found that the monomeric species sediments with a  $s_{20^\circ, w}$  of 1.79 (1.70-1.90) Svedberg with an apparent mass of 35.94 (35.63-36.29) kDa. This corresponds to  $R_H = 4.91$  nm and  $f/f_0 = 2.22$  when analyzed with SEDNTERP. The dimeric species was determined to sediment with a  $s_{20^\circ, w}$  of 3.961 (3.960-3.968) Svedberg, corresponding to  $R_H = 4.43$  nm and  $f/f_0 = 1.60$ . The  $R_H$  and  $f/f_0$  determined for the dimeric species are slightly greater than those found by light-scattering techniques but are in reasonable agreement. The disparity is likely due to the fact that DLS measurements are made in a bulk manner on a polydisperse system and, in turn, are weight-averages. Returned dissociation constants from direct boundary fitting to the reversible monomer-dimer model ranged from 0.23 – 1.03  $\mu\text{M}$ . As described below, rigorous sedimentation equilibrium experiments were performed to substantiate these values.

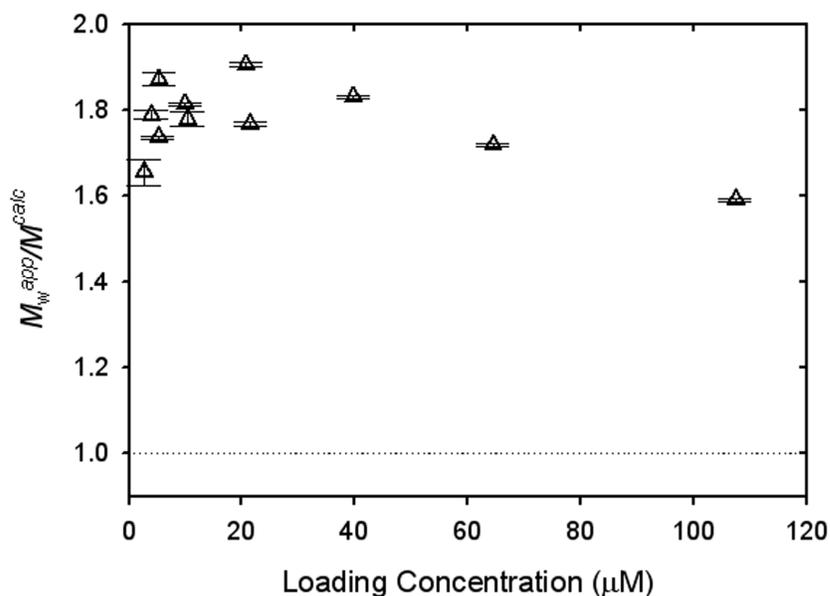


**Figure 4.4.** Analysis of N-HisPur $\beta$  sedimentation velocity by direct fitting of time-resolved concentration difference curves. **A**, Scan-pair subtractions ( $\Delta c$ ) of N-HisPur $\beta$  (13.4  $\mu$ M loading concentration shown) in buffer E at 50,000 rpm were fit to single molecular species with an average of absolute residuals of  $5.3488 \times 10^{-3}$  fringes (various grayscale symbols represent scan-pair subtraction data; lines represent single species fit).  $\Delta c$  curves represent scan-pair subtractions taken every 300 seconds. **B**, The same data set shown fit to a reversible monomer-dimer model with an average of absolute residuals of  $4.3248 \times 10^{-3}$ . The residuals of the fit are shown below the plots. The increase in the average of absolute residuals in **A** vs **B** are also seen in the increase in systematic deviations in the residuals.

The disparities between the experimentally derived and calculated  $R_H$  values, as well as the value of  $f/f_0$  for the dimeric species are suggestive that the shape of dimeric N-HisPur $\beta$  is markedly different from that of a condensed sphere. Molecular shape calculations predict that N-HisPur $\beta$  dimer is elliptical in solution, either prolate or oblate, with axial ratios of either  $a/b = 7.01$  or  $8.05$  for a prolate or oblate ellipsoid respectively. These calculations were made using a degree of hydration of  $0.4294 \text{ g H}_2\text{O/g N-HisPur}\beta$  which represents the predicted maximum degree of hydration based on amino acid composition of the protein using the Teller method (160) which is based on the assumption that all amino acids are maximally hydrated. Given this degree of hydration, the molecular dimensions would be  $23.57 \text{ nm} \times 3.36 \text{ nm}$  ( $2a \times 2b$ ) for the prolate prediction and  $12.90 \text{ nm} \times 1.60 \text{ nm}$  ( $2a \times 2b$ ) for the oblate ellipsoid. However, studies suggest that the actual degree of hydration of proteins is generally lower than maximal values, presumably due to folding and exclusion of water in the hydrophobic core of proteins. Kuntz showed that corrections for folding on the degree of hydration of proteins were on the order of 10% (90% of the maximum calculated value) (159), however values of  $0.30\text{-}0.35 \text{ g H}_2\text{O/g protein}$  are generally used in instances where the actual degree of hydration is not known. Unfortunately, accurate hydrodynamic modeling cannot be accomplished without prior knowledge of either the degree of hydration or the axial ratios of the hydrated molecule (109). If a 10% decrease in hydration of dimeric N-HisPur $\beta$  due to folding is assumed, the resulting hydration value is  $0.349 \text{ g H}_2\text{O/g protein}$ , and predicts axial ratios of  $7.29$  and  $8.43$  for a prolate and oblate ellipsoid, respectively. Molecular dimensions arising from these values would

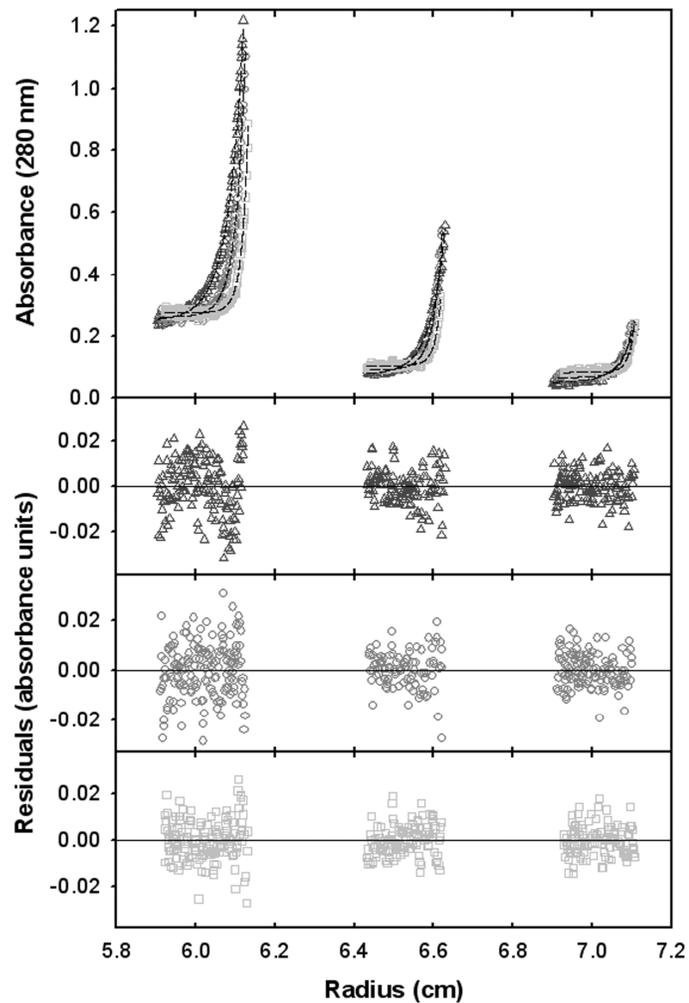
then be 23.91 x 3.28 nm ( $2a \times 2b$ ) for the prolate case, and 12.94 x 1.54 nm ( $2a \times 2b$ ) for the oblate prediction. These values indicate that the projected dimensions differ only modestly with the inclusion of this assumption-based correction.

***Sedimentation equilibrium analysis of recombinant Pur $\beta$***  – In order to validate the monomer-dimer self-association model for N-HisPur $\beta$ , and confirm the equilibrium constant that characterizes this association, sedimentation equilibrium studies were performed. A 10-fold range of concentrations (from 4 – 40  $\mu$ M) of size-fractionated N-HisPur $\beta$  in buffer E were sedimented at three different rotor speeds until equilibrium was attained (representative scans are shown in Figure 4.6). A careful and systematic analysis of experiments in which protein concentration, rotor speeds and buffer conditions were altered indicated that we were limited in the range of conditions, which would result in interpretable data. We were unable to run N-HisPur $\beta$  at concentrations lower than 4  $\mu$ M due to low radial absorbance distributions resulting in values below the signal to noise ratio of the instrument (data not shown). Attempts at lowering the reducing agent concentration to lower baseline absorbance, in an effort to collect low concentration absorbance datasets at 230 nm, resulted in formation of higher order oligomers as assessed by SEC (data not shown). We believe that these were artifactual oxidation products. Protein concentrations and rotor speeds were chosen such that non-ideal solution conditions were avoided. Further, buffer conditions were selected such that N-HisPur $\beta$  solubility was optimal. It has been found by our laboratory that preparations of N-HisPur $\beta$  are considerably less stable in buffered solutions containing less than 200 mM monovalent salt.



**Figure 4.5. Analysis of concentration-dependant N-HisPur $\beta$  solution non-ideality by sedimentation equilibrium.** Solutions of N-HisPur $\beta$  at various concentrations were analyzed by sedimentation equilibrium at rotor speeds of 22,000, 28,000, and 35,000 rpm. Radial concentration distributions for each loading concentration from all three rotor speeds were globally fit to a single-ideal species to ascertain the weight-average apparent molecular weight of the samples. These values (as a ratio of the apparent molecular weight to the known monomeric molecular weight,  $M_w^{App}/M^{Cal}$ ) are shown plotted versus the initial loading concentration. Data indicate that the apparent solution molecular weight approaches that of a dimer at loading concentrations less than 40  $\mu$ M, representative of self-association. Loading concentrations greater than 40  $\mu$ M show a decline in  $M_w^{App}/M^{Cal}$  values, suggestive of electrostatic repulsion non-ideality.

Our logic in choosing an appropriate set of experimental conditions to perform sedimentation equilibrium experiments is as follows. We evaluated the dependence of the ratio of apparent weight-average molecular weight to the calculated monomeric molecular weight ( $M_w^{App}/M^{Calc}$ ). We describe two possible outcomes. First, a systematic decrease in the  $M_w^{App}/M^{Calc}$  ratio as a function of loading concentration is suggestive of either hydrodynamic or electrostatic repulsion solution non-ideality. Both situations can be created by high loading concentrations or sedimentation at high rotor speeds and manifest as perturbations of the observed thermodynamic parameters of the system (163). Alternatively,  $M_w^{App}/M^{Calc}$  ratios greater than 1 and/or positive slopes in plots of this kind are indicative of electrostatic attraction. This latter case was seen at loading concentrations below 40  $\mu\text{M}$  and values of  $M_w^{App}/M^{Calc}$  approached 2, a result suggestive of dimerization (Figure 4.5). We found that the radial concentration distributions of N-HisPur $\beta$  at sedimentation equilibrium using loading concentrations of 1.40, 0.35, and 0.24 mg/ml (39.8, 9.94, and 3.98  $\mu\text{M}$ , respectively) and rotor speeds of 22,000, 28,000, and 35,000 rpm fit best to an ideal monomer-dimer model, as judged by fitting statistics and inspection of residual plots (Figure 4.6 and Table 4.1). The global fit of nine data sets to a reversible monomer- $N$ mer model returned a value of  $N = 2.04 \pm 0.03$  (holding monomer molecular weight = 35,168.6). This result confirmed the monomer-dimer equilibrium model that was indicated by the light scattering and sedimentation velocity data. Attempts at fitting the data to other models, including a single ideal monomer and various association models that incorporated noncompetent monomers or dimers did not result in improved fits (Table 4.1). Holding stoichiometry constant at  $N = 2$ , global fitting



**Figure 4.6. Sedimentation equilibrium analysis of N-HisPur $\beta$ .** Radial absorbance (280 nm) distributions of N-HisPur $\beta$  at three protein concentrations covering a 10-fold molar concentration range were obtained at rotor speeds of 22,000 rpm ( $\Delta$ ), 28,000 rpm ( $\circ$ ) and 35,000 rpm ( $\square$ ). The lines through the data represent a global fit of all 9 datasets to a reversible monomer to dimer equilibrium reaction. This model returned an equilibrium dissociation constant,  $K_d = 1.13 \pm 0.27 \mu\text{M}$ .

**Table 4.1 Sedimentation equilibrium data: Parameters from global analysis**

Model	Root Mean Square Deviations	Converged Parameters
Single, ideal	$8.81 \times 10^{-3}$	$M_w = 66,489 \pm 79$
Single, non-ideal <sup>a</sup>	$9.68 \times 10^{-3}$	$A_2 = 4.96 \times 10^{-7} \pm 4.01 \times 10^{-7} \text{ mlmol/g}^2$
Monomer-Nmer	$8.55 \times 10^{-3}$	$N = 2.04 \pm 0.03 \quad \ln K_a = 13.49 \pm 0.14$
Monomer-Dimer <sup>b</sup>	$8.53 \times 10^{-3}$	$\ln K_a = 13.69 \pm 0.10$
Monomer-Dimer + Incompetent Monomer <sup>c</sup>	$8.56 \times 10^{-3}$	$\ln K_a = 13.37 \pm 0.11 \quad \alpha = 0.00 \pm 0.00$
Monomer-Dimer + Incompetent Dimer <sup>c</sup>	$8.58 \times 10^{-3}$	$\ln K_a = 13.70 \pm 32.11 \quad \alpha = -0.003 \pm 22.79$

<sup>a</sup>  $M$  is constrained to dimeric molecular weight (70,337.2).

<sup>b</sup>  $N$  is constrained to a value of 2.0 (dimer).

<sup>c</sup>  $\alpha$  is equal to the weight fraction of the incompetent species given by  $\alpha = [M'] / ([M'] + [M] + 2[D'])$  in the case of incompetent monomer or by  $\alpha = [D'] / ([M] + 2[D] + 2[D'])$  for incompetent dimer.  $[M]$  is the molar concentration of monomer,  $[D]$  is the molar concentration of dimer,  $[M']$  is the molar concentration of incompetent monomer, and  $[D']$  is the molar concentration of incompetent dimer.

of the data yielded a dissociation constant,  $K_d = 1.13 \mu\text{M} \pm 0.27 \mu\text{M}$  (Table 4.1;  $\pm 1$  standard deviation), which is reasonably consistent with direct sedimentation velocity boundary fitting analyses.

The effects of salt concentration on N-HisPur $\beta$  self-association energetics were not explored in the present study. As mentioned previously, solubility of N-HisPur $\beta$  appears to be an issue when salt concentrations are reduced below 200 mM. Salt effects above 200 mM have yet to be rigorously tested.

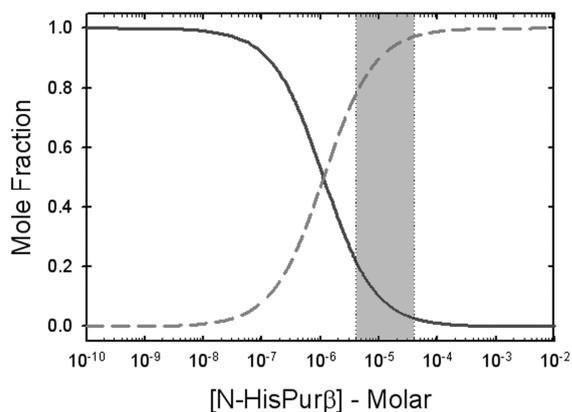
## **DISCUSSION**

In this study, we utilized hydrodynamic techniques to characterize the quaternary structure of N-HisPur $\beta$  in the absence of nucleic acids. The various hydrodynamic and thermodynamic techniques employed here suggest that N-HisPur $\beta$  exists in a monomer-dimer equilibrium characterized by a  $K_d = 1.13 \pm 0.27 \mu\text{M}$ . Preliminary sedimentation studies on N-HisPur $\beta$  showed signs of contaminating incompetent species (data not shown) that appeared to be artifacts, namely disulfide oxidation products, generated by concentrating techniques and consequently it was our feeling that utilization of a variety of robust hydrodynamic and thermodynamic techniques coupled with a preemptive size-fractionation step were warranted to distinguish between a possible mixture of thermodynamically tight dimer with incompetent monomer and a measurably reversible monomer-dimer equilibrium. Omission of the SEC step in preliminary sedimentation

equilibrium experiments resulted in returned  $K_d$  values for dimer dissociation ranging from 3.7 to 19.5  $\mu\text{M}$ , whereas inclusion of this step limited the range to 0.70 to 1.54  $\mu\text{M}$ . Moreover, fitting our data to models that included incompetent dimer did not improve the fitting statistics (Table 4.1). Hence, we believe that we have a fully reversible interacting system. The value of  $K_d$  reported here should be interpreted as an apparent upper limit to the actual value, since all measurements of  $K_d$  were made at concentrations greater than the apparent value of 1  $\mu\text{M}$  (Figure 4.7) due to low optical densities. It is likely that equilibrium measurements at concentrations lower than those utilized here will have to be made by a sensitive orthogonal technique, such as fluorescence polarization spectroscopy (128).

Hydrodynamic radius determinations either made directly by dynamic light scattering or by extrapolation from sedimentation data were in sound agreement. Molecular shape calculations were consistent with an asymmetric shape of dimeric N-HisPur $\beta$  in solution. This finding is not surprising based on the fact that Pur $\beta$  is composed of 22.2% glycine (143), which likely contributes to a lack of secondary structural elements and a somewhat disordered tertiary structure, and may explain difficulties in obtaining higher resolution structural information. A further interesting finding is that the  $f/f_0$  ratio determined for the monomeric species is greater than that of the dimer, which suggests that dimerization results in partial condensation of the overall structure. The implications of this finding are discussed below.

We also addressed the concern that the N-terminal hexahistidine tag present on the recombinant protein might adversely affect the functional activity of Pur $\beta$  by comparing



**Figure 4.7. Species plot of N-HisPurβ.** This simulation was performed assuming a monomer-dimer equilibrium association reaction with a  $K_d = 1.13 \pm 0.27 \mu\text{M}$ , showing the relative mole fractions of monomeric (—) and dimeric (---) species of N-HisPurβ at various concentrations. The concentration at which the two plots cross is indicative of the  $K_d$ . The region demarcated by dotted lines is the concentration range of N-HisPurβ used to determine the equilibrium dissociation constant.

the ssDNA-binding activity of native and recombinant N-HisPur $\beta$  using a quantitative ELISA approach (145). The ssDNA-binding activity of N-HisPur $\beta$  was indistinguishable compared to Pur $\beta$  derived from either mouse embryo fibroblasts or vascular smooth muscle cells (data not shown). Also, non-specific metal ion-mediated dimerization of hexahistidine can be eliminated as a complicating factor since all measurements were made in the presence of 0.5 mM EDTA.

Our interest in Pur proteins stems from their putative involvement in repressing the transcription and translation of genes that mark the phenotypic status of myofibroblasts, vascular smooth muscle cells, and cardiomyocytes (28, 104, 142, 145, 148). Evidence for formation of transient ssDNA structures within the asymmetric purine/pyrimidine tract of the 5'-SM $\alpha$ A promoter (9) has also fueled investigation into the mechanism of DNA-binding by Pur proteins. It has been suggested that Pur proteins bind *PUR*-elements as either hetero- or homodimers (141), although the stoichiometry, mechanism, and thermodynamics of nucleoprotein assembly have not yet been determined by rigorous physical investigation. In the case of Pur $\alpha$ , self-association has been reported to be mediated by an RNA molecule of unknown nature (91). This was an intriguing finding, as it suggested that RNA may influence the ability of Pur $\alpha$  to stably self-associate and to bind to ssDNA. Curiously, despite sharing 70% amino acid sequence identity, we have shown in this study that Pur $\beta$  dimerizes in the absence of any nucleic acid. It is possible that this disparity is due to a distinct functional difference between Pur $\alpha$  and Pur $\beta$  as implied by previous gain-of-function studies (145, 148), or could be reflective of different experimental approaches (e.g. pull-down assay versus sedimentation

equilibrium analysis). Moreover, the apparent binding affinities of Pur $\alpha$  and Pur $\beta$  for different *PUR*-elements are reported to be on the order of 1 nM (145, 310), whereas the affinity of Pur $\beta$  self-association is three orders of magnitude weaker. As such, Pur $\beta$  at concentrations below 100 nM is predicted to be largely monomeric as depicted in the molecular species plot in Figure 4.7. This suggests that other factors (such as an RNA ligand in the specific case of Pur $\alpha$ ) may be required to help facilitate dimerization in a cellular milieu if the concentration of protein is limiting. On the other hand, it remains to be resolved as to what extent Pur $\alpha$  and Pur $\beta$  share similar intrinsic homodimerization ability in the absence of nucleic acid and whether or not heterodimeric complexes associate with enhanced or reduced affinity relative to their homodimeric counterparts.

In conclusion, we report that recombinant Pur $\beta$  participates in a monomer-dimer equilibrium governed by an apparent upper limit dissociation constant of  $\sim 1$   $\mu$ M. The ability to dimerize in a reversible fashion may represent an important regulatory mechanism, allowing mass-action governed self-association to play either a positive or negative role in nucleic acid-binding. As indicated by frictional coefficient ratios, dimerization may result in structural reorganization of N-HisPur $\beta$  that may permit nucleic acid recognition and binding. Self-association of transcription factors as a prerequisite to DNA-binding is not unprecedented. For example, the STAT proteins require phosphorylation-dependent dimerization prior to nuclear localization and binding to cytokine responsive gene promoters (112). Similarly, intracellular estrogen receptor DNA-binding activity is apparently dependent upon ligand-mediated dimerization (80, 282, 295). However, in the case of Pur $\beta$ , dimerization does not appear to require post-

translational modification or ligand binding *per se* as self-association is governed by protein concentration *in vitro*. However, this fact does not rule out the prospect that post-translational modifications may alter the affinity of homodimerization *in vivo*. Although the absolute intracellular concentration is not known, it has been reported that levels of Pur $\beta$  increase dramatically in cardiovascular cell types undergoing phenotypic changes (9). This entices speculation that repressive effects of Pur $\beta$  may depend upon its expression and accumulation within the nucleus to levels that drive self-association and permit ssDNA-binding. Future studies will focus on determining the affinity and stoichiometry of relevant Pur nucleoprotein complexes with the use of quantitative biophysical approaches such as those described herein.

**CHAPTER V. THERMODYNAMIC ANALYSIS OF SEQUENCE-SPECIFIC SINGLE-STRANDED DNA-BINDING BY RECOMBINANT PUR $\beta$  REVEALS A COOPERATIVE MECHANISM OF NUCLEOPROTEIN ASSEMBLY**

*The following is a description of original and unpublished work.*

**INTRODUCTION**

Classical models of transcription regulation imply that assembly and processivity of transcriptional machinery at gene promoter regions is coordinated by the pre-assembly of *trans*-acting factors at specific regions of the promoter, and that this pre-assembly is dictated by a combination of the nucleotide sequence of the promoter and environmental cues. By-and-large, these models propose that binding of specific DNA sequences by transcription factors involves the recognition of particular arrangements of chemical functional groups along the major or minor groove of duplex DNA that coincide with compatible amino acid chemical groups arranged on the surface of the protein. However, the detection of large structural re-arrangements, and altered nuclease sensitivities in promoter regions of genes coincident with apparent transcription factor binding has suggested that chromatin structure and/or DNA conformation also plays a prominent role in regulating gene transcription (for review see (277)). Furthermore, the detection of stable, single-stranded DNA structures, and the identification of interacting sequence-specific ssDNA-binding transcription factors that are critical for regulated transcription of

numerous genes implies that many eukaryotic genes do not conform to classical models of gene regulation, and likely utilize more complex mechanisms that incorporate combinations of changes in chromatin structure, and the binding of protein factors to double-stranded, and transiently formed single-stranded and non-B-DNA structures. Collectively these findings question the view that genomic promoter DNA acts as a static lattice for the coordinated assembly of transcription factors to bind and enforce transcriptionally active or inactive states, and instead it appears likely that, for some genes, the same sequence of promoter DNA is able to bind select sets of transcription factors dependent upon the dynamic interconversion of double-stranded and single-stranded conformations.

Single-stranded promoter character, as well as reliance on sequence-specific ssDNA-binding proteins for regulated transcription is fairly common, and has been observed for such notable genes as *c-myc* (10, 11, 191, 192, 203, 280, 288), major histocompatibility complex class II (69, 183), platelet-derived growth factor A-chain (298, 301, 322), apolipoprotein B (very low density lipoprotein II (263), low density lipoprotein receptor (221), *fas* cell surface receptor (263), androgen receptor (39, 101), epidermal growth factor receptor (130, 312), procollagen  $\alpha 1$  (68), smooth muscle  $\alpha$ -actin (11, 48, 275), and  $\beta$ -myosin heavy chain (129). Common to many of the promoters of these genes are the presence of asymmetrically distributed nucleotide tracts that, in other studies, have been shown to be prone to forming non-B-DNA structures. For instance, polypurine/polypyrimidine, or alternating purine/pyrimidine stretches have been shown to not only adopt Z-DNA (204), and H-DNA (235) conformations, but to cause loss of base-pairing in the neighboring

transition residues (105). Often, single-stranded character in these regions can be detected through the use of ssDNA-specific reagents (9, 130, 192, 229, 235, 274, 285, 298, 301), further suggesting that these conformations are stable entities. The correlation of non-B-DNA-prone sequences and detectable single-stranded DNA in regions of activated transcription suggests that alternate DNA conformations are an important component of transcriptional regulation.

Despite mounting evidence supporting their existence and utility, the manner in which sequence-specific ssDNA-binding proteins are able to function in a predominantly double-stranded genome remains largely unknown. It has been postulated (179), and shown that, in certain instances, negative supercoiling generated by transcriptional machinery-associated helicase activity promotes localized Z-DNA formation or unwinding of upstream sequences harboring atypical nucleotide stretches, as mentioned above. This characteristic coincides with regions of transcriptional activity, but is not universal, as it appears limited to a certain subset of genes (33, 161). A transcription-induced topological strain mechanism has been attributed to the case of the far upstream element (FUSE) of the *c-myc* promoter in which upstream nucleotide unpairing was noted for both supercoiled and linearized plasmid templates in reconstituted transcription assays (153), and *in vivo* (154). Structural interconversion, in turn, permits interaction with several sequence-specific ssDNA-binding proteins including FUSE-binding protein (FBP, (75)), and FBP-interacting repressor (FIR, (178)). Thus, in the case of *c-myc* FUSE-protein interactions, sequence-specific ssDNA-binding by FBP and FIR is apparently opportunistic, relying on pre-formation of single-stranded DNA structures for

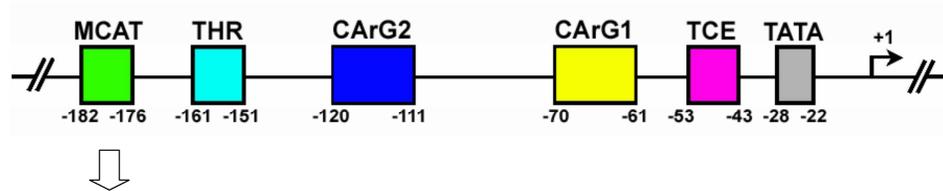
assembly, despite reports that FBP is capable of helix destabilization in linear dsDNA constructs containing FUSE sequences (8). Conceptually, this mechanism may not apply to all promoter regions, especially those in which maintenance of a transcriptionally repressed state requires ssDNA-interacting transcription factors. It is envisioned that in these cases, single-stranded conformations are created and stabilized by binding of protein factors by virtue of their inherent energetics of nucleoprotein complex assembly; i.e., ssDNA nucleoprotein complex formation is thermodynamically competitive to non-B-DNA duplex formation. Transcriptional regulation by sequence-specific ssDNA-binding proteins at promoters containing sequences prone to Z-DNA conformations, or other high-energy states with reduced melting temperatures, suggests that helix destabilization may be a mechanism employed by many ssDNA-dependent transcription factors to achieve promoter occupation and transcriptional regulation. This notion is supported by findings that have shown that many sequence-specific ssDNA-binding proteins bind to supercoiled plasmid DNA and not to linearized plasmids (8, 153, 154, 192, 288, 310).

Pur $\beta$  is a sequence specific ssDNA/RNA- binding transcription/translation factor that, along with Y-box protein MSY1 (mouse YB-1) and Pur protein family co-member Pur $\alpha$ , acts as a repressor of smooth muscle  $\alpha$ -actin (SM $\alpha$ A) gene expression. The protein encoded by this gene is an essential cytoskeletal contractile apparatus protein required for regulation of vascular tone and cellular migration (79, 210, 231, 257). Therefore SM $\alpha$ A and has been identified as being critical for maintenance of the differentiated contractile phenotype of vascular smooth muscle cells (VSMCs) and myofibroblasts (166, 226, 243,

271). The repressive activities of Pur $\beta$ , Pur $\alpha$  and MSY1 on the SM $\alpha$ A promoter have implicated them in facilitating smooth muscle cell dedifferentiation and vascular wall remodeling, processes common to numerous vascular pathologies including atherosclerosis (175). It has been demonstrated that each transcription factor exerts its repressive effect by binding to the proximal muscle-specific CAT (MCAT) enhancer element of the SM $\alpha$ A promoter that has been shown to possess a high level of purine/pyrimidine base asymmetry (48, 82, 143, 275), and to transiently adopt single-stranded conformations in response to dedifferentiating stimuli (9). Biochemical studies have confirmed that Pur $\alpha$  and Pur $\beta$  preferentially bind the purine-rich strand, whereas MSY1 shows specific affinity for the opposing pyrimidine-rich strand, and furthermore, a network of protein-protein interactions between each transcription factor exists (141, 148) as well as nucleic acid independent self-association, in the case of Pur $\beta$  [190]. Current models for the achievement of transcriptional repression by these proteins propose that binding of Pur $\alpha$  and/or Pur $\beta$  to the purine-rich strand and MSY1 to the pyrimidine-rich strand destabilizes B-DNA helical structures and maintains the enhancer region in a single-stranded conformation, thus disrupting dsDNA-dependent transcription enhancer 1 (TEF-1) binding and function (28, 275).

Based on our recent work which showed that recombinant Pur $\beta$  self-dimerizes with a dissociation constant of 1.13  $\mu$ M (222), and previous reports describing marked increases in Pur $\beta$  protein levels in VSMCs undergoing phenotypic changes consistent with SM $\alpha$ A repression (104, 271), it was hypothesized that Pur $\beta$  dimerization may be a prerequisite to ssDNA-binding and serve as a regulatory mechanism for Pur $\beta$  function. In the present

study, our goal was to delineate the mechanism of sequence recognition and ssDNA binding by Pur $\beta$ , which has been shown to be the more dominant Pur protein repressor in the context of SM $\alpha$ A gene expression (19, 145), to a sequence representative of the proximal SM $\alpha$ A MCAT enhancer element. We have used a thermodynamic approach to examine nucleoprotein stoichiometry, mechanism, and binding energetics to gauge the plausibility of helix destabilization as a possible mechanism by which Pur $\beta$  binds specific ssDNA sequences in a generally double-stranded environment. Our results indicate that binding of Pur $\beta$  to the purine-rich strand (promoter element; PE) within the SM $\alpha$ A MCAT enhancer involves facilitated cooperative assembly of Pur $\beta$  monomers (in the 300 pM range) to form higher order nucleoprotein structures with stoichiometries of 2:1 (Pur $\beta$ :PE). The resolved energetics of binding of Pur $\beta$  to the purine rich strand of SM $\alpha$ A enhancer, by way of quantitative ssDNA footprinting, suggests that maintenance of a single-stranded state within this region might require auxiliary activities; either the involvement of Pur $\alpha$ , the concurrent cooperative assembly of MSY1 on the pyrimidine-rich strand, the reduction of local melting temperatures by induction of topological stress, or a combination thereof.



PE32-F: (-195) – GGGAGCAGAACAGAGGAATGCAGTGGAAGAGA – (-164)

**Figure 5.1. Schematic of *cis*-regulatory elements of the SM $\alpha$ A promoter.** Plasticity of SM $\alpha$ A expression is governed by multiple regulatory elements in the region of the promoter 5' to the transcriptional start site. Non-canonical CArG elements mediate potent serum-stimulated activation of gene expression, whereas the MCAT, THR, and TCE confer both activation and repression depending on structural configuration and corresponding transcription factor occupancy. Detectable structural rearrangements and Pur $\alpha$ , Pur $\beta$ , and MSY1 repressor occupancy at the cryptic MCAT enhancer element make this region a focus of study herein. Shown is the purine-rich strand, as represented by oligonucleotide PE32-F, which has been shown to bind Pur $\alpha$  and Pur $\beta$  with high affinity. Numbers indicate nucleotide positions relative to the transcriptional start site. Oligonucleotide SMP382-F, used in footprinting studies, corresponds to nucleotides -323 to +59 encompassing the *cis*-regulatory elements comprising this region.

## MATERIALS AND METHODS

*Chemicals, protein reagents, and oligonucleotide probes.* All chemicals used in this study were of reagent grade or better. Recombinant Pur $\alpha$  and Pur $\beta$  were expressed as amino-terminally labeled hexahistidine tag fusion proteins (referred to in this report as N-HisPur $\alpha$ , and N-HisPur $\beta$ , respectively), purified from *E. coli* expression cultures, and quantified as described in previous publications (149, 222). AKR-2B nuclear extracts were obtained from cell monolayers cultured under exponential growth conditions as described previously (143). Protein concentrations of nuclear extracts were assessed by bicinchonic acid assay (Sigma-Aldrich) using high purity BSA (Boehringer-Mannheim) as a protein standard. Preparation and validation of epitope-specific rabbit polyclonal antibodies against mouse Pur $\alpha$  (anti-Pur $\alpha$  291) and mouse Pur $\beta$  (anti-Pur $\beta$  302) has been described previously (141). Synthetic oligonucleotides were purchased from Sigma-Genosys. Yeast tRNA was purchased from Sigma. Enzyme reagents were purchased from commercial sources; T4 polynucleotide kinase was purchased from New England Biolabs, DNase I, and Accuprime Supermix<sup>TM</sup> (*Taq* polymerase) were purchased from Invitrogen Corp. Fresh stocks of [ $\gamma$ -<sup>32</sup>P]ATP (6000 Ci/mmol) were obtained from Perkin-Elmer. Sequenase dideoxy-NTP sequencing kit was purchased from Unites States Biochemical Corp. SYBR-Gold nucleic acid stain was purchased from Invitrogen.

*Preparation of ssDNA for quantitative DNase I footprinting.* Methods for the purification of ssDNA were based on protocols originally developed and reported by

others (121, 139). A plasmid vector containing the full-length mouse SM $\alpha$ A promoter, termed VSMP8 (296) was used as a template for PCR-based amplification of a 382 base fragment comprising the forward strand of the proximal SM $\alpha$ A promoter, termed SMP382-F (bases -323 to +59 relative to the transcriptional start site). To facilitate both amplification and purification of the forward single strand consisting of the purine-rich sequence of the proximal MCAT enhancer element, PCR primers were designed such that the reverse strand primer was 5'-biotinylated (SMP8p1122s-R-5btn; 5'-biotin-GGCTACTTACCCTGACAGCGACT-3'), whereas the forward strand primer was unmodified (SMP8p741s-F; 5'-TTCTGAGGAATGTGCAAACCGTG-3'). PCR amplification of the 382 bp fragment from 1 ng/ $\mu$ L of VSMP8 template was carried out using Accuprime Supermix<sup>TM</sup> reagent according to the manufacturer's instructions. The incorporation of the biotinylated reverse strand primer allowed for the isolation of useable quantities of SMP382-F through implementation of biotin-streptavidin affinity based techniques. Briefly, double-stranded PCR product (typically 500  $\mu$ L) was applied to 1 mg of Streptavidin MagneSphere paramagnetic particle resin (Promega Corp.), pre-equilibrated in 20 mM Tris pH 8.8, and incubated at 4°C for 16 h. The PCR-product-resin complex was washed twice with 20 mM Tris pH 8.8. Non-biotinylated SMP382-F was eluted by incubating the resin complex in 0.2 N NaOH for 5 min at 20 $\pm$ 1°C. The eluant solution was neutralized by the addition of 1/10 volume of 5 M ammonium acetate. SMP382-F was precipitated by the addition of 2 volumes of isopropanol at -20°C, washed with 70% ethanol, and redissolved in nuclease-free water. Concentration of SMP382-F was approximated by absorbance measurement at 260 nm, assuming an

extinction coefficient of  $3,630,200 \text{ M}^{-1}\text{cm}^{-1}$  as calculated using web-based software (<http://biophysics.idtdna.com>) based on methods described by Tataurov and colleagues (208, 284).

***End-labeling of oligonucleotides with [ $\gamma$ - $^{32}\text{P}$ ]ATP.*** Single-stranded DNA oligonucleotides for use in electrophoretic mobility shift assays (PE32-F), DNase I footprinting assays (SMP382-F), and primer-labeled dideoxy sequencing reactions (SMP8sp741s-F) were enzymatically labeled on their 5' termini with [ $\gamma$ - $^{32}\text{P}$ ]ATP using T4 polynucleotide kinase activity as directed by the manufacturer. Reactions were performed at  $37^\circ\text{C}$  for 90 min. Upon completion, the enzyme was heat inactivated by incubating reaction mixtures at  $70^\circ\text{C}$  for 10 min. Unincorporated [ $\gamma$ - $^{32}\text{P}$ ]ATP was removed by buffer exchange over G-25 Microspin columns (Pharmacia). Extent of labeling was assessed by scintillation counting of purified probes using a Perkin-Elmer Tri-Carb® scintillation counter. For purposes of clarity, oligonucleotides carrying a 5'- $^{32}\text{PO}_4$  radiolabel will be marked with an "\*" (for example, PE32-F\* denotes 5'- $^{32}\text{PO}_4$ -labeled PE32-F).

***Qualitative electrophoretic mobility (super)shift assay.*** Binding reactions for qualitative electrophoretic mobility shift assays (EMSA) were carried out in EMSA binding buffer (50 mM Tris•HCl pH 7.5, 100 mM KCl, 5% (v/v) glycerol, 0.5 mM dithiothreitol, 2  $\mu\text{g}/\text{mL}$  dT<sub>32</sub>, 50  $\mu\text{g}/\text{mL}$  BSA) plus 0.5 mM ethylenediaminetetraacetic acid (EDTA) to prevent nuclease digestion of probe by nuclear extracts. Recombinant proteins and nuclear extracts were diluted to the indicated concentrations prior to addition of PE32-F\* (5'-GGGAGCAGAACAGAGGAATGCAGTGGGAAGAGA-3') to a final

concentration of 1 nM. Binding reactions were incubated for 16 h at 4°C to allow formation of nucleoprotein complexes prior to the addition of super-shifting antibodies, anti-Pur $\alpha$  291 and anti-Pur $\beta$  302, to the indicated concentrations. Final volumes of reactions were 40  $\mu$ L. Reactions were incubated at 4°C for an additional 2 h. Complexes were resolved electrophoretically by loading 10  $\mu$ L of binding reactions (typically 5,000 – 10,000 cpm) on a 1.5 mm thick 10% polyacrylamide gel (75:1 acrylamide:bisacrylamide) cast and run in TBE buffer (25 mM Tris, 25 mM boric acid, 0.5 mM EDTA). Gels were pre-run at 4 Watts per gel for 1 h prior to loading and running for 75 min at 4 Watts at room temperature. Upon completion, gels were disassembled and dried on Whatman filter paper in a Bio-Rad slab gel dryer for 45 min at 65°C. Phosphorimaging of the dried gels was performed by exposing to phosphor storage screens (Molecular Dynamics) for 24-48 h, prior to developing with a Bio-Rad PersonalFX<sup>TM</sup> phosphorimager. In certain cases, dried gels were visualized by exposure to X-Omat film for 6-16 h at -80°C prior to development.

An agarose gel electrophoresis-based assay was used to gauge the extent of binding of recombinant proteins to plasmid DNA, and is based on experiments described previously (310). Briefly, VSMP8 (7.9 kb) and parent plasmid vector pBLCAT3 (4.3 kb) were purified by double cesium chloride gradient centrifugation, and digested at a final DNA concentration of 50 ng/ $\mu$ L, with *Hind*III (Roche) according to the manufacturers instruction. Enzyme was inactivated by incubation at 65°C for 10 min. Digested and supercoiled plasmids were mixed at an original plasmid concentration of 1 nM (2.84 and 5.21 ng/mL for pBLCAT3 and VSMP8, respectively) with 25, 50, 100, or 200 nM

recombinant Pur protein in buffer (50 mM Tris•HCl pH 7.5, 100 mM KCl, 5% (v/v) glycerol, 0.5 mM dithiothreitol, 2 µg/mL dT<sub>32</sub>, 50 µg/mL BSA) for 16 h at 20±1°C. Samples (30 ng total DNA) were electrophoresed on a 0.7% SeaKem LE agarose (Cambrex) gel cast and run in 0.5X traditional TBE Buffer (44.5 mM Tris-Borate pH 8.3, 1 mM EDTA) at 5 V/cm for approximately 2 h at 20±1°C. Gels were stained with SYBR-Gold stain diluted 1:10,000 in 0.5X traditional TBE for 30 min prior to image capture with a Bio-Rad GelDoc XR imaging system.

***Fluorescence anisotropy measurements.*** Fluoresceinated (3') oligonucleotide probe PE32-F (termed PE32-F-3FLC, 5'-GGGAGCAGAACAGAGGAATGCAGTGGGAAGAGA-Fluorescein-3') was titrated with N-HisPurβ in buffer consisting of 50 mM HEPES pH 7.8, 300 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 2 mM β-mercaptoethanol, and allowed to equilibrate for approximately 16 h at 20±1 °C in the absence of light. Probe concentrations were varied. Steady-state fluorescence anisotropy measurements were made on a Quantamaster-6 spectrofluorometer (Photon Technologies International, South Brunswick, NJ) equipped with a 75 W xenon arc lamp excitation source, excitation and emission monochromators, and automatic excitation and emission polarizers in a T-format. Slit-widths were varied to maintain constant observed emission intensities between samples with different probe concentrations. Parallel and perpendicular emission intensities were collected with horizontally polarized exciting light in order to first calculate the instrument correction factor,  $G$ , given by  $G = I_{HV}/I_{HH}$  where  $I_{HV}$  and  $I_{HH}$  are the intensities measured through the vertical and horizontal polarizers when excited with horizontally polarized light.

Observed anisotropy values,  $r_{obs}$ , were calculated by the expression  $r_{obs} = (I_{VH} - GI_{VH})/(I_{VV} + 2GI_{VH})$ , where  $I_{VV}$  and  $I_{VH}$  are the intensities measured through the vertical and horizontal polarizers when excited with vertically polarized light. Changes in emission intensity as a function of protein concentration were not observed.

For the completely cooperative reaction,  $nP + D \rightleftharpoons P_nD$ , where  $P$  is the protein ligand,  $D$  is the DNA lattice and  $n$  is the stoichiometry of the terminal complex; the constant  $K_r$  can be defined as

$$K_r = \frac{[P][D]}{[P_nD]} \quad (\text{Equation 5.1})$$

Note that  $K_r$  is not a true equilibrium constant in cases where  $n$  is greater than unity. Substituting terms describing the laws of mass action,  $K_r$  can be written as

$$K_r = \frac{([P_t] - n[P_nD])([D_t] - [P_nD])}{[P_nD]} \quad (\text{Equation 5.2})$$

Where  $[P_t]$  and  $[D_t]$  represent the total concentrations of protein and DNA, respectively. Solving for  $[P_nD]$  gives the quadratic

$$[P_nD] = \frac{([P_t] + n[D_t] + K_r) \pm (([P_t] + n[D_t] + K_r)^2 - 4n[P_t][D_t])^{1/2}}{2n} \quad (\text{Equation 5.3})$$

It has been described elsewhere (180, 292), that for a mixture of molecular species with different anisotropies (but with the same fluorescence intensities), the measurable anisotropy of the solution is the sum of the mole fractions of the individual species multiplied by their inherent anisotropic values

$$r_{obs} = f_1r_1 + f_2r_2 + \dots + f_i r_i \quad (\text{Equation 5.4})$$

where  $f_i$  and  $r_i$  are the mole fraction (and quantum yield) and the anisotropy of the  $i^{th}$  species, respectively. Equation 5.4 can be rearranged and expressed in terms of fraction bound,  $f_B$

$$f_B = \frac{[P_n D]}{[D_i]} = \frac{(r_{obs} - r_f)}{(r_b - r_f)} \quad (\text{Equation 5.5})$$

Substituting equation 5.5 into equation 5.4 and applying terms that accommodate nonspecific binding gives:

$$r_{obs} = b(r_b - r_f) \frac{1}{[D_i]} \left\{ \frac{(R_{P/D}[D_i] + n[D_i] + K_r) \pm ((R_{P/D}[D_i] + n[D_i] + K_r)^2 - 4nR_{P/D}[D_i]^2)^{1/2}}{2n} \right\} + r_f + mR_{P/D}[D_i] \quad (\text{Equation 5.6})$$

where  $R_{P/D}$  is the molar ratio,  $[P_i]/[D_i]$ , and  $b$  and  $m$  are non-specificity terms that permit both the fluctuation of the equivalence saturation point and the slope of the plateau region due to non-specific binding, respectively. Obtained anisotropy values,  $r_{obs}$ , for titrations

of N-HisPur $\beta$  against fixed concentrations of PE32-F-3FLC were plotted against known values of  $R_{PD}$ , and fit to equation 5.6 to obtain best-fit values of  $n$ . Fitting was performed with Prism 5 software (GraphPad Software, Inc., San Diego, CA).

***Quantitative electrophoretic mobility shift assay.*** Binding reactions for quantitative EMSAs were carried out in EMSA binding buffer and allowed to equilibrate for 16 h at  $20 \pm 1$  °C. Methods of quantitative EMSA for the estimation of binding stoichiometry are based on protocols published previously (84, 85, 223, 224) termed the serial-dilution method. Briefly, solutions of N-HisPur $\beta$  (5 nM) and PE32-F\* (1 nM) were serially-diluted 1.1:1 fold in EMSA binding buffer to obtain a complete dilution series. The serial-dilution method relies on the redistribution of nucleoprotein complex concentration, as a result of serial-dilution, based on the ideal law of chemical equilibrium without changing the ratio of total protein concentration to total DNA concentration. This redistribution is then monitored by native gel electrophoresis as described above for qualitative EMSA, with the only exception being that gels are run for 45 min to prevent streaking due to complex dissociation. For the general cooperative binding mechanism  $nP + D \rightleftharpoons P_nD$ , the macroscopic association constant  $K_a$  is defined as:

$$K_a = [P_nD]/[P_{free}]^n[D] \quad (\text{Equation 5.7})$$

In these expressions  $P$  represents N-HisPur $\beta$ ,  $D$  represents the probe PE32-F\*, and  $n$  represents the stoichiometry of the nucleoprotein complex. Rearrangement of the definition of  $K_a$  gives:

$$\ln([P_nD]/[D]) = n\ln[P_{free}] + \ln K_a \quad (\text{Equation 5.8})$$

Densitometry of phosphorimages was carried out using ImageQuant software (Molecular Dynamics) to determine  $[P_nD]$  and  $[D]$ , whereas  $[P_{free}]$  was estimated from the relationship:

$$[P_{free}] = [P_{input}] - n[P_nD] \quad (\text{Equation 5.9})$$

In which case  $[P_{input}]$  is the input concentration (a known quantity) and  $n$  is an integer estimate of the stoichiometry. Measured values of  $[P_nD]$ ,  $[D]$ , and estimated values of  $[P_{free}]$  were used to calculate a value of  $n$  (slope from regression of the linear plot of  $\ln([P_nD]/[D])$  versus  $\ln[P_{free}]$ ). The estimated integer estimate value of  $n$  was iteratively changed until the integer estimate and returned regression value of  $n$  converged. The value of  $\ln K_a$  was also estimated from the intercept of the converged linear plot. Linear regression was performed using Prism 5 software.

Direct titration methods were used to estimate macroscopic binding affinities and for the detection of cooperativity of N-HisPur $\beta$  binding to PE32-F\*. Briefly, 2X solutions of N-HisPur $\beta$  were prepared in EMSA binding buffer by 2/3 fold serial dilution of a 20 nM

master stock solution. An equal volume of 2X PE32-F\* in EMSA binding buffer was added to N-HisPurβ solutions so that final concentrations of each were at 1X with PE32-F\* concentration being constant across all binding reactions. In order to maintain validity of the assumption  $[P_{free}] \approx [P_{total}]$ , PE32-F\* concentration was kept at 25 pM for EMSA used for rigorous thermodynamic investigations. Free and bound probe was separated by native gel electrophoresis as described above, with 5-10 μL of reaction mixture (usually 700 – 2000 cpm) loaded in each lane. Dried gels were exposed 72-96 h to phosphor storage screens. Quantification of binding was carried out by measuring the optical density of each electrophoretic species (band) using ImageQuant software (Molecular Dynamics). Species density values were then used to determine fractional species saturation,  $\Theta_i$ , where  $i$  is equal to the number of protein ligands bound to the ssDNA lattice ( $i = 0, 1, \text{ or } 2$  for a system with a finite stoichiometry of 2:1) by applying the following expression:

$$\Theta_i = I_i / \sum I_i \quad (\text{Equation 5.10})$$

where  $I_i$  refers to the integrated optical density of the  $i$ th species, and the summation is over all of the bands in a particular lane of the gel. Estimation of binding parameters was carried out by a statistical mechanical method described previously by Brenowitz and Senear (250), and makes use of the two-site model depicted in Table 5.1. The probability of a particular species existing can be expressed by

$$f_s = \frac{e^{(-\Delta G_s)/RT} [P_{free}]^n}{\sum_{s=1}^n e^{(-\Delta G_s)/RT} [P_{free}]^n} \quad (\text{Equation 5.11})$$

where  $\Delta G_s$  is the relative free energy change observed upon formation of the  $s$  configuration compared to the reference state,  $R$  is the gas constant,  $T$  is the absolute temperature,  $n$  is the number of N-HisPur $\beta$  monomers bound in the  $s$  configuration, and the summation is over all species. Microscopic association constants are related to their corresponding microscopic free energies through the familiar relationship,  $\Delta G_i = -RT \ln k_i$ . Thus for a two-site system, the fractional saturation,  $\Theta_i$ , of a given species can be expressed as the sum of the probabilities of configurations in which the adopted stoichiometry equals  $i$ , giving

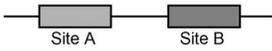
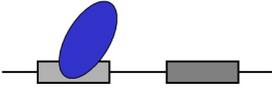
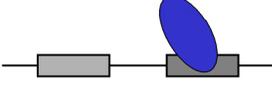
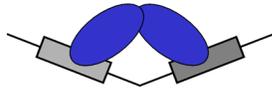
$$\Theta_0 = \frac{1}{1 + (k_1 + k_2)[P_{free}] + k_1 k_2 k_c [P_{free}]^2} \quad (\text{Equation 5.12})$$

$$\Theta_1 = \frac{(k_1 + k_2)[P_{free}]}{1 + (k_1 + k_2)[P_{free}] + k_1 k_2 k_c [P_{free}]^2} \quad (\text{Equation 5.13})$$

$$\Theta_2 = \frac{(k_1 k_2 k_c)[P_{free}]^2}{1 + (k_1 + k_2)[P_{free}] + k_1 k_2 k_c [P_{free}]^2} \quad (\text{Equation 5.14})$$

As indicated by Brenowitz and Senear, the microscopic constants  $k_1$ ,  $k_2$ , and  $k_c$  only appear in two combinations in all three equations and can be replaced by substituting macroscopic constants  $K_1 = (k_1 + k_2)$ , and  $K_2 = k_1 k_2 k_c$ . Thus, global fitting of species-specific isotherms resolves only the macroscopic constants  $K_1$  and  $K_2$ , from which microscopic constants can only be extracted in instances when cooperativity is

**Table 5.1. Microscopic configurations and corresponding interaction free energies for a two site system.**

Two-site Binding			Free Energy Contribution <sup>2</sup>	Microscopic Constant
Species Representation <sup>1</sup>				
$s = 1$		$i, n = 0$	Reference state	-
$s = 2$		$i, n = 1$	$\Delta G_1$	$k_1$
$s = 3$		$i, n = 1$	$\Delta G_2$	$k_2$
$s = 4$		$i, n = 2$	$\Delta G_1 + \Delta G_2 + \Delta G_c$	$k_1 k_2 k_c$

<sup>1</sup>Species diagrams represent possible microscopic configurations and do not necessarily reflect structural perturbations as a result of ligation, only that interaction between sites occurs in the case of cooperativity. <sup>2</sup>Changes in intrinsic free energy for each species configuration are relative to the unligated reference state, and are related to microscopic constants through the relationship  $\Delta G_i = -RT \ln k_i$ .

nonexistent or binding sites are identical. Fractional saturation curves were also fit to the well known phenomenological Hill equation (305) to gauge the extent of cooperative interactivity present in nucleoprotein complex formation. For this purpose complete fractional saturation was calculated by:

$$\bar{Y} = 1 - ([D_{free}]/[D_{input}]) \quad (\text{Equation 5.15})$$

and fit to the familiar equation:

$$\bar{Y} = [P_{free}]^{\alpha_H} / (K_d^{\alpha_H} + [P_{free}]^{\alpha_H}) \quad (\text{Equation 5.16})$$

where  $\alpha_H$  is the Hill coefficient and  $K_d$  represents the macroscopic dissociation constant ( $K_d = K_a^{-1}$ ). Nonlinear least-squares fitting was performed using Prism software. Goodness of fit was assessed by visual inspection of residuals and by monitoring of fitting statistics.

***Quantitative DNase I ssDNA footprinting.*** To monitor binding of N-HisPur $\beta$  monomers to individual sites of the purine-rich strand of the SM $\alpha$ A MCAT enhancer element, quantitative DNase I footprinting was performed based on methods described by Ackers and coworkers (1, 22, 23, 251), with the following modifications. All binding reactions were carried out in buffer containing 50 mM Tris-HCl pH 7.5, 100 mM KCl, 1 mM dithiothreitol, 1 mM CaCl<sub>2</sub>, 2.5 mM MgCl<sub>2</sub>, 2  $\mu$ g/mL dT<sub>32</sub>, and 50  $\mu$ g/mL BSA at 20  $\pm$  1°C and allowed to equilibrate for 16 h. Each reaction contained 20,000 cpm of freshly

labeled SMP382-F\* template, at a final concentration estimated to be well enough below N-HisPur $\beta$  binding affinity to maintain validity of the  $[P_{free}] \approx [P_{total}]$  assumption ( $< 25$  pM). N-HisPur $\beta$  was added to each reaction at the indicated concentrations to cover a range from approximately  $10^{-13}$  to  $10^{-8}$  M in a final volume of 200  $\mu$ L. After equilibrium had been reached, template digestion was initiated by the addition of 5  $\mu$ L of a DNase I solution in assay buffer to reach a final concentration of 1.0 Units/mL. Digestion was allowed to proceed for 2 min at  $20 \pm 1^\circ\text{C}$  and was stopped by the addition of 700  $\mu$ L of stop solution (97% ethanol, 0.57 M ammonium acetate, 50  $\mu$ g/mL yeast tRNA) and incubated in a dry ice-ethanol bath for 30 min. Single-stranded DNA was pelleted by centrifugation, washed twice with 70% ethanol, and air dried. Pellets were dissolved in 5  $\mu$ L of buffer containing 98% formamide, 10 mM EDTA, 0.1% bromophenol blue, and 0.1% xylene cyanol. Samples were heated to  $95^\circ\text{C}$  for 5 min prior to being electrophoresed for 130 min at 65 Watts on a 0.4 mm thick sequencing gel consisting of 8% polyacrylamide (29:1 acrylamide:bisacrylamide), and 8M urea, cast and run in TBE buffer. Gels were pre-run at 65 Watts for at least 2 h or until a gel temperature of  $\geq 50^\circ\text{C}$  was attained. End-labeled primer dideoxy-NTP sequencing reactions performed with double-stranded SMP382 PCR product as a template and SMP8p741s-F\* as the extension primer were also electrophoresed in order to identify sequences of interest on the resulting footprints. Dideoxy-NTP sequencing reactions were performed according to the manufacturer's instructions for end-labeled primer sequencing (USB Corp.), and typically 10,000-15,000 cpm per reaction were loaded on the sequencing gels. Following electrophoresis, gels were dried in a Bio-Rad slab gel dryer on Whatman filter paper at

75°C for 45 min. Dried gels were exposed to phosphor storage screens for 72-96 h and phosphorimaged as described above. Densitometry of phosphorimages was performed using ImageQuant software (Molecular Dynamics) in order to determine values for fractional protection ( $F_p$ ) of a given sequence:

$$F_p = 1 - [(I_{q, site} / I_{q, control})] / [(I_{r, site} / I_{r, control})] \quad (\text{Equation 5.17})$$

where  $I$  is again the relative densitometric intensity,  $q$  refers to any lane of the gel with finite N-HisPur $\beta$  concentration,  $r$  refers to the reference lane containing no protein,  $site$  refers to the ssDNA-binding site in question, and  $control$  refers to a region of the gel whose intensity is independent of protein ligand concentration (bases -218 to -210 of SMP382-F\*). First line, crude analysis of fractional protection data was done by fitting data to the familiar Langmuir isotherm, primarily to determine the upper and lower endpoints of fractional protection since binding of protein ligands at specific sites, even at saturating conditions, does not provide complete protection:

$$f = u \cdot \{k[P_{free}] / (1 + k[P_{free}])\} + l \quad (\text{Equation 5.18})$$

Where  $k$  refers to the microscopic association constant (assuming no interaction between sites),  $u$  and  $l$  refer to the upper and lower endpoints, respectively. Fractional protection values were converted to values of fractional saturation,  $\bar{Y}$ , using the following expression and the values of  $u$  and  $l$  from equation 5.19:

$$\bar{Y} = (f - l) / (u - l) \quad (\text{Equation 5.19})$$

***Resolution of binding mechanism and microscopic association and interaction free energies.*** Constructions of individual site binding expressions were done using a statistical thermodynamics approach that has been described and applied previously (1). Briefly the probability of the MCAT enhancer element existing in any one of the microscopic configurations depicted in Table 5.2,  $f_s$ , can be expressed as indicated in equation 5.11. Microscopic association constants are related to microscopic free energies through the relationship,  $\Delta G_i = -RT \ln k_i$ . The fractional saturation,  $\bar{Y}$ , of a given site can be expressed as the sum of the probabilities of configurations in which the indicated site is occupied. Applying this treatment for each of the models depicted in Table 2 gives expressions describing fractional saturation at each site, A and/or B, in terms of their microscopic association constants and N-HisPur $\beta$  monomer concentration:

$$\text{Obligate Dimer} \quad \bar{Y}_{A,B} = \frac{k_1 k_{dil} [P_{free}]^2}{1 + k_1 k_{dil} [P_{free}]^2} \quad (\text{Equation 5.20})$$

$$\text{Identical, independent} \quad \bar{Y}_{A,B} = \frac{k_1 [P_{free}] + k_1^2 [P_{free}]^2}{1 + 2k_1 [P_{free}] + k_1^2 [P_{free}]^2} \quad (\text{Equation 5.21})$$

$$\text{Identical, interacting} \quad \bar{Y}_{A,B} = \frac{k_1 [P_{free}] + k_1^2 k_c [P_{free}]^2}{1 + 2k_1 [P_{free}] + k_1^2 k_c [P_{free}]^2} \quad (\text{Equation 5.22})$$

Nonidentical, independent  $\bar{Y}_A = \frac{k_1[P_{free}] + k_1k_2[P_{free}]^2}{1 + (k_1 + k_2)[P_{free}] + k_1k_2[P_{free}]^2}$  (Equation 5.23)

$\bar{Y}_B = \frac{k_2[P_{free}] + k_1k_2[P_{free}]^2}{1 + (k_1 + k_2)[P_{free}] + k_1k_2[P_{free}]^2}$  (Equation 5.24)

Nonidentical, interacting  $\bar{Y}_A = \frac{k_1[P_{free}] + k_1k_2k_c[P_{free}]^2}{1 + (k_1 + k_2)[P_{free}] + k_1k_2k_c[P_{free}]^2}$  (Equation 5.25)

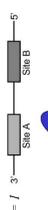
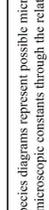
$\bar{Y}_B = \frac{k_2[P_{free}] + k_1k_2k_c[P_{free}]^2}{1 + (k_1 + k_2)[P_{free}] + k_1k_2k_c[P_{free}]^2}$  (Equation 5.26)

Individual site isotherms were globally fit to equations describing each model using Prism 5 software (GraphPad Software, Inc., San Diego, CA). Monte Carlo error simulations for the estimate of model confidence and parameter constraints were also performed using Prism 5.

## RESULTS

*Extent of Pur $\beta$  binding to the SM $\alpha$ A proximal MCAT enhancer element.* Specific binding of sequence-specific ssDNA binding proteins, namely Pur $\alpha$ , Pur $\beta$  and MSY1, to the proximal MCAT enhancer element of the mouse SM $\alpha$ A gene promoter represents an important regulatory event for the repression of SM $\alpha$ A gene expression in response to repressive stimuli in a variety of cell types (28, 47, 48, 82, 141, 144, 148, 275). In the present study we sought to examine the nature of nucleoprotein complex interactions

**Table 5.2. Microscopic configurations and corresponding interaction free energies for multiple models of two site-binding.**

Species Representation <sup>1</sup>	Obligate Dimer		Identical Independent		Identical Interacting		Non-identical Independent		Non-identical Interacting	
	Free Energy Contribution <sup>2</sup>	Microscopic Constant	Free Energy Contribution <sup>2</sup>	Microscopic Constant						
$s = 1$ 3 	Reference state	-	Reference state	-						
$s = 2$ 	-	$k_1$	$\Delta G_1$	$k_1$						
$s = 3$ 	-	$k_1$	$\Delta G_1$	$k_1$	$\Delta G_1$	$k_1$	$\Delta G_1$	$k_2$	$\Delta G_2$	$k_2$
$s = 4$ 	-	$k_1^2$	$2\Delta G_1$	$k_1^2$	-	-	$\Delta G_1 + \Delta G_2$	$k_1 k_2$	-	-
$s = 5$ 	$\Delta G_{11} + \Delta G_1$	$k_{11} k_1$	-	-	$2\Delta G_1 + \Delta G_c$	$k_1^2 k_c$	-	-	$\Delta G_1 + \Delta G_2 + \Delta G_c$	$k_1 k_2 k_c$

<sup>1</sup>Species diagrams represent possible microscopic configurations and do not necessarily reflect structural perturbations as a result of ligation, only that facilitation of binding occurs in the case of interacting sites. <sup>2</sup>Changes in free energy are related to microscopic constants through the relationship  $\Delta G = -RT \ln k$ , where  $R$  is the gas constant and  $T$  is the absolute temperature. The fractional saturation,  $\gamma$ , of each site is given as the sum of probabilities,  $f_s$ , of monomer occupation at each site for each microscopic species,  $s$ , where

$$f_s = \frac{e^{-\Delta G(s, n_1)} P_{s, n_1}^m}{\sum_{s'} e^{-\Delta G(s', n_1)} P_{s', n_1}^m}$$

formed between Pur $\beta$ , which has been shown to be the dominant Pur protein repressor of SM $\alpha$ A at the proximal MCAT enhancer element in relevant cell-culture models (148, 149), and the purine-rich strand of the fore mentioned enhancer element. Towards this aim, we used a combination of techniques that examine the effects of N-HisPur $\beta$  concentration on the solution and electrophoretic properties of the representative single-stranded oligonucleotide probe PE32-F. Both the electrophoretic mobility and fluorescence anisotropy of PE32-F were cumulatively affected by increasing concentrations of N-HisPur $\beta$  (Figure 5.2, panels A and B) indicative of binding. Electrophoretic analysis of N-HisPur $\beta$ :PE32-F\* nucleoprotein complexes suggest that N-HisPur $\beta$  binds to PE32-F\* (2 nM) in a sequential manner, as indicated by the appearance of at least three discrete molecular species, and shifting of the higher mobility species to complexes with lower mobility upon titration of N-HisPur $\beta$  over a concentration range of 0.41 nM to 400 nM (Figure 5.2, panel A). The co-appearance of the two highest mobility species (speculated as adopting 1:1 and 2:1 complex stoichiometries, respectively) at lower protein concentrations suggests that the 2:1 complex represents a high affinity complex, possibly relying on cooperative facilitated assembly on a preformed 1:1 complex, as the higher mobility complex disappears at higher protein concentrations, and the 2:1 species predominates over the applied N-HisPur $\beta$  concentration range. The appearance of higher molecular weight species (3:1 and greater) at the highest N-HisPur $\beta$  concentrations shows that these species likely represent low affinity, and possibly non-



binding of N-HisPur $\beta$  to fluorescent probe PE32-F-3FLC (50 nM). Data was fit to equation 5.6 to determine the equivalency point (dashed line). Fixing of  $K_r$  to near-zero values (infinite affinity, solid line) verified the equivalency transition at an  $R_{P/D}$  value of 2:1. Symbols represent titrations from different preparations of N-HisPur $\beta$ , indicating similar number of active binding sites. **C**, Qualitative electrophoretic mobility (super)shift assay of PE32-F\* nucleoprotein complexes. The relative electrophoretic mobilities of nucleoprotein complexes containing PE32-F\* and either recombinant, or nuclear extract-derived Pur proteins were compared to identify the predominant and presumably biologically relevant species. The predominant species observed with 5 nM N-HisPur $\beta$  and 1 nM PE32-F\* (lane 2) displays mobility similar to that seen from nuclear extracts (lanes 4-6), as verified by specific supershifting with anti-Pur $\beta$  302 antibody (lanes 13 and 14). Similar findings were observed for Pur $\alpha$  (compare lane 7 with lanes 4-6, and supershift lanes 16 and 17). The indicated identities of observed complexes are hypothetical.

specific interactions of N-HisPur $\beta$  with free ssDNA sites and/or protein-protein interaction sites on preformed nucleoprotein complexes.

Interpretations of the preliminary electrophoretic mobility shift results described above were verified by experiments in which binding of N-HisPur $\beta$  to fluorescein-labeled PE32-F (PE32-F-3FLC) at saturating conditions (probe concentration >50 times the reported dissociation constant of approximately 1 nM) was monitored by changes in the measured anisotropy of the probe. The results of this analysis are shown in panel B of Figure 5.2. Under saturating binding conditions as applied in this analysis, it is evident that N-HisPur $\beta$  can saturate PE32-F-3FLC to a specific terminal stoichiometry of 2:1, as indicated by the returned stoichiometry values from non-linear least-squares regression of both data sets to equation 5.6, and the equivalency transition point when the apparent affinity value,  $K_r$ , is fixed at values approaching zero (near-infinite affinity, solid line). Curvature in the best fit line at values of  $R_{P/D}$  near the equivalence point when  $K_r$  is not fixed (dashed line) supports the notion that either the total concentration of probe is not sufficiently high enough to ensure that every protein molecule is bound to the probe at lower protein concentrations, a condition that is only satisfied when  $D_{total} / K_d \gg 10$ , in the case of a completely cooperative system (228), that a non-specific binding event occurs at higher concentrations leading to higher order stoichiometries, or both. Three lines of evidence support the latter possibility: 1) Affinities of Pur $\beta$  for single-stranded *PUR*-elements have been reported to be in the 0.5 to 1.5 nM range (148, 310), whereas in the experiments performed here, PE32-F-3FLC were held at 50 nM. 2) A positive slope in the plateau region indicates that additional ligand binding is occurring at higher protein

concentrations. This observation could also be explained by an increase in the solution viscosity as a function of protein concentration, however an anisotropy change of PE32-F-3FLC in the presence of increasing concentrations of BSA was not observed nor did the anisotropy of PE32-F3FLC change in the context of a double-stranded probe with increasing concentrations of N-HisPur $\beta$  (data not shown), both suggesting that protein concentration viscosity gradients are not significant in these concentration ranges. 3) The electrophoretic mobility shift profile of PE32-F\* at high N-HisPur $\beta$  concentrations displayed in Figure 5.2A shows that species formed at these conditions are likely of higher order (greater than 2:1), as indicated by the presence of two faster moving species at lower protein concentrations. Collectively these data show that N-HisPur $\beta$  has the capacity to form complexes with PE32-F or PE32-F-3FLC beyond a 1:1 stoichiometry, and the extent of binding follows, most stably, the formation of a 2:1 complex but does not appear terminable; however the extent of binding that is biologically relevant remains uncertain.

To assess the extent of PE32-F binding by Pur $\beta$  in the context of the nuclear environment, and to compare this to levels observed with recombinant N-HisPur $\beta$  under controlled conditions, comparative, qualitative EMSAs were performed. Figure 5.2C demonstrates the electrophoretic mobilities of PE32-F\*-containing nucleoprotein complexes derived from equilibration of PE32-F\* (1 nM) with recombinant N-HisPur $\beta$ , N-HisPur $\alpha$ , and AKR-2B MEF nuclear extract-derived proteins. Consistent with observations made in panel A, N-HisPur $\beta$ , at a concentration of 5 nM shifts PE32-F\* into two visible bands (lane 2), speculated to be N-HisPur $\beta$ <sub>1</sub>:PE32-F\* (lower band) and N-

HisPur $\beta_2$ :PE32-F\* (upper band) based on findings presented above. N-HisPur $\alpha$  (5 nM, lane 7) shows a single shifted band dissimilar to N-HisPur $\beta$ , but likely corresponds to N-HisPur $\alpha_2$ :PE32-F\* despite slight differences in mobilities compared to the N-HisPur $\beta$  equivalents. Mobility disparities are consistent with differences in the monomeric molecular weights and calculated charges at pH 8. AKR-2B MEF nuclear extracts shift PE32-F\* into a lower band, and an upper group of closely migrating bands that do not resolve well under the electrophoretic conditions applied here (lanes 4-6). Longer runs suggest that, indeed, this upper grouping is composed of several closely migrating bands, however, their complete resolution is hindered due to streaking of bands with longer gel-run times (data not shown). Incubation of nuclear extract binding reactions with an antibody that specifically binds Pur $\beta$  (anti-Pur $\beta$  302, lanes 13 and 14) clearly identifies the lower band as the Pur $\beta$ :PE32-F\* complex, as indicated by the complete disappearance of this band in lanes 13 and 14, and the appearance of supershifted complexes, SS1 and SS2. Additionally, the presence of anti-Pur $\beta$  302 causes some loss of the faster migrating portion of the upper doublet, suggesting that this species might be heteromeric Pur $\alpha/\beta$ :PE32-F\* complexes. Incubation of nuclear extract binding reactions with an antibody that specifically binds Pur $\alpha$  (anti-Pur $\alpha$  291, lanes 15 and 16) further divulge the compositions of the slower migrating bands generated by nuclear extract material. The upper band(s) in lanes 4-6 likely correspond to Pur $\alpha$ :PE32-F\*, or heteromeric Pur $\alpha/\beta$ :PE32-F\* complexes due to, again, loss of the faster migrating portion of the doublet, and some loss of the slower migrating portion upon antibody addition, and the appearance of supershifted complexes. Persistence of some slower migrating species

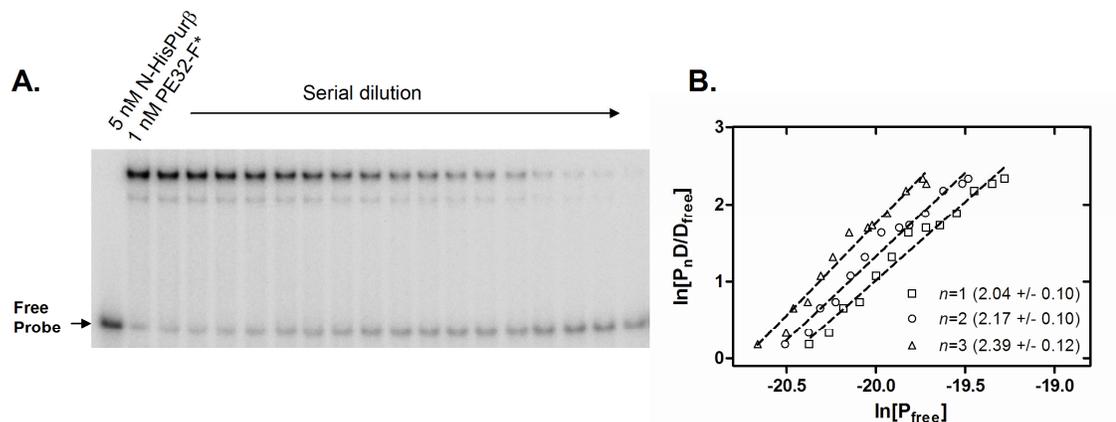
in the presence of 1  $\mu\text{g}/\text{mL}$  anti-Pur $\alpha$  291 suggests that these bands might represent nucleoprotein complexes of unknown composition, or that anti-Pur $\alpha$  291 possesses lower affinity for Pur $\alpha$ -containing nucleoprotein complexes than does anti-Pur $\beta$  302 for Pur $\beta$ -containing nucleoprotein complexes, a possibility supported by supershifts conducted on recombinant Pur:PE32-F complexes (compare lanes 11 and 19). Previous studies have suggested that Pur $\beta$  is the dominant Pur protein repressor in the context of SM $\alpha$ A expression (145, 148), despite the fact that the intense slower moving complex observed contains Pur $\alpha$ , as indicated by reactivity with anti-Pur $\alpha$  291 antibody. This finding suggests that either Pur $\alpha$  is more abundant than Pur $\beta$  in asynchronous AKR-2B nuclear extracts, or displays higher affinity for PE32-F. Previous studies have shown that Pur $\alpha$  binds with higher affinity than Pur $\beta$  to both PE32-F (148), and to a 24mer representative of the *c-myc* PUR element (310).

Comparing the predominant N-HisPur $\beta$  shifted complex of PE32-F\* (speculated as N-HisPur $\beta_2$ :PE32-F\*, lane 2) and the lower, Pur $\beta$ -containing band generated by equilibration with nuclear extract-derived proteins (lower band, lanes 4-6), it is further conjectured that the latter is likely Pur $\beta_2$ :PE32-F\*. This inference is based on the similar relative mobilities of the two complexes. Slight differences in electrophoretic mobility between the recombinant and endogenous nucleoprotein complexes are consistent with molecular weight and charge differences between N-HisPur $\beta$  (35,168.6 Da, -5.05) , and Pur $\beta$  (33,901.3 Da, -5.15). In conclusion, the predominant shifted species observed when 5 nM N-HisPur $\beta$  is equilibrated with 1 nM PE32-F\* is equivalent to the nuclear extract

derived-endogenous Pur $\beta$ :PE32-F\* complex. However, the exact stoichiometry of this complex can only be inferred from anisotropy experiments, described above, as being 2:1.

Direct assessment of the predominant electrophoretic nucleoprotein complex generated from equilibration of 5 nM N-HisPur $\beta$  with 1 nM PE32-F\* was performed using a serial dilution-coupled EMSA. Briefly, a reaction mixture containing N-HisPur $\beta$  and PE32-F\* at the forementioned concentrations, was serially diluted at a ratio of 1.1 to 1 with buffer to create a series of reaction samples with the same molar ratio of components, but at differing concentrations, and a distribution of reversibly interacting species governed by laws of mass-action. Quantification of molecular species by densitometric methods subsequent to electrophoretic separation permits determination of the system stoichiometry by way of a value convergence method, as described in the methods section of this paper. Implementation of this method, as shown in Figure 5.3, shows that the persistent, and stable high-affinity nucleoprotein complex formed between N-HisPur $\beta$  and PE32-F\* adopts a stoichiometry of 2 to 1, as convergence between estimated values of  $n$  and those obtained by linear regression of the  $\ln[P_n D/D_{free}]$  vs  $\ln[P_{free}]$  plot occurred at  $n = 2$ .

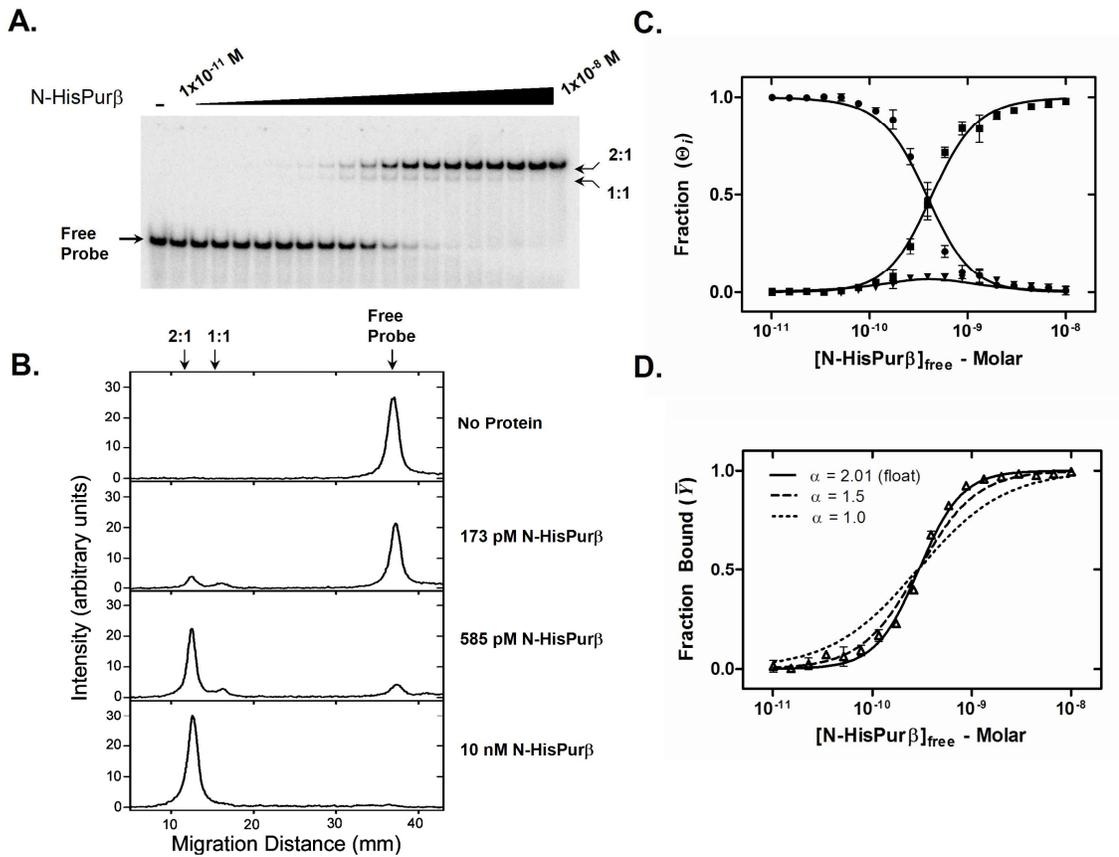
***Affinity of specific N-HisPur $\beta$ /ssDNA interactions.*** The reported ssDNA binding characteristics that unify Pur proteins, is their preference for interaction with purine-rich sequences, but also their affinity for such sequences (148, 310). However, an extensive and thermodynamically rigorous assessment of the free energy of ssDNA-binding by Pur proteins has not yet been reported. To address this fact, we performed direct titrations of



**Figure 5.3 Serial dilution EMSA determination of the N-HisPur $\beta$ :PE32-F\* nucleoprotein complex stoichiometry.**

**A.** Limited serial dilution of a mixture of N-HisPur $\beta$  and PE32-F\* (5 nM and 1 nM, respectively, serially diluted 1.1:1.0) was performed and subjected to quantitative EMSA to determine the stoichiometry of the N-HisPur $\beta$ :PE32-F\* complex, as exemplified in Figure 5.2. Intensity of the *Free Probe* band was quantified and standardized to known quantities of PE32-F\* (lanes 1 and 20) to determine the concentration of free probe,  $[D_{free}]$ . The concentration of the nucleoprotein complex  $[P_nD]$  with stoichiometry  $n$ , was determined from the known total concentration of PE32-F\* loaded in each lane using the relationship  $[P_nD] = [D_{total}] - [D_{free}]$ . The concentration of free protein was estimated using the relationship  $[P_{free}] = [P_{total}] - n[P_nD]$ , in which  $n$  is an integer estimation. Isotherms with varying values of  $n$  were plotted as shown in **B**. Each point represents the mean of duplicate experiments. Error bars were omitted for clarity. Dashed lines represent the least-squares regressions of each data set to the relationship  $\ln[P_nD/D_{free}] = n \cdot \ln[P_{free}] + \ln K_a$ , where  $K_a$  represents the macroscopic association constant for the general equilibrium  $nP + D \rightleftharpoons P_nD$ . Numbers in parentheses reflect the returned regression fit value of  $n \pm s.d.$

N-HisPur $\beta$  against 25 pM PE32-F\*, a condition that maintains validity of the assumption  $[P_{free}] \approx [P_{total}]$ , which is necessary for mathematical modeling of binding reactions (228). Molecular species distributions at equilibrium were separated by native gel electrophoresis and quantified as described in the methods section. Results of this analysis are shown in Figure 5.4. Panel A shows the electrophoretic mobility shift profile of PE32-F\* as a result of binding in the presence of increasing concentrations of N-HisPur $\beta$ . Visualization and densitometric analysis confirms the presence of three separate bands or pixel intensity peaks (Figure 5.4B) corresponding to free probe, and two shifted complexes that are (N-HisPur $\beta$ )<sub>1</sub>:PE32-F\* (1:1 complex) and (N-HisPur $\beta$ )<sub>2</sub>:PE32-F\* (2:1 complex) as identified in the previous section. Of special interest is the transient nature of the 1:1 complex with respect to N-HisPur $\beta$  concentration. Comparing lane intensity profiles in Figure 5.4B for increasing N-HisPur $\beta$  concentrations shows that the 1:1 complex is not a protein preparation contaminant as the peak intensity reaches a maxima at moderate N-HisPur $\beta$  concentrations and drops-off as PE32-F\* becomes saturated. Furthermore, this peak pattern has been theoretically assigned to systems that adopt cooperative, two-site binding mechanisms (27). Peak integration for each species provided the isotherms observed in Figure 5.4C. Species specific isotherms  $\Theta_0$ ,  $\Theta_1$ , and  $\Theta_2$  were globally fit to equations 5.12, 5.13 and 5.14, respectively, to resolve values for the macroscopic association constants  $K_1$ , and  $K_2$ . This approach yielded  $K_1 = 3.43 (\pm 0.368) \times 10^8 \text{ M}^{-1}$ , and  $K_2 = 6.06 (\pm 0.191) \times 10^{18} \text{ M}^{-1}$ . It has been previously demonstrated that in instances in which the nature of two-site binding of protein ligands to DNA is unknown (i.e.,



**Figure 5.4. Quantitative electrophoretic mobility shift assay of N-HisPurβ binding to PE32-F\*.** **A**, N-HisPurβ was titrated over a concentration range of  $10^{-11}$  to  $10^{-8}$  M and equilibrated with 25 pM PE32-F\* prior to subjecting reaction mixtures to quantitative EMSA. The presence and transient intermediate nature of the visible (N-HisPurβ)<sub>1</sub>:PE32-F\* (1:1) complex upon titration suggests that binding of N-HisPurβ to PE32-F\* involves a sequential mechanism. **B**, Densitometric analysis of lanes of gel shown in **A** verify the presence of three pixel intensity peaks indicative of separate electrophoretic species. Titration of N-HisPurβ confirms that the 1:1 complex does not accumulate significantly compared to the free probe or 2:1 peaks, suggestive of a

sequential and cooperative binding mechanism. **C**, Individual band intensities are plotted as a function of  $[N\text{-HisPur}\beta_{free}]$  (●, free Probe; ▼, 1:1 complex; ■, 2:1 complex). Each point represents the *mean*  $\pm$  *s.d.* of quadruplicate experiments. Lines are global nonlinear least squares fits of individual species isotherms to equations 5.12-5.14. **D**, Band intensity data from **A** presented as *Fraction Bound*,  $\bar{Y}$ , ( $\bar{Y} = ([D_{total}] - [D_{free}]) / [D_{total}]$ ) versus  $[N\text{-HisPur}\beta_{free}]$ , were fit to the phenomenological Hill equation (equation 5.16). The Hill coefficient,  $\alpha_H$ , was held constant at values of 1.5 (dashed line) and 1.0 (dotted line) to reflect the dependency of this variable on goodness of fit.

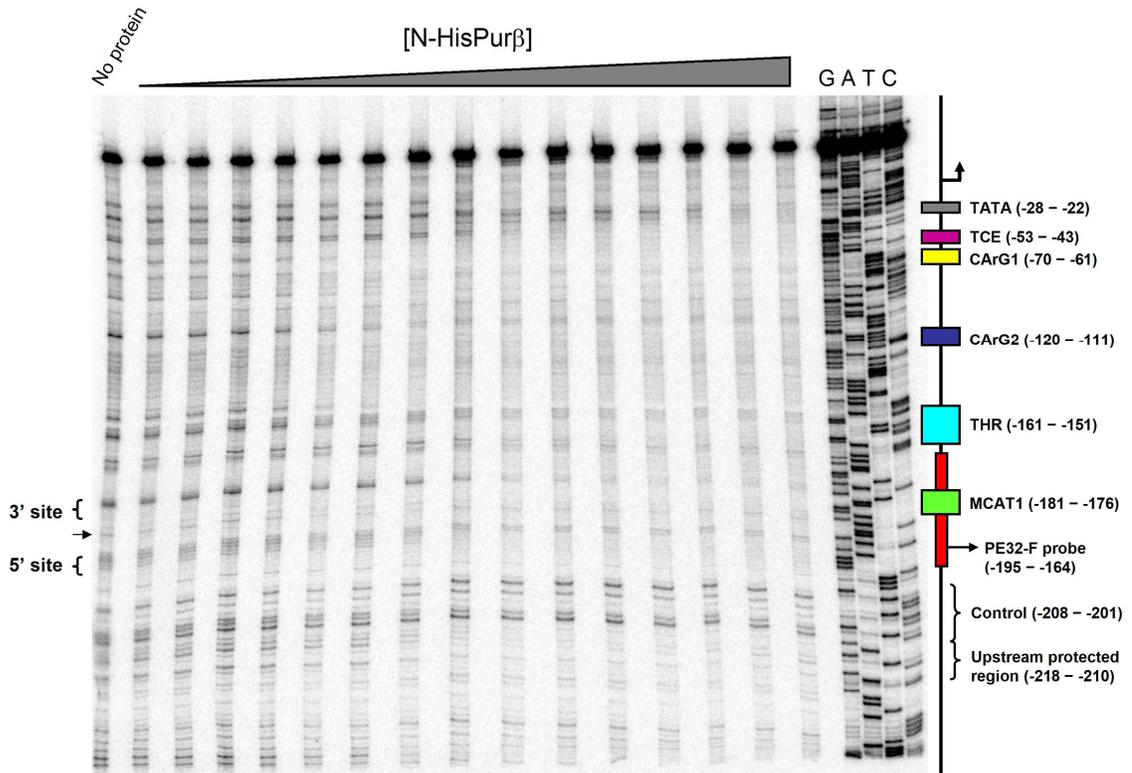
identical versus non-identical sites, positive or negative cooperativity) microscopic equilibrium constants cannot be definitively determined from the resolved macroscopic terms, regardless of their precision (250). Despite this inability, inferences on the nature of binding can be made. For instance, Seneor and Brenowitz showed that whenever  $K_2 > K_1^2/4$ , as is observed here, it can be inferred that either ligand binding exhibits positive cooperativity ( $k_c > 1$ ) if binding sites are identical ( $k_1 = k_2$ ), binding sites are non-identical ( $k_1 \neq k_2$ ), or a combination of the two. EMSA cannot independently discern these possibilities. Evidence of positive cooperativity is also observed by analyzing quantitative EMSA data by way of fitting the fractional saturation data to the Hill equation (equation 5.16). In order to circumvent quantification issues arising from the presence of multiple shifted complexes and the streaking of bands due to system reversibility, the extent of binding was determined by the amount of free probe in each lane, which likely represents the extent of binding at equilibrium prior to electrophoresis (27). This approach resulted in the isotherm presented in Figure 5.4, panel D. Non-linear least-squares fitting of the generated isotherm to the Hill equation returned a macroscopic dissociation constant,  $K_d$ , of approximately 300 pM, which is in close agreement to previously reported values for Pur $\beta$  (148, 310). The returned Hill coefficient,  $\alpha_H$ , ( $2.01 \pm 0.07$ ), reflects that binding of N-HisPur $\beta$  to PE32-F\* is cooperative, since  $\alpha_H$  converges at a value close to the value of  $n$ , as we have described in this report. It has been shown that values of  $\alpha_H$  approach the system stoichiometry only in cases where positive cooperativity is present (305). Fits of the data to the Hill equation in which  $\alpha_H$  was fixed

at values of 1.5 and 1.0 (Figure 5.4D) show the dependence of the goodness of fit on this variable, and that positive cooperativity is present.

Collectively, these data suggest that N-HisPur $\beta$  binds to the purine-rich strand of the SM $\alpha$ A proximal MCAT enhancer element with an apparent affinity in the sub-nanomolar range, and that binding of the protein likely proceeds via a sequential and cooperative mechanism. However, the exact mechanism cannot be resolved by the methods described above, due mainly to the fact that it is currently unknown whether or not the sequence represented by PE32-F\* contains two identical and independent sites, identical and interacting sites, non-identical and independent sites, or non-identical and interacting sites.

*Resolution of binding mechanism and microscopic interaction free energies.* As suggested above, despite our insight into the apparent affinity of purine-rich ssDNA binding by Pur $\beta$ , a concise description of the binding mechanism and energetics is elusive due to the inability of simple quantitative binding assays to delineate these points. To resolve these issues, we employed a quantitative DNase I footprinting technique to measure site-specific fractional saturation of the forward strand of the SM $\alpha$ A promoter corresponding to bases -323 to +59 (relative to the transcription start site, termed SMP382-F) in response to N-HisPur $\beta$  concentration. The power offered by this method is its ability to examine the quantitative nature of binding of protein to sites of interest on the DNA template, thus providing a means to determine microscopic binding constants and to discriminate between possible binding mechanisms.

The effects of N-HisPur $\beta$  titration on the DNase I protection profile of SMP382-F\* is shown in Figure 5.5. Numerous sites of protection, which can be inferred as binding, can be seen. Worthy of noting are regions of the SM $\alpha$ A promoter that have been reported to be responsive to Pur protein interaction. These regions correspond to the TGF $\beta$ 1 control element, TCE (272), the TGF $\beta$ 1 hypersensitive region, THR (47), CArG element 1 and 2 (103) and the proximal MCAT enhancer element (28, 48, 141, 144, 148, 275), as represented by PE32-F. Also worth noting, is protection of regions that have thus far gone unreported. Specifically, a region adjacent to, and downstream of the TATA and an upstream protected region in the vicinity of bases -218 to -210 show levels of protection. The extents of actual binding by Pur $\beta$  to these regions *in vivo*, as well as the functional consequences of these interactions, are unknown at this point. It must be submitted that the observed binding of N-HisPur $\beta$  to these regions could be an artifact of non-specific binding created by complete single-strandedness of the template (opportunistic, non-specific binding), despite apparent high affinity. It is unlikely that the observed protection is an artifact generated by induced secondary structures in the template DNA coupled with DNase I substrate preferences, as it has been reported previously that DNase I shows a substrate preference and enhanced cleavage rates for dsDNA over ssDNA, presumably due to the enzyme's requirement for interaction with the minor groove of B-DNA (73, 273, 276). Thus, hypersensitive regions likely correspond to self-complementary regions adopting double-stranded conformations. The impacts these structures have on binding of N-HisPur $\beta$  to regions of interest are unknown.



**Figure 5.5. Quantitative DNase I footprinting of N-HisPurβ:SMP382-F\* interactions.** Representative footprint titration analysis of N-HisPurβ binding to SMP382-F\* shows two regions of protection adjacent to the core MCAT enhancer element and within the region represented by PE32-F\* (marked as 3' site and 5' site). Regulatory cis-elements are mapped next to lanes containing dideoxy-sequencing reactions (G, A, T, and C). The protected sites within the PE32-F region are separated by a band with N-HisPurβ concentration-independent pixel intensity (arrow), when normalized to pixel intensity of the control region (-208 to -201) suggesting this intervening region is not protected by N-HisPurβ-binding. Other sites of protection are noted within THR, TCE, and CArG boxes 1 & 2, as previously described, as well as a

previously uncharacterized upstream region (-218 to -210), and a region adjacent to the TATA box.

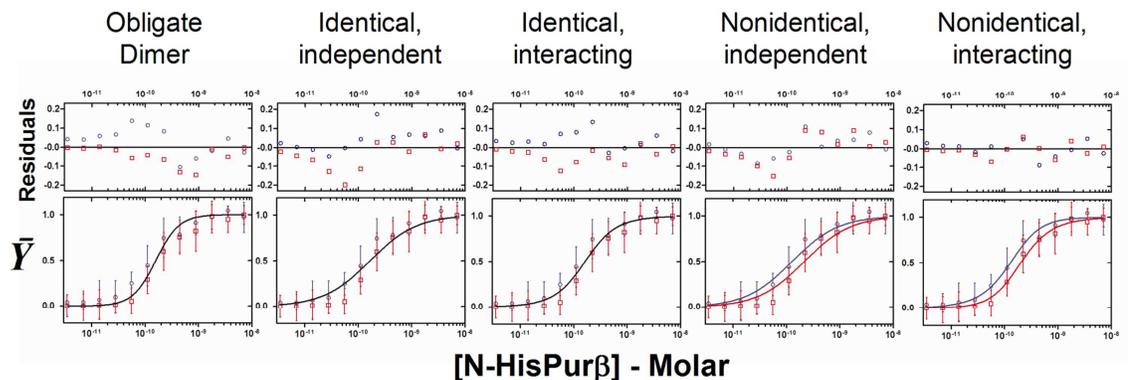
A goal of our approach was to quantitatively interrogate binding of N-HisPur $\beta$  to the region encompassing the proximal MCAT enhancer (bases -195 to -164). Close examination of this region by densitometric methods yields two sites of protection afforded by N-HisPur $\beta$  titration. These sites have been termed the 3', and 5' sites, and are labeled as such in Figure 5.5. It should be noted that these sites are separated by a region that is not protected when compared to the control region (-208 to -201), and is marked by an arrow. Based on our knowledge of the stoichiometry of the nucleoprotein complex formed between N-HisPur $\beta$  and this sequence, this nature of protection appears appropriate.

Mathematical expressions describing various models of interactions between N-HisPur $\beta$  and the SM $\alpha$ A MCAT enhancer element deemed appropriate as based on the measured stoichiometry of the N-HisPur $\beta$ :PE32-F nucleoprotein complex were generated by methods outlined previously (1). For complexes with a saturating stoichiometry of 2:1 five possible models can be proposed. 1) The first possibility is that of a preformed (obligate) dimer assembling on a single binding site. Other possible models involve sequential assembly of monomers on the ssDNA lattice in which binding sites are: 2) identical and independent, 3) identical and interacting, 4) non-identical and independent, and 5) non-identical and interacting. The various macromolecular configurations for each ligation state allowed by the restrictions of each binding model are depicted in Table 2 along with the corresponding free energy contributions and equilibrium constants used for constructing expressions of N-HisPur $\beta$  binding.

Several assumptions have been made in order to constrain and mathematically define the models described above: A) Binding of N-HisPur $\beta$  to the MCAT enhancer region (bases -195 to -164 of the SM $\alpha$ A promoter) is independent of binding of N-HisPur $\beta$  to other regions of the promoter, outside of this vicinity. It should be submitted that this assumption may not be valid based on the protection profile observed, in which multiple sites display binding. However, at this time the complete complex stoichiometries obtained at these sites are unknown. B) DNA-independent self association of N-HisPur $\beta$ , under the conditions used here, is defined by the pre-determined equilibrium constant of  $k_{di} = 884955 \text{ M}^{-1}$  (222), and is negligible in cases of sequential monomeric assembly, where half saturation of sites is in the subnanomolar range (see results section). It should be noted that the reported dimerization constant for N-HisPur $\beta$  was determined under different solution conditions (higher ionic strength) than those applied in the present study, and the absolute effects of salt identity and concentration on self-association have not been explored thoroughly, due mainly to solubility problems associated with N-HisPur $\beta$  at concentrations necessary for measurement by analytical centrifugation. However, since we have observed that sodium chloride concentrations as high as 1M do not alter the hydrodynamic properties of N-HisPur $\beta$  as judged by size-exclusion chromatography (149), we have assumed that self-dimerization is not significantly perturbed under the conditions applied here. C) Where applicable, identical binding sites exhibit equal intrinsic binding free energy changes upon ligand binding, as illustrated with  $\Delta G_I$  designated as this free energy change for both sites. D) Likewise, non-identical sites exhibit non-equivalent intrinsic binding free energy changes upon ligand

binding, and as such, are designated with individual parameters  $\Delta G_1$  and  $\Delta G_2$  to reflect this prediction. E) In binding pathways possessing intersite interaction, the difference in the change in free energy between binding of each monomer to a ssDNA site individually and the total free energy change observed upon facilitated binding is represented as  $\Delta G_c$ . Microscopic species-specific free energy changes and microscopic equilibrium constant assignments are shown in Table 2. The construction of expressions describing site specific fractional saturation was performed as described in the methods section.

Densitometric quantification of N-HisPur $\beta$  binding to the identified 3', and 5' sites of the proximal MCAT enhancer element was performed to generate individual site binding isotherms, and are displayed in Figure 5.6. As can be seen, isotherms generated by binding of N-HisPur $\beta$  by these sites are non-identical, suggesting that binding does not proceed through an obligate dimer pathway, and that sites are also non-identical. Global fitting of individual site isotherms to the model-specific expressions outlined in the methods section (equations 5.20-5.26), provided a means to further discriminate between models based on goodness of fit. From this approach, it can be seen that binding of N-HisPur $\beta$  to the 3' and 5' sites of the proximal MCAT enhancer element proceeds through a sequential mechanism, in which sites are non-identical and interacting, as judged by random distribution of residuals and fit statistics. More specifically, binding of the 3' site occurs with highest affinity, and that the 3' and 5' sites are non-identical and interacting, with binding to the 5' site being cooperatively facilitated. This model is in line with the results of the qualitative and quantitative electrophoretic mobility shifts presented in previous sections, which suggested that binding of N-HisPur $\beta$  proceeds via a sequential



**Figure 5.6. N-HisPur $\beta$  binds to the SM $\alpha$ A MCAT enhancer element via a cooperative 3' to 5' mechanism.** Individual site binding isotherms showing differential N-HisPur $\beta$ -binding affinity were systematically and globally fit to equations describing various two-site models as illustrated in Table 5.2. (equations 5.20-5.26). Blue symbols represent N-HisPur $\beta$  binding to the 3' site; red symbols represent binding to the 5' site (*mean  $\pm$  standard deviation*). Each point represents the *mean  $\pm$  s.d.* of five independent experiments. Lines represent best fit isotherms. Residual analysis and fit statistics verify that N-HisPur $\beta$  binds to nonidentical sites of the MCAT enhancer element via a 3' to 5' cooperative mechanism.

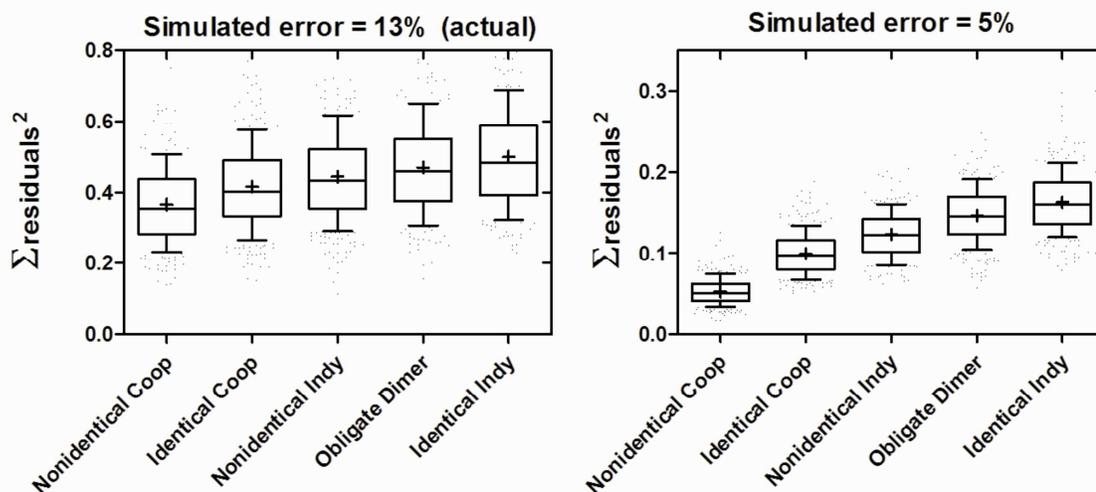
cooperative mechanism. Examination of the sequence represented by PE32-F shows no sequence redundancy, also suggesting that binding sites within PE32-F are nonidentical.

The resolved binding parameters obtained from fitting of individual site isotherms to a non-identical interacting model are as follows (67% confidence intervals are noted in parentheses): change in free energy for binding to the 3' site ( $\Delta G_1$ ) is -12.82 (-12.91 to -12.70) kcal/mole, change in free energy for binding to the 5' site ( $\Delta G_2$ ) is -11.97 (-12.24 to -11.45) kcal/mole, and the change in cooperative free energy ( $\Delta G_c$ ) is -1.457 (-1.768 to -0.748) kcal/mole. These values indicate that despite the high affinity of binding of N-HisPur $\beta$  to individual sites, intersite cooperativity is comparable to values reported for other transcription factors that adopt similar mechanisms (1, 22, 51, 113, 114, 167, 251). The broad confidence intervals obtained for  $\Delta G_2$ , and  $\Delta G_c$  are a consequence of experimental error and parameter cross correlation and are addressed below.

Despite the experimental power afforded by quantitative footprinting techniques, a major shortcoming is the low level of precision that is attainable. Typical footprint titrations yield precision in the  $\pm 10\%$  range (21-23, 113). This trend appears to be amplified in experiments utilizing ssDNA templates, as performed here, and settle at around  $\pm 13\%$ . The reasons for this are unknown, but likely correspond to differences in preference of the nuclease for single-stranded and double-stranded substrates. Due to the error level obtained, we opted to assess our confidence in our ability to discriminate between the possible binding models. To do this, we used a Monte Carlo simulation approach to test the effects of randomly introduced (Gaussian distributed) error at a level of  $\pm 13\%$  to a set of isotherms describing a non-identical interacting model, with binding

parameters identical to those obtained with our actual data, on the goodness of global fits of the resulting isotherms. This process was performed with 1000 iterations, and the goodness of fit to each model was judged by the sum of the square of the residuals of fits for each iteration. A box-and-whisker representation of this analysis is shown in Figure 5.7. Monte Carlo error simulation shows that accurate model estimation for a non-identical interacting system is possible and likely when the possessed binding energetics are similar to those observed here, based on significant differences in the mean and median values of the sum of the residuals squared for each model. However, estimation of the incorrect model is also very possible, a conclusion based on the considerable overlap in the 25-75<sup>th</sup> percentile boxes for all of the models. It can be seen that complete resolution of these binding models requires a simulated error as low as  $\pm 5\%$ , a level that may not be attainable using ssDNA templates.

High degrees of error also introduce higher levels of uncertainty in fit parameters. As mentioned above, the parameter estimates from global fitting of individual site isotherms carry with them very broad confidence intervals, especially in parameters that are cross correlated by virtue of the mathematical expressions from which they are obtained. Estimate distributions for  $\Delta G_1$ ,  $\Delta G_2$ , and  $\Delta G_c$  in the context of the non-identical interacting site model were again determined by Monte Carlo simulating the observed level of error ( $\pm 13\%$ ), and observing the returned fit parameters for each of 1000 iterations. The distributions of the returned parameters are depicted as histograms in Figure 5.8. The cross-correlation between  $\Delta G_2$  and  $\Delta G_c$  can be seen by the mirror image, biphasic distributions of the two parameters, as well their sheer broadness when error is

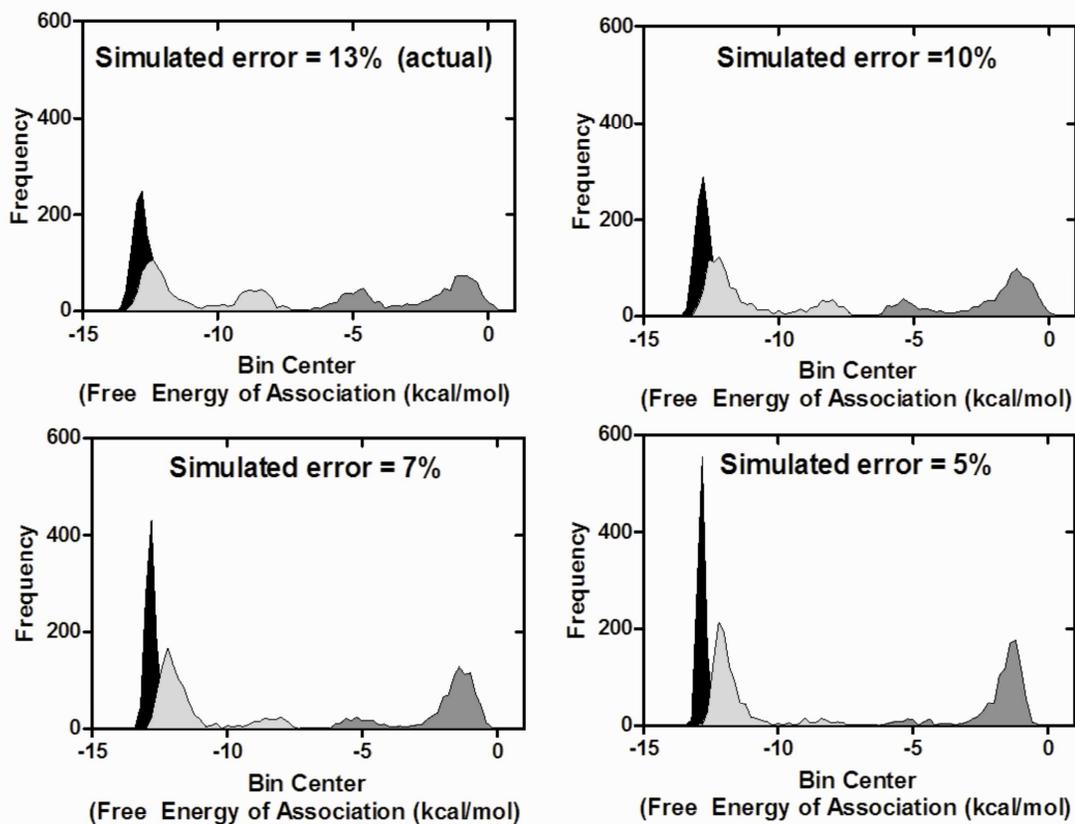


**Figure 5.7. Monte Carlo error simulations to assess model confidence.** Reiterative error simulations (1000 iterations) were performed on individual site isotherms shown in Figure 5.6 by Monte Carlo methods to yield error-incorporated isotherms that were then globally fit to the various two site models as described in Figure 5.6. Box-and-whisker plots showing distributions of fitting statistics ( $\Sigma\text{residuals}^2$ ) for each model are shown when error is introduced at the level observed in experiments described herein ( $\pm 13\%$ , left panel). Reducing error to  $\pm 5\%$  leads to higher model confidence as indicated by resolution of box-and-whisker plots (right panel). Boxes represent 25-75 percentiles, whiskers represent 10-90 percentiles. Median is marked by line across the box, and mean is denoted as (+).

simulated at  $\pm 13\%$ . The biphasicity persists until error is reduced to  $\pm 5\%$ , however the width of the distributions is significant for all levels of introduced error. It has been shown previously that resolution of cross-correlated values requires the isolated determination of one of the parameters by implementation of reduced valency templates (1). Unfortunately, construction of templates with reduced numbers of binding sites requires extensive sequence knowledge of the binding site in question, and that binding to this site can be abolished by deletion or mutation of the template. The absolute sequence identity of a Pur protein binding site remains speculative at this time. Footprints obtained by nuclease digestion often overestimate the size of binding sites due to steric effects (reviewed in (108)), and thus do not provide enough resolution to unequivocally identify the N-HisPur $\beta$  nucleotide binding site observed here, except to say that one exists on the 3' and 5' ends of the PE32-F region. The matter of binding site characterization is under current investigation by our laboratory.

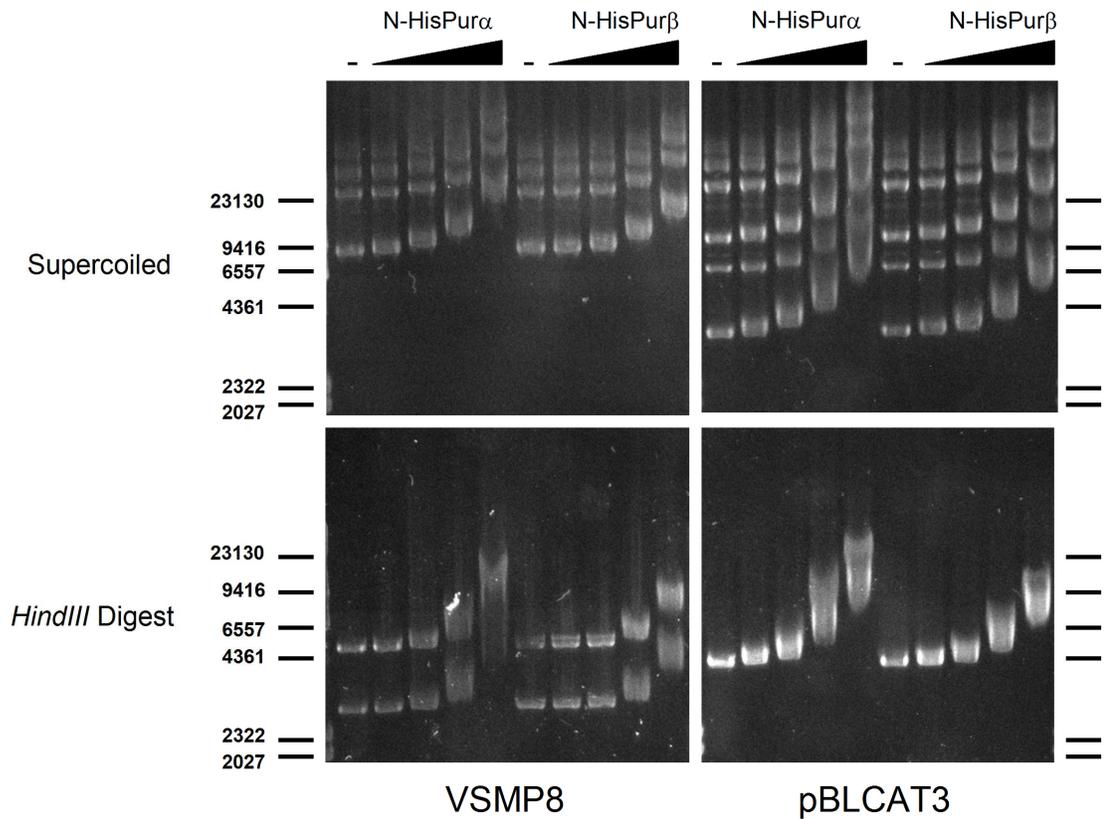
***Binding of N-HisPur $\beta$  to supercoiled DNA.*** The inability of purified recombinant Pur $\beta$  to bind 32 base-pair double-stranded representations of the SM $\alpha$ A proximal MCAT enhancer element (dsPE32-F) has been previously demonstrated (148, 149). Calculations of the free energy of strand separation (melting) of dsPE32-F by the nearest-neighbor method proposed by Breslauer (24), using thermodynamic values reported by SantaLucia, Jr. (239), suggest that the process requires 58.4 kcal/mole. This value is in accordance with experimental values obtained by our group (data not shown) using methods previously described (215). We have reported here that additive binding of N-HisPur $\beta$  to a sequence corresponding to PE32-F liberates ( $\Delta G_{Total} = \Delta G_1 + \Delta G_2 + \Delta G_c$ ) 26.25

kcal/mole (26.92 to 24.90 kcal/mole at 66.7% confidence interval). These findings suggest that occupation of the forward strand of the SM $\alpha$ A proximal MCAT enhancer sequence by Pur $\beta$  cannot proceed in a pure thermodynamic competition with the complementary reverse strand, but must require auxiliary factors that induce localized helix destabilization. Supercoiling requirements for binding of sequence-specific ssDNA-binding proteins have been demonstrated for FBP binding to FUSE (8, 153, 154, 192, 288), and recombinant glutathione-S-transferase fused (GST)-fused Pur $\beta$  binding to the *c-myc* promoter *PUR* element (310). We sought to examine the ability of N-HisPur $\beta$  to bind double-stranded sequences of DNA containing the proximal MCAT enhancer of SM $\alpha$ A in the context of a supercoiled versus relaxed environment, and compare these characteristics to those of N-HisPur $\alpha$ , for which dsDNA-binding properties have already been examined (310). To achieve this, we utilized an agarose gel-based mobility shift assay, and employed the full-length SM $\alpha$ A promoter:reporter construct VSMP8 in both a supercoiled and linearized (*HindIII* digested) configuration, to determine if torsional stress created by supercoiling creates localized destabilization of sequences, in turn making them accessible for binding by recombinant Pur proteins. For comparison, the parent-vector of VSMP8, termed pBLCAT3, was used to determine if binding events observed were specific for SM $\alpha$ A promoter elements. Figure 5.9 shows the results of this analysis. As can be seen in the upper-left panel of Figure 5.9, both N-HisPur $\beta$  and N-HisPur $\alpha$  bind to all supercoiled catamers of VSMP8, although N-HisPur $\alpha$  shows slightly greater affinity than does N-HisPur $\beta$  as judged by comparing the level of shifting



**Figure 5.8. Monte Carlo error simulations to assess converged parameter confidence.** Error-simulated individual site isotherms were globally fit to equations describing a non-identical, interacting two-site model (equations 5.25 and 5.26). Converged free energy parameters were plotted as frequency histograms from 1000 error simulations. Error levels of  $\pm 13\%$  (the level experimentally observed) led to broad distributions of parameters  $\Delta G_1$  (black),  $\Delta G_2$  (light grey), and  $\Delta G_c$  (dark grey), indicating low levels of value confidence. Systematically reducing the error from  $\pm 13\%$  to  $\pm 5\%$  restricts the distribution of  $\Delta G_1$ , but not  $\Delta G_2$  or  $\Delta G_c$ . The mirror-image distributions of  $\Delta G_2$  and  $\Delta G_c$  at all tested error levels are indicative of parameter cross-correlation.

observed in corresponding lanes containing the same amount of each protein. Binding of N-HisPur $\beta$  and N-HisPur $\alpha$  to supercoiled DNA was not dependent upon SM $\alpha$ A promoter components, as, both proteins bound to catamers of pBLCAT3 (Figure 5.9, upper-right panel) with similar affinity as VSMP8. A similar trend was observed for GST-Pur $\alpha$  binding to pUC19 with affinities similar to a pUC19-derived construct containing the *c-myc PUR*-element (310). To test if binding was supercoiling-dependent, we performed the same binding analysis with *Hind*III-digested plasmids. Digestion of VSMP8 yields two observable fragments, a ~4.9 kbp fragment containing mostly pBLCAT3-derived sequence plus ~800 bp of SM $\alpha$ A intron 1 sequence, and a ~2.9 kbp fragment containing ~1kb of upstream elements of the SM $\alpha$ A gene promoter, as well as ~1.9 kb of exon 1 and a portion of intron 1. As with supercoiled dsDNA, both recombinant Pur proteins bound SM $\alpha$ A promoter-derived sequences and parent vector-derived sequences with affinities similar to what was observed for supercoiled sequences (Figure 5.9, lower panels), albeit with a much lower affinity than what was observed for ssDNA sequences.



**Figure 5.9.** N-HisPur $\alpha$  and N-HisPur $\beta$  bind supercoiled and linearized plasmids VSMP8 and pBLCAT3. Agarose gel EMSAs were performed on equilibrium binding reactions containing titrating concentrations of N-HisPur $\alpha$  or N-HisPur $\beta$  in the presence of supercoiled or *HindIII*-linearized VSMP8 or pBLCAT3. Recombinant Pur proteins were titrated (25, 50, 100, 200 nM) in the presence of 1 nM plasmid DNA (2.84 and 5.21 ng/mL for pBLCAT3 and VSMP8, respectively). Complexes of bound DNA corresponding to 30 ng of total DNA were separated by electrophoresis as described in the methods section. The lengths of DNA markers (in bps) are indicated on the left.

## DISCUSSION

The modulation of gene transcription by duplex-to-single-strand interconversions represents an emerging regulatory scheme for eukaryotic gene transcription, as it permits exclusive binding of distinct sets of *trans*-acting protein factors to particular conformations that in turn regulate expression in a positive or negative fashion. In substantiation of this notion, the involvement of sequence-specific ssDNA-binding proteins has been identified to be of vital importance for the coordinated transcription of genes involved in a wide array of biological functions, including tissue growth and development, immune response, cell cycle progression, as well as cancer development, as described in the introduction.

In the present study, we have focused on examining the mechanism of sequence-specific ssDNA binding by Pur $\beta$ , a transcriptional regulator of, most notably, muscle-specific genes in mammals, such as SM $\alpha$ A and  $\alpha$ MHC (48, 82, 103, 143, 275). A culmination of promoter deletion, gain of function, and loss of function analyses has identified binding of Pur $\beta$  to the purine-rich strand of the SM $\alpha$ A proximal MCAT enhancer element as a critical step in repression of SM $\alpha$ A expression (148). Hence, a complete consideration of the enhancer element-binding mechanism employed by Pur $\beta$  will aid our understanding of how this apparently sequence-specific ssDNA-binding protein contributes to SMC transdifferentiation. Recent studies by our lab have indicated

that Pur $\beta$  self-associates with a dimerization constant of approximately 1  $\mu$ M (222), leading us to hypothesize that dimerization, representing a regulatory mechanism, is a prerequisite to ssDNA-binding.

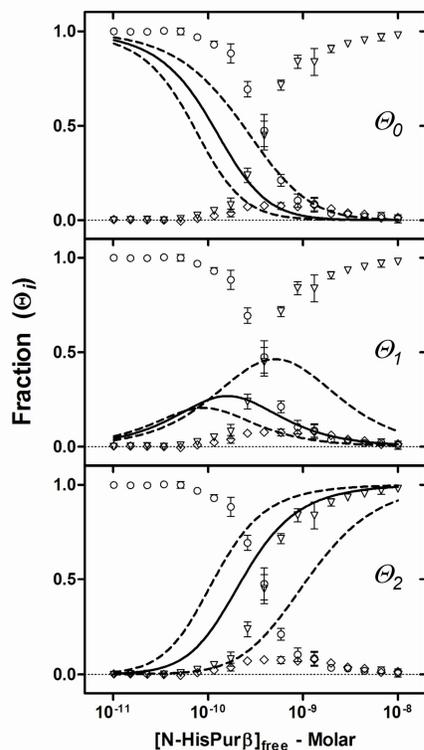
To test our hypothesis, we sought first to determine the binding stoichiometry of a physiologically relevant high affinity nucleoprotein complex formed between recombinant Pur $\beta$  (N-HisPur $\beta$ ) and a 32mer oligonucleotide representative of the purine rich strand of the SM $\alpha$ A proximal MCAT enhancer element (PE32-F) that has exhibited high affinity binding to Pur $\beta$  (28, 141, 144, 148, 275). In line with our hypothesis, we measured the stoichiometry of a N-HisPur $\beta$ :PE32-F high affinity nucleoprotein complex to be 2:1 (Figure 5.3). Furthermore, this complex was deemed to be physiologically relevant based on mobility comparisons with nuclear extract-derived Pur protein complexes (Figure 5.2). Previous reports investigating nucleic acid-binding properties of Pur $\alpha$  have found similarly that high affinity ribo- and deoxyribonucleoprotein complexes containing recombinant forms of Pur $\alpha$  adopt stoichiometries greater than 1:1 (91). However, studies investigating mechanisms of GST-Pur $\alpha$  and GST-Pur $\beta$  binding to a *c-myc* promoter *PUR*-element derived ssDNA oligonucleotide (24mer, termed MF0677) have indicated that both recombinant Pur proteins bind this element with a 1:1 stoichiometry (310). It is currently unknown to what extent nucleotide sequence and lattice length dictate Pur protein nucleoprotein complex stoichiometry, or if fusion domains interfere with facilitated assembly of higher order complexes.

Independent experiments aimed at delineating the mechanism of PE32-F binding by N-HisPur $\beta$  indicated that nucleoprotein assembly proceeds via a sequential binding

mechanism, and that binding is cooperative. These findings contradict our hypothesis which stated that ssDNA binding was preceded by homodimer formation. Quantitative EMSA binding experiments (Figure 5.4D) revealed that half-saturation of PE32-F by N-HisPur $\beta$  occurs at monomer concentrations of approximately 0.3 nM. Accordingly, nucleic acid-free self-association (based on a  $K_d$  of  $\sim 1$   $\mu$ M) at this monomer concentration would give rise to dimer concentrations of approximately 75 fM. This level of affinity would be quite high for any reversibly associating system. Close inspection of N-HisPur $\beta$  EMSA titration data shows not only the existence of an intermediary complex at concentrations near the half-saturation point, also an obligate dimer binding model, but that the species distribution patterns (Figure 5.4B) closely follow predicted distributions of a cooperative system (27). A statistical mechanics approach, assuming a two-site system (based on the determined 2:1 stoichiometry) also detected the presence of positive inter-site cooperativity (Figure 5.4C). Resolution of thermodynamic binding parameters was impossible by this method due to the fact that prior knowledge of the nature of the bindings sites was absent, except to say that cooperative free energy ( $\Delta G_c$ ) was negative (i.e.  $k_c > 1$ ).

To better resolve the binding energetics of N-HisPur $\beta$  for the SM $\alpha$ A proximal MCAT enhancer element, we used quantitative DNase I footprinting of a 382 base fragment of the forward strand of the SM $\alpha$ A promoter. The high utility of this assay comes from its ability to singly measure fractional occupation of a particular binding site. As such we were able to distinguish between binding mechanisms with limited confidence, due primarily to the level of precision obtained with this type of analysis

(Figure 5.7). However, coupling of the footprinting approach to quantitative EMSA analyses which detected positive cooperativity of binding, and analysis of the proximal MCAT enhancer element which shows no sequence redundancy, collectively points to a mechanism in which N-HisPur $\beta$  binds to the indicated sequence via a cooperative mechanism in which the sites are non-identical. As with model confidence, resolved parameters show very broad value constraints which are dictated, again, by low experimental precision, as well as parameter cross-correlation (Figure 5.8). This latter issue persists in all multisite cooperative systems, and is typically dealt with by implementation of DNA templates with deleted sites, so as to remove uncertainty caused by cooperativity (22, 23, 250). Applying values of footprinting resolved parameters, assuming their accuracy, to equations 5.12-5.14 shows that these values only marginally define the species-specific saturation curves obtained from quantitative EMSA (shown in Figure 5.10). Reasons for these disparities are not completely clear, but may be a reflection of experimentally induced disruption of equilibrium during performance of mobility shift assays; specifically dissociation of nucleoprotein complexes upon loading of equilibrated reaction in to gel wells or during electrophoresis. This effect would manifest most drastically at concentrations below and near the half-saturation value. Other possibilities for differences in resolved parameters would be due to increased affinity for interacting multisite templates compared to that of the isolated sites in PE32-F. Affinity of Pur $\alpha$  for long purine-rich ssDNA sequences, such as the bovine papilloma virus type I origin (100mer) has been measured on the order of  $1 \times 10^{-10}$  M, despite knowledge of the complex stoichiometry (136). Similarly, poor constraints on



**Figure 5.10. Comparison of quantitative footprinting-resolved parameters to individual quantitative electrophoretic mobility shift species isotherms.** Substitution of values of  $k_1$ ,  $k_2$ , and  $k_c$ , as determined by quantitative footprinting, into equations 5.12-5.14 describing fractional distribution of electrophoretic species  $\Theta_0$ ,  $\Theta_1$ , and  $\Theta_2$  are shown with data from Figure 5.4,C. Solid lines represent best fit parameters; dashed lines represent 67% confidence limits of best fit parameters.

microscopic values of  $k_2$ , and  $k_c$  have repercussions in the inference of macroscopic constant  $K_1$  and  $K_2$ . As is shown in Figure 5.8, these parameters have very wide estimation distributions. Furthermore, it has been shown by Brenowitz and colleagues (22) that without microscopic values determined by way of reduced valency promoters, resolved cooperative free energy terms represent a lower limit to the actual cooperative free energy of the system, and the greater the cooperative free energy that exists in a system, the more difficult it is to resolve individual site interaction free energies. These trends underlie the need to define the nucleotide binding site determinants for Pur proteins such that mutant promoters can be constructed.

Cooperative binding mechanisms are common for ssDNA-binding proteins, particularly those involved in DNA replication, recombination, and RNA transcription, such as T4 bacteriophage gp32 (4), *E.coli* SSB protein (236), and the Adenovirus DNA-binding protein (198, 323). By virtue of their abilities to cooperatively assemble nucleoprotein filaments, these proteins thus possess helix destabilizing activities necessary for their cellular functions. Cooperativity for sequence-specific ssDNA binding proteins, on the other hand, is minimally described in the literature at this point. Reasons for this are unclear, but may be due to the relative rarity of sequence-specific ssDNA-binding proteins, or the limited number of thermodynamically rigorous assays able to resolve site-specific binding isotherms and detect cooperativity. Binding of the yeast telomere protection protein Pot1 has been ascribed a cooperative mechanism, and similarly to what we have determined for N-HisPur $\beta$  reported here, sequential monomer binding of Pot1 to specific telomeric sites proceeds in a 3' to 5' direction (167).

Functional cooperative binding of Pot1 to yeast telomeres has been deemed necessary for telomeric nucleoprotein filament assembly and, in turn, protection of the chromosome ends from cellular nuclease damage. Cooperative binding of genomic promoter ssDNA sequences by Pur proteins may serve a similar helix destabilizing function, or to impart a potent functional response over a small change in cellular concentration.

Models of SM $\alpha$ A proximal MCAT enhancer repression have proposed that formation and maintenance of a single-stranded non-B-DNA conformation by Pur $\alpha$ , Pur $\beta$ , and MSY1 binding to nucleotides adjacent to the MCAT sequence causes disruption of a double-stranded TEF-1 binding site, and prevents binding of this transcriptional activator (28, 275). Enforcement of the repressive state must then require destabilization of the helix, either by thermodynamic favorability of ssDNA-binding protein occupation, or by environmental factors such as topological stress of the duplex DNA that permit occupation. To test this prediction, we examined the binding properties of N-HisPur $\beta$  to supercoiled and linear sequences of dsDNA (Figure 5.9). Unexpectedly, N-HisPur $\beta$  displayed binding to specific and non-specific sequences of supercoiled and linearized dsDNA. Clearly this type of assay cannot detect if binding is occurring at specific sites within the DNA, but it does suggest that, under these conditions, sequence selectivity of N-HisPur $\beta$  is low. Furthermore, the resolution of this assay does not permit us to differentiate between two possible binding modes for either N-HisPur $\beta$  or N-HisPur $\alpha$ ; dsDNA binding versus ssDNA binding. The lower level of affinity observed in these experiments may indicate that dsDNA binding is a secondary activity for Pur proteins, or that ssDNA-binding/helix destabilization is occurring with a lower level of observable

affinity due to a need to force the equilibrium in favor of ssDNA-binding by increasing protein concentration and activity, overcoming annealing favorability.

Previous studies cast doubt on the notion of dsDNA binding by Pur proteins, primarily reports by Darbinian and colleagues (56) which showed that GST-Pur $\alpha$  possesses strand displacement capabilities in the context of a short 15mer (26.7% guanine) oligonucleotide annealed to M13 single-stranded plasmids. Wortman and colleagues (310) showed that GST-Pur $\alpha$  is capable of displacing short (16mer) pyrimidine-rich strands of telomeric-repeats DNA, and that displacement occurs by virtue of contacts between the protein and the (24mer) purine-rich strand, and that binding of GST-Pur $\alpha$  increased potassium permanganate sensitivity to dsDNA, and permits binding of gp32, both suggestive of helix unwinding. Both groups localized strand displacement activity to the core DNA-binding domain of Pur $\alpha$ , which is highly homologous to Pur $\beta$  (143, 145), consisting of alternating basic-aromatic Class I and acidic leucine-rich Class II repeats. It is unknown whether or not strand displacement capabilities observed by these groups were reliant upon the ssDNA overhangs present in both sets of experiments, as strand displacement with blunt-ended fragments was not demonstrated. Wortman, et al (310) showed that binding of GST-Pur $\alpha$  to linear dsDNA required the C-terminal domain which consists of the psycho motif, as well as the glutamine- and glutamate-rich domains. Glutamine-rich domains have been historically implicated in transcriptional activation by transcription factors possessing them (97), however other studies have implicated glutamine-rich tracts in DNA distortion and helix unwinding activity in *Drosophila melanogaster* GAGA factor (GAF) (307). The means by which helix

unwinding activity is afforded by polyglutamine tracts is unknown, but may be attributable to the hydrogen-bonding capacity contributed by the high local concentrations of amides inherent to these peptides resulting in reduced melting temperatures of stretches of dsDNA in a manner similar to what is observed for formamide (14). Interestingly, human Pur $\beta$ , which is divergent from human Pur $\alpha$  in that it lacks the C-terminal glutamine-rich domain, failed to bind linearized pUC19 in experiments described by Wortman, et al. (310). Our results differ in this regard, as we have shown here that recombinant mouse Pur $\beta$  binds to linearized plasmid DNA (Figure 5.9, lower panels), despite also lacking a C-terminal glutamine-rich domain. Discrepancies in dsDNA-binding activity may arise from primary amino acid sequence differences in the N-terminus of human versus mouse Pur $\beta$ , namely the presence of two polyglycine tracts in mouse Pur $\beta$  that are absent in the human homolog (Figure 1.2). Polyglycine, and glycine-rich domains have been found in numerous proteins with observable helix-destabilization character including the UP1 subunit of heterogeneous ribonucleoprotein A1 (46, 86), and nucleolin (99). However, mouse Pur $\alpha$ , which also contains an N-terminal polyglycine tract (different in length and position compared to mouse Pur $\beta$ ), loses considerable linear dsDNA-binding activity when the C-terminal domain containing the psycho motif, glutamine-rich, and glutamate-rich domains is deleted. Thus terminal domains of mouse Pur $\alpha$  and Pur $\beta$  may direct helix-destabilizing properties differently.

The question remains how Pur proteins bind to ssDNA sequences in a double-stranded environment. Ourselves and others before us have proposed that binding of

sequence-specific ssDNA-binding proteins to dsDNA is facilitated by either pre-formation of single-stranded structures or a reduction in the annealing free energy of a localized region that make thermodynamic competition for site occupation feasible. This theory has been substantiated previously for FBP, a transcriptional activator, which requires transcription-induced negative supercoiling and unwinding in FUSE for FBP binding (153, 154). Models of repression of MCAT enhancer element-dependent transcription of SM $\alpha$ A suggest that this mechanism may not be feasible for Pur protein occupation, as transcription is presumably silent in the Pur protein occupied state, and thus so is transcription-induced negative supercoiling. Moreover, binding of recombinant Pur proteins to relaxed linear dsDNA supports this argument. If binding of ssDNA in the context of a dsDNA lattice by N-HisPur $\beta$  proceeds via a thermodynamic competition (bubble formation) target sequences would have to possess less than ~26 kcal/mole of annealing free energy. Average base-pairing free energy for a single base-pair in dsDNA of infinite length is on the order of -1.8 kcal/mole at 20°C (239). Thus, average sequences of approximately 14 basepairs might be prone to strand displacement by direct thermodynamic competition with N-HisPur $\beta$ . Localized melting of subdomains in large linear DNA fragments have been detected by microscopy techniques at temperatures as low as 64°C (234, 262) compared to a measured melting temperature of ~83°C for dsPE32-F (data not shown). Unfortunately, nearest-neighbor melting temperature predictions would estimate the melting temperature of a dsDNA of infinite length to also be infinite. Hence it is thought that localized melting is due to DNA subdomain architecture. Depressed melting temperatures in linear dsDNA are routinely believed to

be due to richness in A/T basepairs (187); however size and relaxed topology has also been shown to have inverse effects on diffusion coefficients of dsDNA (230). Whether or not limited Brownian motion (degrees of translational, rotational, and vibrational freedom) as a consequence of extended topology dictates localized melting temperature depression of linear dsDNA is not known, but transient base unpairing, or breathing, within localized regions may be a means of escaping this entropic limitation. DNA breathing has been detected at temperatures well below the melting temperature of an oligonucleotide (102, 156), but the transient nature of these occurrences make opportunistic binding by Pur proteins to opened duplexes unlikely, as binding is stably observed despite being reversible. Presumed non-specific binding of Pur proteins to ssDNA with varying degrees of affinity has been documented previously (56, 136, 145, 148, 310). These possibilities have the potential to explain non-specific binding of N-HisPur $\beta$  to linear dsDNA as observed here (Figure 5.9, lower panels). Nevertheless, the nature of dsDNA binding by Pur proteins needs to be investigated, most suitably by nuclease and chemical footprinting techniques that are sensitive to ssDNA formation.

As mentioned previously, repression of MCAT enhancer dependent expression of SM $\alpha$ A relies on the ssDNA-binding activities of Pur $\alpha$ , Pur $\beta$ , and MSY1, as well as a network of protein-protein interactions between the three factors (28, 141, 148). Recessed, pyrimidine strand displacement has been described for Pur $\alpha$  (56, 310), whereas MSY1 (YB-1) has exhibited strand displacement of blunt-ended, Y-box-containing, short double-stranded oligonucleotides, cisplatin-modified double-stranded oligonucleotides, as well as engineered fork and bubble structures (95, 123). Strand

separation activity for YB-1, although independent of ATP, has been shown to be elevated in the presence of ATP. Furthermore, self-association becomes limited from a possible monomer-trimer-hexamer-dodecamer equilibrium to that of primarily monomer-dimer upon addition of ATP (95). These results suggest that ATP has allosteric effects on MSY1 (YB-1) activity and protein-protein interactions. Furthermore, sequence specificity of N-HisPur $\beta$  has been shown to be modulated by MSY1 (145). Hence, ssDNA-recognition, binding activity, strand separation, ATP hydrolysis and transcriptional repression achieved by Pur $\alpha$ , Pur $\beta$ , and MSY1 *in vivo* may be the result of collaborative and cooperative activities and interactions at the MCAT enhancer element of the SM $\alpha$ A promoter that have yet to be meticulously defined, and will likely require high-resolution biophysical and structural studies to do so. Worth noting is that collaboration of Pur $\alpha$  and MSY1 has been found at gene promoters other than SM $\alpha$ A (36, 37, 238).

In conclusion, we have found that recombinant, purified Pur $\beta$  binds to the purine-rich strand of the SM $\alpha$ A proximal MCAT enhancer element in a 3' to 5' sequential and cooperative manner, with a stoichiometry of 2:1. Apparent affinity (half-saturation) is on the order of 0.3 nM, while rigorous thermodynamic interrogation has shown that free energy of binding is -26.25 kcal/mole. This free energy value suggests that Pur $\beta$  is only able to thermodynamically compete for binding to short stretches of oligonucleotides if at all. The involvement of co-repressors Pur $\alpha$  and MSY1 in maintaining repressive single-stranded conformations of the SM $\alpha$ A MCAT enhancer element are the goals of future studies, as well as determining specific nucleotide binding determinants in this element,

as well as other Pur protein-responsive sequences. The intended outcome of these efforts is a better understanding of how transcriptional repression is enforced by these factors.

# **CHAPTER VI. TOWARDS THE CHARACTERIZATION OF THE MINIMAL SINGLE-STRANDED DNA BINDING SITE OF PUR $\beta$ WITHIN THE SM $\alpha$ A PROXIMAL MCAT ENHANCER ELEMENT**

*The following is a description of original and unpublished work.*

## **INTRODUCTION**

The identification of *trans*-acting regulatory proteins involved in the regulation of gene expression has facilitated the elucidation of gene regulatory mechanisms in both prokaryotes and eukaryotes. The involvement of specific proteins in the spatial and temporal regulation of multiple genes has suggested that subsets of tissue-specific genes may utilize similar combinations of *trans*-, and *cis*-acting regulatory elements of transcriptional control to ensure proper timing and levels of expression. The ability to identify additional gene targets of transcription factors that enact tissue-specific or choreographed transcriptional regulation strengthens our capacity to understand developmental process and disease progression by better describing gene regulatory networks and identifying additional target genes. This ability relies, however, on clear definitions of transcription factor binding site in terms of nucleotide sequences.

Transcription factor binding site definitions are typically represented by consensus sequences; nucleotide sequences that signify the statistical predominance of nucleotides at given positions within the binding site. The process of defining a nucleic acid-binding

protein consensus sequence requires the accumulation of numerous pieces of information, most importantly the binding site size or footprint, and the sequences of all demonstrated targets of binding. Stringent nucleotide preferences of transcription factors furnish robust consensus sequences that are very useful for identifying possible interaction targets, whereas diverse sequence tolerance and promiscuous binding can render derived consensus sequences as inaccurate depictions of nucleotide preferences whose utilization can be misleading. Inaccuracy of consensus sequences has led to the use of more sophisticated algorithms of defining and visualizing sequence preference profiles for transcription factors. Sequence logos have circumvented problems associated with nucleic acid binding proteins exhibiting relatively indiscriminate binding properties for which consensus sequences are inappropriate (245, 246). Nevertheless, transcription factor binding site leniency remains problematic for defining nucleotide preferences and identifying possible binding sites.

The goal of the work presented in this chapter was to characterize the minimal ssDNA binding site of the sequence specific SSB Pur $\beta$ . Pur $\beta$  is one of three known sequence-specific SSB transcription factors that repress expression of SM $\alpha$ A, an important cytoskeletal contractile protein whose expression is important for cellular contractile functions (243) and for the differentiation status of numerous cell types (127, 241). The abundance of SM $\alpha$ A protein levels in differentiated SMCs versus SMC undergoing phenotypic modulation towards fibroblastic cell character has made SM $\alpha$ A a hallmark of SMC differentiation (210, 315). Downregulation of SM $\alpha$ A has been implicated as a contributing factor towards dysfunctional vessel wall remodeling (175)

and the vulnerability of atheroma in relation to plaque rupture (88, 151).  $SM\alpha A$  promoter deletion analysis has identified a cryptic enhancer element in the 5' region of the promoter (28, 48, 82, 270, 275) that possesses a high degree of purine/pyrimidine asymmetry and exhibits transient structural interconversions in response to stimulus that activates  $SM\alpha A$  expression (9). This element contains a core MCAT motif shown to bind TEF-1 when in double-stranded configurations which is necessary for gene activation. Binding sites for sequence-specific SSBs Pur $\beta$ , Pur protein family member, Pur $\alpha$ , and Y-box protein MSY1 have been detected on opposing strands of this element, with Pur $\alpha$  and/or Pur $\beta$  occupying the purine-rich strand and MSY1 occupying the pyrimidine-rich strand (28). Recent gain-and-loss of function studies (148) and chromatin immunoprecipitation techniques (149) have pointed to Pur $\beta$  as the dominant Pur protein repressor in the context of  $SM\alpha A$  gene regulation in certain cell types.

The critical nature of Pur $\beta$  functions in regard to  $SM\alpha A$  repression and its phenotypic consequences have yielded a need to identify other targets of Pur $\beta$  repression. Both Pur $\alpha$  and Pur $\beta$  have been implicated in the regulation of numerous genes, both at the transcriptional and translational levels, and in the case of Pur $\alpha$ , the regulation of cellular and viral DNA replication and cell cycle progression (92, 132). Diverse functions in pathological blood vessel remodeling, cancer, and viral pathogenesis has revealed the need for a full description of Pur protein regulatory objectives. Projection of possible genomic and transcriptomic targets of Pur $\beta$  has been hampered by the lack of a defined consensus sequence for this multifunctional protein.

In the present study we have examined the *in vitro* binding characteristics of recombinant mouse Pur $\beta$  to oligonucleotides representative of the SM $\alpha$ A MCAT enhancer element to map the binding sites within this region. Recent reports by our lab have detected two binding sites within the region encompassing nucleotides -195 to -164 of the SM $\alpha$ A promoter in relation to the transcription start site with low resolution (Chapter V). Here we show that these two sites map to nucleotides -195 to -190 (5' site) and -171 to -166 (3' site) of the SM $\alpha$ A promoter. Furthermore, these sites resemble consensus sequences reported previously for Pur $\alpha$  which is 5'-GGGAGA-3' (10, 11, 310), with only slight degeneracy at each site.

## **MATERIALS AND METHODS**

***Chemicals, protein reagents, and oligonucleotide probes.*** All chemicals used in this study were of reagent grade or better. Recombinant Pur $\beta$  was expressed as an amino-terminally labeled hexahistidine tag fusion proteins (referred to in this report as N-HisPur $\beta$ ), purified from *E. coli* expression cultures, and quantified as described previously in Chapter IV and in a previous publication (222). Preparation and validation of the epitope-specific rabbit polyclonal antibody against mouse Pur $\beta$  (anti-Pur $\beta$  302) has been described previously (141). Synthetic oligonucleotides were purchased from Sigma-Genosys.

***Competitive ssDNA-binding assay.*** To monitor the extent of recombinant Pur $\beta$  binding to PE32-F in the presence of various oligonucleotide competitors, a

discontinuous solid-phase DNA-binding assay was performed as previously described with some minor modifications (145). Biotinylated PE32-F was immobilized on streptavidin-coated microtiter wells (Streptawells™, Roche) by application of 100 µL of 1 nM 3'-biotinylated PE32-F (PE32-bF) in buffer consisting of 25 mM tris(hydroxymethyl)aminomethane (Tris)-HCl pH 7.5, 150 mM NaCl, 1.5 mM MgCl<sub>2</sub> at 20 ± 1°C for 1 h with moderate shaking. Solutions were removed and wells were washed three times with 300 µL wash buffer consisting of 25 mM Tris-HCl pH 7.5, 150 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 0.05% (v/v) Tween 20. Blocking of non-specific binding was accomplished by incubation of 250 µL of blocking buffer consisting of 25 mM Tris-HCl pH 7.5, 150 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 20 µg/ml bovine serum albumin (BSA), at 20 ± 1°C for 1 h with moderate shaking. Wells were again washed three times with 300 µL of wash buffer. Competitive binding solutions consisting of 1 nM N-HisPurβ and titrated concentrations of competing oligonucleotides in buffer consisting of 50 mM Tris-HCl pH 7.5, 100 mM KCl, 0.5 mM dithiothreitol, 2 µg/ml polydeoxythymidine (dT<sub>32</sub>), 50 µg/ml BSA were added to the wells (100 µL/well) and incubated overnight at 4°C. The sequences of competing oligonucleotides are listed in Table 6.1. The next morning wells were again washed as before. Quantities of N-HisPurβ nucleoprotein complexes remaining after competition were detected by addition of 100 µL of primary antibody solution containing 1 µg/ml anti-Purβ 302 polyclonal antibody in binding buffer consisting of 25 mM Tris-HCl pH 7.5, 150 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 0.05% Tween 20, 2 µg/ml BSA, for 1 h at 20 ± 1°C. This was followed by triplicate washing and addition

of 100  $\mu\text{L}$  of secondary antibody solutions containing goat-anti-rabbit horseradish peroxidase conjugate (ExtraAvidin-HRP, Sigma) diluted 1:10,000 in binding buffer to each well and incubation for 1 h at  $20 \pm 1^\circ\text{C}$ . Wells were washed three times and 100  $\mu\text{l}$  of 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) substrate solution (ABTS, Chemicon) was added. After satisfactory color development by incubation at room temperature (for approximately 5 min), 100  $\mu\text{L}$  of 1% (w/v) sodium dodecylsulfate was added to stop the reactions. Solution absorbance readings at 405 nm were obtained with a microplate reader. Self-competitor controls (PE32-F) were used to verify complete competition absorbance values ( $A_{min}$ ), whereas wells with no competitor were used to obtain maximum absorbance values ( $A_{max}$ ). These controls were included on each plate to permit normalization of absorbance values necessary for comparison of results from multiple plates. Normalized absorbance values ( $A_{Norm}$ ) were calculated using the expression  $A_{Norm, i} = (A_i - A_{min}) / (A_{max} - A_{min})$ , where  $A_i$  is the absorbance of well  $i$ . Determinations of competitor concentrations necessary for 50% inhibition of complex formation,  $IC_{50}$ , were performed by nonlinear least-squares fitting to the following expression:

$$A_{Norm} = \frac{1}{1 + 10^{(\text{Log}IC_{50} - \text{Log}[\text{Competitor}])\alpha_H}} \quad (\text{Equation 6.1})$$

where  $\alpha_H$  is the Hill coefficient which permits variability of the slope of the transition. Non-linear least-squares fitting was performed using Prism 5 software (GraphPad Software, Inc., San Diego, CA).

**Table 6.1. Oligonucleotides used in this study as fluid-phase competitors in Pur $\beta$  ssDNA-binding functional ELISA.**

	Fluid Phase Competitor <sup>a</sup>	Sequence (5' – 3')
<i>Truncation Series</i>	PE32-F (-195/-164) <sup>b</sup>	GGGAGCAGAACAGAGGAATGCAGTGGGAAGAGA
	-195/-166	GGGAGCAGAACAGAGGAATGCAGTGGGAAGA
	-195/-168	GGGAGCAGAACAGAGGAATGCAGTGGAA
	-195/-170	GGGAGCAGAACAGAGGAATGCAGTGG
	-195/-172	GGGAGCAGAACAGAGGAATGCAGT
	-195/-174	GGGAGCAGAACAGAGGAATGCA
	-195/-176	GGGAGCAGAACAGAGGAATG
	-195/-178	GGGAGCAGAACAGAGGAA
	-195/-180	GGGAGCAGAACAGAGG
	-195/-182	GGGAGCAGAACAGA
	-195/-184	GGGAGCAGAACA
	-195/-186	GGGAGCAGAA
	-195/-188	GGGAGCAG
	-195/-190	GGGAGC
	-193/-164	GAGCAGAACAGAGGAATGCAGTGGGAAGAGA
	-191/-164	GCAGAACAGAGGAATGCAGTGGGAAGAGA
	-189/-164	AGAACAGAGGAATGCAGTGGGAAGAGA
	-187/-164	AACAGAGGAATGCAGTGGGAAGAGA
	-185/-164	CAGAGGAATGCAGTGGGAAGAGA
	-183/-164	GAGGAATGCAGTGGGAAGAGA
	-181/-164	GGAATGCAGTGGGAAGAGA
	-179/-164	AATGCAGTGGGAAGAGA
	-177/-164	TGCAGTGGGAAGAGA
	-175/-164	CAGTGGGAAGAGA
	-173/-164	GTGGAAGAGA
-171/-164	GGAAGAGA	
-169/-164	AAGAGA	
<i>Scanning Series</i>	PE32-F (-195/-164) <sup>b</sup>	GGGAGCAGAACAGAGGAATGCAGTGGGAAGAGA
	-183/-164	GAGGAATGCAGTGGGAAGAGA
	-185/-166	CAGAGGAATGCAGTGGGAAGA
	-187/-168	AACAGAGGAATGCAGTGGAA
	-189/-170	AGAACAGAGGAATGCAGTGG
	-191/-172	GCAGAACAGAGGAATGCAGT
	-193/-174	GAGCAGAACAGAGGAATGCA
	-195/-176	GGGAGCAGAACAGAGGAATG

<sup>a</sup> Numbers reflect position relative to transcriptional start site of the SMAA promoter  
<sup>b</sup> Full length oligonucleotide probe used for self-competition control

## RESULTS

Investigations into genomic and RNA targets of Pur proteins have been impeded by the inability of researchers to definitively describe a consensus binding site for these proteins. The detection of Pur protein binding to a variety of ss/dsDNA and RNA sequences has only made this pursuit more difficult, due to the fact that Pur $\alpha$  and Pur $\beta$  display extraordinary sequence promiscuity. This apparent promiscuity likely arises from a combination of factors but may arise artifactually, however, due to a poor understanding of the nature of Pur protein nucleoprotein complexes with respect to stoichiometry and affinity, as it is envisioned that sequence dictates these parameters. Therefore, we sought to systematically examine the base specificity of Pur $\beta$  in the context of a system that has been thermodynamically predefined (described in Chapter V). Towards this end we have examined N-HisPur $\beta_2$ :PE32-F nucleoprotein complex stability in the presence of oligonucleotide competitors designed to systematically dissect nucleotide stability contributions. We have employed a discontinuous solid-phase DNA-binding functional enzyme-linked immunosorbent assay (ELISA) to address this goal, as this assay permits simultaneous testing of competitive oligonucleotides in a high-throughput fashion, and direct comparison of results to gauge the effects of each competitor.

To identify nucleotides that contribute to overall nucleoprotein complex stability, we designed three series of competitive oligonucleotides for use in our functional ELISA. Table 6.1 details the sequence identity of the competitors and the rationale of our design.

Briefly, we designed a series of binucleotide truncation mutants of PE32-F (-195 to -164 of the SM $\alpha$ A promoter) that have deletions proceeding from the 3' end (-195 series), and from the 5' end (-164 series). Recent studies have shown the importance of nucleotides near positions -195 to -192 and -171 to -164 of the SM $\alpha$ A promoter in both transfection-based reporter assays (28, 48) and direct ssDNA-binding assays (275). Therefore, truncations from either end of PE32-F (-195/-164) should interfere with the ability of the oligonucleotide to compete for N-HisPur $\beta$  binding. The results obtained by incorporation of these oligonucleotide competitors in functional ELISAs are shown in Figure 6.1. Panel A and B show the competition isotherms generated by this approach for the -195 series and -164 series, respectively. Nonlinear least-squares fitting of these isotherms to a phenomenological inhibition equation (Equation 6.1) yields  $IC_{50}$  values for comparing relative affinities for N-HisPur $\beta$ . We used self-competition (PE32-F) as a positive control and  $IC_{50}$  reference point for comparison. As shown in Figure 6.1, panel C, deletions from either the 5' or 3' end of PE32-F results in loss of affinity for N-HisPur $\beta$ , and the extent of truncation directly coincides with loss of affinity, as indicated by increasing  $IC_{50}$  values. Consistent with the notion that N-HisPur $\beta_2$ :PE32-F nucleoprotein complexes are stabilized by cooperative interactions between sites (Chapter V), deletion of either putative binding site results in a loss of the competitors ability to displace N-HisPur $\beta$  from the solid phase. Interestingly, deletions from the 5' end appear to be better tolerated than those from the 3' end. This contradicts previous findings that N-HisPur $\beta$  shows higher affinity for the 3' site of PE32-F than for the 5' site (Chapter V). However,

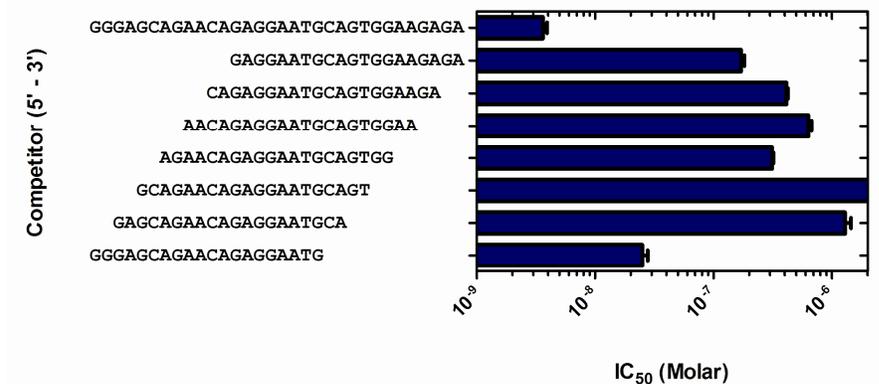


it may suggest that 5' nucleotides of the putative *PUR*-consensus (GGGAGA) are more critical to binding than are those in the 3' end.

Based on results shown in Figure 6.1, it was found that oligonucleotides with a minimum length of 18-20 nt consistently conferred the best competitive response in our functional ELISA. This finding, coupled with a need to confirm sequence specificity and discount possible length effects associated with observed  $IC_{50}$  profiles, led us to design a series of competitive oligonucleotides of consistent length (20 nt) that scan the entirety of PE32-F (Table 6.1). Using these oligonucleotides in the competitive functional ELISA provided the results shown in Figure 6.2. Similar to what was observed for the truncation series, oligonucleotides representing the ends of PE32-F competed the best for N-HisPur $\beta$  binding. This result suggests that the competitive effect we observed in Figure 6.1 is sequence-specific and is not purely dependent upon oligonucleotide length. This data also shows that the two binding sites are not equivalent. However, this data also contradicts previous findings which show the 3' site to possess greater inherent affinity for N-HisPur $\beta$  binding than the 5' site, as discussed in Chapter V.

## **DISCUSSION**

Binding of Pur proteins to nucleic acids has been shown to potentiate extensive cellular consequences ranging from cell cycle arrest to transcriptional and translational regulation. The diverse functional aspects of Pur proteins make them important molecular target candidates for the prevention of viral replication, cardiovascular disease and cancer (132). However, a fundamental understanding of mechanisms employed by



**Figure 6.2. Analysis of N-HisPurβ:PE32-F nucleotide interactions by scanning oligonucleotide competition.** Results of fluid-phase competitor titrations in a ELISA-based ssDNA-binding assay are shown for a series of 20 nt oligonucleotides that scan PE32-F. Competition isotherms (not shown) were fit to equation 6.1 to resolve  $IC_{50}$  values for each competitor, plotted as best fit value  $\pm$  67% confidence interval.  $IC_{50}$  value bars extending past the axis frame are indicative of very low affinity.

Pur proteins in regulating these processes have not been achieved, due in part to an inadequate description of Pur protein target sequences.

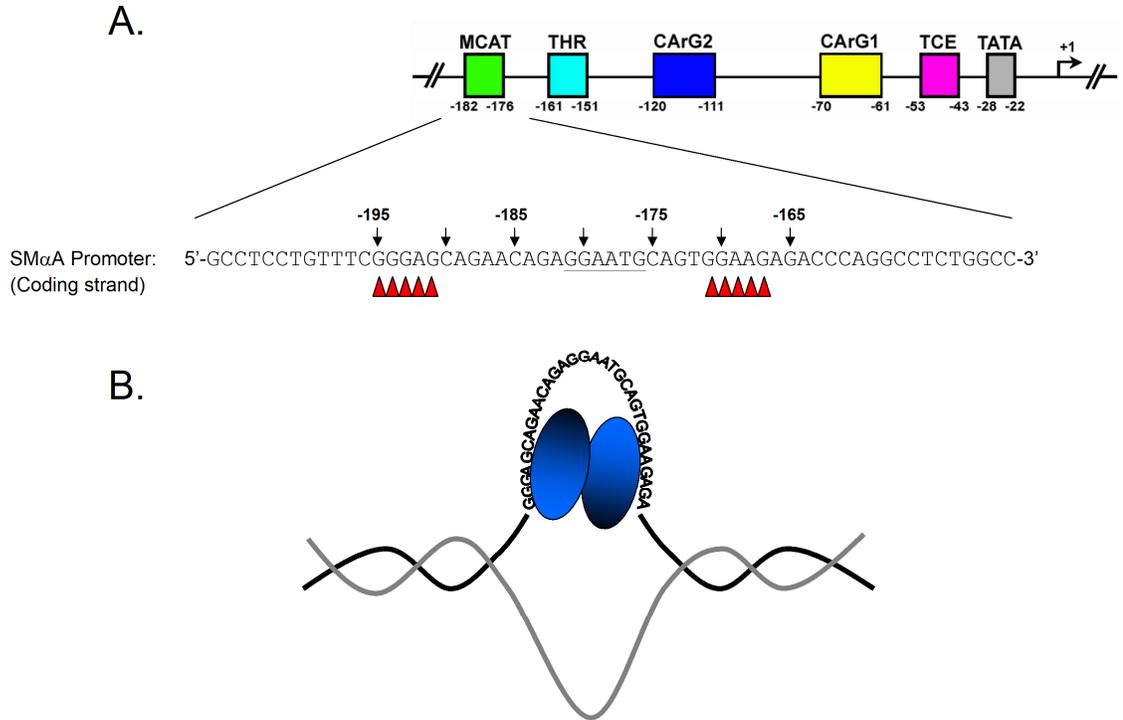
The diversity of nucleotide sequences that have been used to probe for Pur protein binding and involvement, as well as differences in interaction affinities associated with the use of these probes (Table 1.1) illustrates the difficulties associated with definitive binding site identification. Furthermore, a lack of nucleoprotein complex stoichiometries reported for these probes, with few exceptions, makes estimation of the number of binding sites in each sequence difficult to estimate. The possibility of multiple heterogeneous binding sites existing on a single probe obviously exists. Therefore a systematic approach to delineation of binding site identity of any proteins requires preexisting knowledge regarding complex stoichiometry, affinity, and/or footprint information. Few systematic approaches to identifying nucleotide sequence components critical for Pur proteins have been armed with any of this information. One such study, aimed at the identification of a Pur $\alpha$  consensus sequence made use of this information in regards to the c-myc-associated *PUR*-element (10, 11, 310). In this series of papers, Edward Johnson and colleagues systematically showed that Pur $\alpha$  binds to a core consensus sequence of GGGAGA with what is likely a 1:1 stoichiometry (reviewed in Chapter I).

Armed with equivalent thermodynamic data regarding the binding of recombinant N-HisPur $\beta$  to the proximal MCAT enhancer element of the SM $\alpha$ A promoter (Chapter V), we sought to test the consensus sequence prediction of Johnson and colleagues in regards to Pur $\beta$  nucleotide preferences, and to see if they differ from those of Pur $\alpha$ . Previous

reports have implicated nucleotides flanking the core MCAT sequence of the SM $\alpha$ A promoter as being necessary for Pur protein-mediated repression of reporter expression (28, 48) and for direct binding of Pur $\alpha$  and Pur $\beta$  to oligonucleotides representative of this region (28, 275). Figure 6.3 summarizes the results of these findings. In the present study, we used a convenient enzyme-linked immunosorbent ssDNA-binding assay to gauge the effects of competitive fluid-phase mutant oligonucleotides to on N-HisPur $\beta_2$ :PE32-F nucleoprotein complex assembly.

Our results indicate that nucleotides on either end of PE32-F are critical to nucleoprotein complex stability. This finding is based on the competition profiles of 5' and 3' truncation series of oligonucleotides harboring successive doublet deletions (Figure 6.1). The findings were further corroborated by employment of a series of scanning oligonucleotides, where the intention was to discount length effects that could possibly explain the competition results of the truncation series. Indeed, these oligonucleotides also suggest the involvement of terminal nucleotides in N-HisPur $\beta_2$ :PE32-F nucleoprotein complex stability (Figure 6.2). The presence of degenerate GGGAGA core consensus sequences in each end of PE32-F suggest that these regions indeed represent N-HisPur $\beta$  binding sites and the N-HisPur $\beta$  shares nucleotide preferences with those reported for Pur $\alpha$ .

Previous finding by our lab have shown regions of DNase I protection afforded by N-HisPur $\beta$  that correspond to the 5' and 3' regions of the MCAT enhancer element spanning from nucleotides -195 to -164 of the SM $\alpha$ A promoter when in a single-stranded conformation (Figure 5.5). Unfortunately, the resolution of this footprinting approach



**Figure 6.3. Recognition and binding of N-HisPur $\beta$  to the purine-rich strand of the SM $\alpha$ A cryptic MCAT enhancer element.** *A*, Region and sequence of the purine-rich strand of the cryptic MCAT enhancer of the SM $\alpha$ A promoter. Numbers indicate nucleotide positions relative to the transcriptional start site. Red arrows indicate positions and identities of nucleotides deemed important for stabilization of (N-HisPur $\beta$ )<sub>2</sub>:PE32-F nucleoprotein complexes shown here and in previous studies, and are indicative of two Pur protein bindings sites exhibiting slight degeneracy of the *PUR*-motif, GGGAGA. *B*, Hypothetical mechanism of maintained directional binding in a cooperatively assembled nucleoprotein complex. Looping of the purine-rich strand permits cooperative interactions between binding sites while maintaining 5' to 3' directionality of the binding site in relation to the interaction interface of N-HisPur $\beta$ .

did not allow for precise determination of ligating nucleotides within these sites. The results did indicate that N-HisPur $\beta$  binds to these sites in a successive 3' to 5' and cooperative manner suggesting that these sites are non-equivalent, a finding consistent with sequences analysis showing non-redundancy of these sites. Results from the present study also show that the 5' and 3' binding sites within PE32-F are non-equivalent with respect to N-HisPur $\beta$  binding; however our results also indicate that N-HisPur $\beta$  binds to the 5' site with greater affinity than to the 3' site, in contrast to results obtained by quantitative footprinting. It is possible that this discrepancy arises from isolation of the two binding sites by the use of the 32 nt oligonucleotide PE32-F. Quantitative footprinting experiments described in Chapter V were performed using 382 nt template representations of the SM $\alpha$ A promoter region containing other known sites of Pur protein interaction (47, 148, 272). It is possible binding of N-HisPur $\beta$  to the 5' and 3' sites of the region encompassing the MCAT enhancer element is thermodynamically linked to binding at other sites that may influence the interaction energetics observed in these experiments. N-HisPur $\beta$ -mediated nuclease protection at sites outside of nucleotides -195 to -164 indicates that this might be a possibility.

Identification of N-HisPur $\beta$  binding sites within the MCAT enhancer element of SM $\alpha$ A raises structural questions of how nucleoprotein assembly is accomplished. Numerous dimeric transcription factors typically display palindromic binding site sequences. For example bHLH dimers bind palindromic E-boxes (5'-NCANTGN-3') by monomer recognition of major groove C-G basepair constituents on successive half-turns of the B-DNA helix (202). The sequence represented by PE32-F contains direct repeats

of the putative N-HisPur $\beta$  binding sites. Therefore, the manner by which two monomers of N-HisPur $\beta$  cooperatively bind two direct repeats of the identified binding sequence represents a structural dilemma. Figure 6.3, panel B shows how binding might occur and suggests that looping of the intervening ssDNA sequences may transpire. This is purely speculation, however, and requires high resolution structural data for validation.

In conclusion, the results described herein provide direct information regarding the nucleotide contributions stabilizing nucleoprotein complexes formed between recombinant Pur $\beta$  and oligonucleotide representations of the proximal MCAT enhancer element of the SM $\alpha$ A gene promoter. These results provide insight into possible experimental methods of reducing Pur $\beta$  binding affinity in cultured cells and *in vivo* by mutating putative nucleotide contacts, as well as in the detection of additional Pur $\beta$  binding targets genome-wide. These approaches may not only aid in our understanding of mechanisms of SM $\alpha$ A repression involved in phenotypic reprogramming of VSMCs, but also unveil the extents of Pur protein function in physiological events and disease progression, and provide a means to prevent or manage clinical consequences of Pur protein-ssDNA interactions.

## CHAPTER VII. CONCLUDING REMARKS

The studies described in this dissertation have helped to characterize the mechanism implemented by Pur $\beta$  to assemble a repressive nucleoprotein component of SM $\alpha$ A transcriptional regulation as it might occur *in vivo* and in cell-culture models of myofibroblast transdifferentiation and VSMC dedifferentiation. Equally as important, these studies have also shown that Pur protein-mediated repression of SM $\alpha$ A transcription can serve as a model system for examining the cause and effect relationship between sequence-specific SSB transcription factors and structural interconversions in regulating gene expression.

Generally speaking, the results of these studies have shown that Pur $\beta$ , the dominant Pur protein repressor of SM $\alpha$ A transcription in MEFs, utilizes cooperative interactions to facilitate and stabilize nucleoprotein assembly at the proximal MCAT enhancer element of the SM $\alpha$ A promoter. It remains to be established whether or not this mechanism is accurate in the context of a repressive nucleoprotein entity, one that enacts repression of SM $\alpha$ A transcription *in vivo*, as the mechanism described herein was delineated with purified recombinant Pur $\beta$ , in the absence of co-repressors Pur $\alpha$  and MSY1. Loss-of-function studies described here, suggest that full-repression of SM $\alpha$ A expression in MEFs requires collaborative, synergistic activities of both Pur $\alpha$  and Pur $\beta$ . Furthermore, the inability of recombinant Pur $\beta$  to solely establish ssDNA-binding in the presence of short complementary strands (strand displacement) also suggests that Pur $\beta$  is an

opportunistic SSB transcription factor, and that co-repressor involvement might be necessary to destabilize base-pairing within the SM $\alpha$ A MCAT enhancer element, to allow for repressive nucleoprotein complex assembly.

Future efforts in continuation of the studies described in this dissertation would be aimed towards delineating the roles of Pur $\alpha$  and MSY1 in nucleoprotein complex assembly and helix destabilization as mentioned above. A combination of biophysical assessment of isolated nucleoprotein assemblies for Pur $\alpha$  and MSY1 with their respective ssDNA strands, and footprinting techniques involving dsDNA templates in the presence of Pur $\alpha$ , Pur $\beta$ , and MSY1 (in isolation and combination) would likely reveal the thermodynamic parameters stabilizing nucleoprotein assembly and detail how assembly occurs in the context of a double-stranded environment. Similarly, probing of the secondary structure of the SM $\alpha$ A promoter region with ssDNA-sensitive reagents, in the contexts of supercoiled and linearized plasmid dsDNA constructs, may help to explain if topological stress facilitates stable non-B-DNA structures that promote sequence-specific SSB recruitment and occupation, events crucial to negative regulation of SM $\alpha$ A expression. Collectively, these future studies would help deduce the cause and effect relationship between sequence-specific SSB repressors and promoter structural interconversion in the regulation of SM $\alpha$ A gene transcription and other genes that rely on ssDNA-binding transcription factors for regulated expression. In addition, they would provide a technical foundation for examining these systems.

## COMPREHENSIVE BIBLIOGRAPHY

1. Ackers GK, Johnson AD, Shea MA (1982) Quantitative model for gene regulation by lambda phage repressor. *Proc Natl Acad Sci U S A.*79(4):1129-33.
2. Aikawa M, Rabkin E, Voglic SJ, Shing H, Nagai R, Schoen FJ, Libby P (1998) Lipid lowering promotes accumulation of mature smooth muscle cells expressing smooth muscle myosin heavy chain isoforms in rabbit atheroma. *Circ Res.*83(10):1015-26.
3. Aikawa M, Sakomura Y, Ueda M, Kimura K, Manabe I, Ishiwata S, Komiyama N, Yamaguchi H, Yazaki Y, Nagai R (1997) Redifferentiation of smooth muscle cells after coronary angioplasty determined via myosin heavy chain expression. *Circulation.*96(1):82-90.
4. Alberts BM, Frey L (1970) T4 bacteriophage gene 32: a structural protein in the replication and recombination of DNA. *Nature.*227(5265):1313-8.
5. Autieri MV, Keleman SE, Wendt KW (2003) AIF-1 is an actin-polymerizing and Rac1-activating protein that promotes vascular smooth muscle cell migration. *Circ Res.*92:1107-14.
6. Ban C, Chung S, Park DS, Shim YB (2004) Detection of protein-DNA interaction with a DNA probe: distinction between single-strand and double-strand DNA-protein interaction. *Nucleic Acids Res.*32:e110 1-8.
7. Barr SM, Johnson EM (2001) Ras-induced colony formation and anchorage-independent growth inhibited by elevated expression of Puralpha in NIH3T3 cells. *J Cell Biochem.*81(4):621-38.

8. Bazar L, Meighen D, Harris V, Duncan R, Levens D, Avigan M (1995) Targeted melting and binding of a DNA regulatory element by a transactivator of c-myc. *J Biol Chem.*270(14):8241-8.
9. Becker NA, Kelm RJ, Jr., Vrana JA, Getz MJ, Maher III LJ (2000) Altered sensitivity to single-strand-specific reagents associated with the genomic vascular smooth muscle  $\alpha$ -actin promoter during myofibroblast differentiation. *J Biol Chem.*275:15384-91.
10. Bergemann AD, Johnson EM (1992) The HeLa pur factor binds single-stranded DNA at a specific element conserved in gene flanking regions and origins of DNA replication. *Mol Cell Biol.*12:1257-65.
11. Bergemann AD, Ma ZW, Johnson EM (1992) Sequence of cDNA comprising the human *pur* gene and sequence-specific single-stranded-DNA-binding properties of the encoded protein. *Mol Cell Biol.*12:5673-82.
12. Bergwerff M, Verberne ME, DeRuiter MC, Poelmann RE, Gittenberger-de Groot AC (1998) Neural crest cell contribution to the developing circulatory system: implications for vascular morphology? *Circ Res.*82(2):221-31.
13. Bissell DM, Wang SS, Jarnagin WR, Roll FJ (1995) Cell-specific expression of transforming growth factor-beta in rat liver. Evidence for autocrine regulation of hepatocyte proliferation. *J Clin Invest.*96(1):447-55.
14. Blake RD, Delcourt SG (1996) Thermodynamic effects of formamide on DNA stability. *Nucleic Acids Res.*24(11):2095-103.

15. Blank RS, McQuinn TC, Yin KC, Thompson MM, Takeyasu K, Schwartz RJ, Owens GK (1992) Elements of the smooth muscle alpha-actin promoter required in cis for transcriptional activation in smooth muscle. Evidence for cell type-specific regulation. *J Biol Chem.*267(2):984-9.
16. Blank RS, Owens GK (1990) Platelet-derived growth factor regulates actin isoform expression and growth state in cultured rat aortic smooth muscle cells. *J Cell Physiol.*142(3):635-42.
17. Blank RS, Thompson MM, Owens GK (1988) Cell cycle versus density dependence of smooth muscle alpha actin expression in cultured rat aortic smooth muscle cells. *J Cell Biol.*107(1):299-306.
18. Bloomfield VA (1981) Quasi-elastic light scattering applications in biochemistry and biology. *Annu Rev Biophys Bioeng.*10:421-50.
19. Bochkarev A, Bochkareva E (2004) From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold. *Curr Opin Struct Biol.*14(1):36-42.
20. Bockman DE, Sohal GS (1998) A new source of cells contributing to the developing gastrointestinal tract demonstrated in chick embryos. *Gastroenterology.*114(5):878-82.
21. Brenowitz M, Senear DF, Kingston RE (2001) DNase I footprint analysis of protein-DNA binding. *Current protocols in molecular biology* / edited by Frederick M Ausubel [et al.Chapter 12:Unit 12 4.
22. Brenowitz M, Senear DF, Shea MA, Ackers GK (1986) "Footprint" titrations yield valid thermodynamic isotherms. *Proc Natl Acad Sci U S A.*83(22):8462-6.

23. Brenowitz M, Senear DF, Shea MA, Ackers GK (1986) Quantitative DNase footprint titration: a method for studying protein-DNA interactions. *Methods Enzymol.*130:132-81.
24. Breslauer KJ, Frank R, Blocker H, Marky LA (1986) Predicting DNA duplex stability from the base sequence. *Proc Natl Acad Sci U S A.*83(11):3746-50.
25. Brunelli S, Tagliafico E, De Angelis FG, Tonlorenzi R, Baesso S, Ferrari S, Niinobe M, Yoshikawa K, Schwartz RJ, Bozzoni I, Ferrari S, Cossu G (2004) Msx2 and Necdin combined activities are required for smooth muscle differentiation in mesoangioblast stem cells. *Circ Res.*94:1571-8.
26. Bucala R, Spiegel LA, Chesney J, Hogan M, Cerami A (1994) Circulating fibrocytes define a new leukocyte subpopulation that mediates tissue repair. *Molecular medicine (Cambridge, Mass.)*1(1):71-81.
27. Cann JR (1989) Phenomenological theory of gel electrophoresis of protein-nucleic acid complexes. *J Biol Chem.*264(29):17032-40.
28. Carlini LE, Getz MJ, Strauch AR, Kelm RJ, Jr. (2002) Cryptic MCAT enhancer regulation in fibroblasts and smooth muscle cells. *J Biol Chem.*277:8682-92.
29. Carroll SL, Bergsma DJ, Schwartz RJ (1986) Structure and complete nucleotide sequence of the chicken alpha-smooth muscle (aortic) actin gene. An actin gene which produces multiple messenger RNAs. *J Biol Chem.*261(19):8965-76.
30. Carroll SL, Bergsma DJ, Schwartz RJ (1988) A 29-nucleotide DNA segment containing an evolutionarily conserved motif is required in cis for cell-type-restricted

repression of the chicken alpha-smooth muscle actin gene core promoter. *Mol Cell Biol.*8(1):241-50.

31. Carson JA, Fillmore RA, Schwartz RJ, Zimmer WE (2000) The smooth muscle gamma-actin gene promoter is a molecular target for the mouse bagpipe homologue, mNkx3-1, and serum response factor. *J Biol Chem.*275(50):39061-72.

32. Cavaluzzi MJ, Borer PN (2004) Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. *Nucleic Acids Res.*32(1):e13

33. Champ PC, Maurice S, Vargason JM, Camp T, Ho PS (2004) Distributions of Z-DNA and nuclear factor I in human chromosome 22: a model for coupled transcriptional regulation. *Nucleic Acids Res.*32(22):6501-10.

34. Chang CF, Gallia GL, Muralidharan V, Chen NN, Zoltick P, Johnson E, Khalili K (1996) Evidence that replication of human neurotropic JC virus DNA in glial cells is regulated by the sequence-specific single-stranded DNA-binding protein Pur alpha. *J Virol.*70(6):4150-6.

35. Chaponnier C, Goethals M, Janmey PA, Gabbiani F, Gabbiani G, Vandekerckhove J (1995) The specific NH2-terminal sequence Ac-EEED of alpha-smooth muscle actin plays a role in polymerization in vitro and in vivo. *J Cell Biol.*130(4):887-95.

36. Chen NN, Chang CF, Gallia GL, Kerr DA, Johnson EM, Krachmarov CP, Barr SM, Frisque RJ, Bollag B, Khalili K (1995) Cooperative action of cellular proteins YB-1

and Pur $\alpha$  with the tumor antigen of the Human JC Polyomavirus determines their interaction with the viral lytic control element. *Proc Natl Acad Sci USA*.92:1087-91.

37. Chen NN, Khalili K (1995) Transcriptional regulation of human JC polyomavirus promoters by cellular proteins YB-1 and Pur alpha in glial cells. *J Virol*.69(9):5843-8.

38. Chen S, Lechleider RJ (2004) Transforming growth factor- $\beta$ -induced differentiation of smooth muscle from a neural crest stem cell line. *Circ Res*.94:1196-202.

39. Chen S, Supakar PC, Vellanoweth RL, Song CS, Chatterjee B, Roy AK (1997) Functional role of a conformationally flexible homopurine/homopyrimidine domain of the androgen receptor gene promoter interacting with Sp1 and a pyrimidine single strand DNA-binding protein. *Mol Endocrinol*.11(1):3-15.

40. Chen Y, Kelm RJ, Jr., Budd RC, Sobel BE, Schneider DJ (2004) Inhibition of apoptosis and caspase-3 in vascular smooth muscle cells by plasminogen activator inhibitor type-1. *J Cell Biochem*.92(1):178-88.

41. Chepenik LG, Tretiakova AP, Krachmarov CP, Johnson EM, Khalili K (1998) The single-stranded DNA binding protein, Pur-alpha, binds HIV-1 TAR RNA and activates HIV-1 transcription. *Gene*.210:37-44.

42. Cho A, Graves J, Reidy MA (2000) Mitogen-activated protein kinases mediate matrix metalloproteinase-9 expression in vascular smooth muscle cells. *Arterioscler Thromb Vasc Biol*.20(12):2527-32.

43. Clement S, Hinz B, Dugina V, Gabbiani G, Chaponnier C (2005) The N-terminal Ac-EEED sequence plays a role in alpha-smooth-muscle actin incorporation into stress fibers. *Journal of cell science*.118(Pt 7):1395-404.
44. Clement S, Stouffs M, Bettiol E, Kampf S, Krause KH, Chaponnier C, Jaconi M (2007) Expression and function of alpha-smooth muscle actin during embryonic-stem-cell-derived cardiomyocyte differentiation. *Journal of cell science*.120(Pt 2):229-38.
45. Clouthier DE, Comerford SA, Hammer RE (1997) Hepatic fibrosis, glomerulosclerosis, and a lipodystrophy-like syndrome in PEPCK-TGF-beta1 transgenic mice. *J Clin Invest*.100(11):2697-713.
46. Cobianchi F, SenGupta DN, Zmudzka BZ, Wilson SH (1986) Structure of rodent helix-destabilizing protein revealed by cDNA cloning. *J Biol Chem*.261(8):3536-43.
47. Cogan JG, Subramanian SV, Polikandriotis JA, Kelm RJ, Jr., Strauch AR (2002) Vascular smooth muscle  $\alpha$ -actin gene transcription during myofibroblast differentiation requires Sp1/3 protein binding proximal to the MCAT enhancer. *J Biol Chem*.277:36433-42.
48. Cogan JG, Sun S, Stoflet ES, Schmidt LJ, Getz MJ, Strauch AR (1995) Plasticity of vascular smooth muscle  $\alpha$ -actin gene transcription. *J Biol Chem*.270:11310-21.
49. Cole JL (2004) Analysis of heterogeneous interactions. *Methods Enzymol*.384:212-32.
50. Coles LS, Bartley MA, Bert A, Hunter J, Polyak S, Diamond P, Vadas MA, Goodall GJ (2004) A multi-protein complex containing cold shock domain (Y-box) and

polypyrimidinetract binding proteins forms on the vascularendothelial growth factor mRNA. *Eur J Biochem.*271:648-60.

51. Connaghan-Jones KD, Heneghan AF, Miura MT, Bain DL (2007) Thermodynamic analysis of progesterone receptor-promoter interactions reveals a molecular model for isoform-specific function. *Proc Natl Acad Sci U S A.*104(7):2187-92.

52. Corjay MH, Thompson MM, Lynch KR, Owens GK (1989) Differential effect of platelet-derived growth factor- versus serum-induced growth on smooth muscle alpha-actin and nonmuscle beta-actin mRNA expression in cultured rat aortic smooth muscle cells. *J Biol Chem.*264(18):10501-6.

53. Creazzo TL, Godt RE, Leatherbury L, Conway SJ, Kirby ML (1998) Role of cardiac neural crest cells in cardiovascular development. *Annual review of physiology.*60:267-86.

54. Da Silva N, Bharti A, Shelley CS (2002) hnRNP-K and Pur(alpha) act together to repress the transcriptional activity of the CD43 gene promoter. *Blood.*100(10):3536-44.

55. Damon DH (2005) Sympathetic innervation promotes vascular smooth muscle differentiation. *Am J Physiol Heart Circ Physiol.*288:H2785-H91.

56. Darbinian N, Gallia GL, Khalili K (2001) Helix-destabilizing properties of the human single-stranded DNA- and RNA-binding protein Pur $\alpha$ . *J Cell Biochem.*80:589-95.

57. Darbinian N, Gallia GL, King J, Del Valle L, Johnson EM, Khalili K (2001) Growth inhibition of glioblastoma cells by human Pur(alpha). *J Cell Physiol.*189(3):334-40.

58. Darbinian N, Gallia GL, Kundu M, Shcherbik N, Tretiakova A, Giordano A, Khalili K (1999) Association of Pura $\alpha$  and E2F-1 suppresses transcriptional activity of E2F-1. *Oncogene*.18:6398-402.
59. Darbinian N, Sawaya BE, Khalili K, Jaffe N, Wortman B, Giordano A, Amini S (2001) Functional interaction between cyclin T1/cdk9 and Puralpha determines the level of TNFalpha promoter activation by Tat in glial cells. *Journal of neuroimmunology*.121(1-2):3-11.
60. Darby I, Skalli O, Gabbiani G (1990) Alpha-smooth muscle actin is transiently expressed by myofibroblasts during experimental wound healing. *Laboratory investigation; a journal of technical methods and pathology*.63(1):21-9.
61. Davies MJ (1996) Stability and instability: two faces of coronary atherosclerosis. The Paul Dudley White Lecture 1995. *Circulation*.94(8):2013-20.
62. Davis-Smyth T, Duncan RC, Zheng T, Michelotti G, Levens D (1996) The far upstream element-binding proteins comprise an ancient family of single-strand DNA-binding transactivators. *J Biol Chem*.271(49):31679-87.
63. Desmouliere A, Chaponnier C, Gabbiani G (2005) Tissue repair, contraction, and the myofibroblast. *Wound Rep Reg*.13:7-12.
64. Desmouliere A, Geinoz A, Gabbiani F, Gabbiani G (1993) Transforming growth factor-beta 1 induces alpha-smooth muscle actin expression in granulation tissue myofibroblasts and in quiescent and growing cultured fibroblasts. *J Cell Biol*.122(1):103-11.

65. Desmouliere A, Redard M, Darby I, Gabbiani G (1995) Apoptosis mediates the decrease in cellularity during the transition between granulation tissue and scar. *The American journal of pathology*.146(1):56-66.
66. Desmouliere A, Rubbia-Brandt L, Abdiu A, Walz T, Macieira-Coelho A, Gabbiani G (1992) Alpha-smooth muscle actin is expressed in a subpopulation of cultured and cloned fibroblasts and is modulated by gamma-interferon. *Experimental cell research*.201(1):64-73.
67. Desmouliere A, Rubbia-Brandt L, Grau G, Gabbiani G (1992) Heparin induces alpha-smooth muscle actin expression in cultured fibroblasts and in granulation tissue myofibroblasts. *Laboratory investigation; a journal of technical methods and pathology*.67(6):716-26.
68. Dhalla AK, Ririe SS, Swamynathan SK, Weber KT, Guntaka RV (1998) chk-YB-1b, a Y-box binding protein activates transcription from rat alpha1(I) procollagen gene promoter. *Biochem J*.336 ( Pt 2):373-9.
69. Didier DK, Schiffenbauer J, Woulfe SL, Zacheis M, Schwartz BD (1988) Characterization of the cDNA encoding a protein binding to the major histocompatibility complex class II Y box. *Proc Natl Acad Sci U S A*.85(19):7322-6.
70. Ding Y, Osugi T, Kuo CH, Tanaka H, Do E, Irie Y, Miki N (1997) Characterization of a nuclear factor that enhances DNA binding activity of SSCRE-BP/PUR alpha, a single-stranded DNA binding protein. *Neurochem Int*.31(1):45-54.

71. Dixon JL, Stoops JD, Parker JL, Laughlin MH, Weisman GA, Sturek M (1999) Dyslipidemia and vascular dysfunction in diabetic pigs fed an atherogenic diet. *Arterioscler Thromb Vasc Biol.*19(12):2981-92.
72. Doucet C, Brouty-Boye D, Pottin-Clemenceau C, Canonica GW, Jasmin C, Azzarone B (1998) Interleukin (IL) 4 and IL-13 act on human lung fibroblasts. Implication in asthma. *J Clin Invest.*101(10):2129-39.
73. Drew HR, Travers AA (1984) DNA structural variations in the E. coli tyrT promoter. *Cell.*37(2):491-502.
74. Du Q, Tomkinson AE, Gardner PD (1997) Transcriptional regulation of neuronal nicotinic acetylcholine receptor genes. A possible role for the DNA-binding protein Puralpha. *J Biol Chem.*272(23):14990-5.
75. Duncan R, Bazar L, Michelotti G, Tomonaga T, Krutzsch H, Avigan M, Levens D (1994) A sequence-specific, single-strand binding protein activates the far upstream element of c-myc and defines a new DNA-binding motif. *Genes Dev.*8(4):465-80.
76. Engelse MA, Lardenoye JH, Neele JM, Gimbergen JM, de Vries MR, Lamfers MLM, Pannekok H, Quax PHA, de Vries CJM (2002) Adenoviral activin A expression prevents intimal hyperplasia in human and murine blood vessels by maintaining the contractile smooth muscle phenotype. *Circ Res.*90:1128-34.
77. Evdokimova V, Ruzanov P, Anglesio MS, Sorokin AV, Ovchinnikov LP, Buckley J, Triche TJ, Sonenberg N, Sorensen PH (2006) Akt-mediated YB-1 phosphorylation activates translation of silent mRNA species. *Mol Cell Biol.*26(1):277-92.

78. Fabunmi RP, Sukhova GK, Sugiyama S, Libby P (1998) Expression of tissue inhibitor of metalloproteinases-3 in human atheroma and regulation in lesion-associated cells: a potential protective mechanism in plaque stability. *Circ Res.*83(3):270-8.
79. Fatigati V, Murphy RA (1984) Actin and tropomyosin in smooth muscles. *J Biol Chem.*259:14383-8.
80. Fawell SE, White R, Hoare S, Sydenham M, Page M, Parker MG (1990) Inhibition of estrogen receptor-DNA binding by the "pure" antiestrogen ICI 164,384 appears to be mediated by impaired receptor dimerization. *Proc Natl Acad Sci U S A.*87(17):6883-7.
81. Folta-Stogniew E, Williams KR (1999) Determination of Molecular Masses of Proteins in Solution: Implementation of an HPLC Size Exclusion Chromatography and Laser Light Scattering Service in a Core Laboratory. *J Biomol Techniques.*10:51-63.
82. Foster DN, Min B, Foster LK, Stoflet ES, Sun S, Getz MJ, Strauch AR (1992) Positive and negative cis-regulatory elements mediate expression of the mouse vascular smooth muscle  $\alpha$ -actin gene. *J Biol Chem.*267:11995-2003.
83. Frid MG, Kale VA, Stenmark KR (2002) Mature vascular endothelium can give rise to smooth muscle cells via endothelial-mesenchymal transdifferentiation: in vitro analysis. *Circ Res.*90(11):1189-96.
84. Fried M, Crothers DM (1981) Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.*9(23):6505-25.
85. Fried MG, Daugherty MA (1998) Electrophoretic analysis of multiple protein-DNA interactions. *Electrophoresis.*19(8-9):1247-53.

86. Fukuda H, Katahira M, Tsuchiya N, Enokizono Y, Sugimura T, Nagao M, Nakagama H (2002) Unfolding of quadruplex structure in the G-rich strand of the minisatellite repeat by the binding protein UP1. *Proc Natl Acad Sci U S A.*99(20):12685-90.
87. Fukuda T, Ashizuka M, Nakamura T, Shibahara K, Maeda K, Izumi H, Kohno K, Kuwano M, Uchiumi T (2004) Characterization of the 5'-untranslated region of YB-1 mRNA and autoregulation of translation by YB-1 protein. *Nucleic Acids Res.*32(2):611-22.
88. Gabbiani G, Kocher O, Bloom WS, Vandekerckhove J, Weber K (1984) Actin expression in smooth muscle cells of rat aortic intimal thickening, human atheromatous plaque, and cultured rat aortic media. *J Clin Invest.*73(1):148-52.
89. Gabbiani G, Ryan GB, Majne G (1971) Presence of modified fibroblasts in granulation tissue and their possible role in wound contraction. *Experientia.*27(5):549-50.
90. Gallia GL, Darbinian N, Jaffe N, Khalili K (2001) Single-stranded nucleic acid-binding protein, Pur $\alpha$ , interacts with RNA homologous to 18S ribosomal RNA and inhibits translation in vitro. *J Cell Biochem.*83:355-63.
91. Gallia GL, Darbinian N, Johnson EM, Khalili K (1999) Self-association of Pur $\alpha$  is mediated by RNA. *J Cell Biochem.*74:334-48.
92. Gallia GL, Johnson EM, Khalili K (2000) Pur $\alpha$ : a multifunctional single-stranded DNA- and RNA-binding protein. *Nucleic Acids Res.*28:3197-205.

93. Gallia GL, Safak M, Khalili K (1998) Interaction of the single-stranded DNA-binding protein Puralpha with the human polyomavirus JC virus early protein T-antigen. *J Biol Chem.*273(49):32662-9.
94. Gan Q, Yoshida T, Li J, Owens GK (2007) Smooth muscle cells and myofibroblasts use distinct transcriptional mechanisms for smooth muscle alpha-actin expression. *Circ Res.*101(9):883-92.
95. Gaudreault I, Guay D, Lebel M (2004) YB-1 promotes strand separation in vitro of duplex DNA containing either mispaired bases or cisplatin modifications, exhibits endonucleolytic activities and binds several DNA repair proteins. *Nucleic Acids Res.*32(1):316-27.
96. Geary RL, Wong JM, Rossini A, Schwartz SM, Adams LD (2002) Expression profiling identifies 147 genes contributing to a unique primate neointimal smooth muscle cell phenotype. *Arterioscler Thromb Vasc Biol.*22(12):2010-6.
97. Gerber HP, Seipel K, Georgiev O, Hofferer M, Hug M, Rusconi S, Schaffner W (1994) Transcriptional activation modulated by homopolymeric glutamine and proline stretches. *Science.*263(5148):808-11.
98. Getz MJ, Elder PK, Benz J, B.W., Stephens RE, Moses HL (1976) Effect of cell proliferation on levels and diversity of poly(A)-containing mRNA. *Cell.*7:255-65.
99. Ghisolfi L, Joseph G, Amalric F, Erard M (1992) The glycine-rich domain of nucleolin has an unusual supersecondary structure responsible for its RNA-helix-destabilizing properties. *J Biol Chem.*267(5):2955-9.

100. Gordon D, Reidy MA, Benditt EP, Schwartz SM (1990) Cell proliferation in human coronary arteries. *Proc Natl Acad Sci U S A.*87(12):4600-4.
101. Grossmann ME, Tindall DJ (1995) The androgen receptor is transcriptionally suppressed by proteins that bind single-stranded DNA. *J Biol Chem.*270(18):10968-75.
102. Gueron M, Kochoyan M, Leroy JL (1987) A single mode of DNA base-pair opening drives imino proton exchange. *Nature.*328(6125):89-92.
103. Gupta M, Sueblinvong V, Gupta MP (2007) The single-strand DNA/RNA-binding protein, Purbeta, regulates serum response factor (SRF)-mediated cardiac muscle gene expression. *Canadian journal of physiology and pharmacology.*85(3-4):349-59.
104. Gupta M, Sueblinvong V, Raman J, Jeevanandam J, Gupta MP (2003) Single-stranded DNA-binding proteins Pur $\alpha$  and Pur $\beta$  bind to a purine-rich negative regulatory element of the  $\alpha$ -myosin heavy chain gene and control transcriptional and translational regulation of gene expression. *J Biol Chem.*278:44935-48.
105. Ha SC, Lowenhaupt K, Rich A, Kim YG, Kim KK (2005) Crystal structure of a junction between B-DNA and Z-DNA reveals two extruded bases. *Nature.*437(7062):1183-6.
106. Haas S, Steplewski A, Siracusa LD, Amini S, Khalili K (1995) Identification of a sequence-specific single-stranded DNA binding protein that suppresses transcription of the mouse myelin basic protein gene. *J Biol Chem.*270(21):12503-10.
107. Haas S, Thatikunta P, Steplewski A, Johnson EM, Khalili K, Amini S (1995) A 39-kD DNA-binding protein from mouse brain stimulates transcription of myelin basic protein gene in oligodendrocytic cells. *J Cell Biol.*130(5):1171-9.

108. Hampshire AJ, Rusling DA, Broughton-Head VJ, Fox KR (2007) Footprinting: a method for determining the sequence selectivity, affinity and kinetics of DNA-binding ligands. *Methods*.42(2):128-40.
109. Harding SE (1995) On the hydrodynamic analysis of macromolecular conformation. *Biophys Chem*.55:69-93.
110. Hautmann MB, Madsen CS, Owens GK (1997) A transforming growth factor beta (TGFbeta) control element drives TGFbeta-induced stimulation of smooth muscle alpha-actin gene expression in concert with two CArG elements. *J Biol Chem*.272(16):10948-56.
111. Hautmann MB, Thompson MM, Swartz EA, Olson EN, Owens GK (1997) Angiotensin II-induced stimulation of smooth muscle alpha-actin expression by serum response factor and the homeodomain transcription factor MHox. *Circ Res*.81(4):600-10.
112. Hebenstreit D, Horejs-Hoeck J, Duschl A (2005) JAK/STAT-dependent gene regulation by cytokines. *Drug News Perspect*.18(4):243-9.
113. Heneghan AF, Connaghan-Jones KD, Miura MT, Bain DL (2006) Cooperative DNA binding by the B-isoform of human progesterone receptor: thermodynamic analysis reveals strongly favorable and unfavorable contributions to assembly. *Biochemistry*.45(10):3285-96.
114. Heneghan AF, Connaghan-Jones KD, Miura MT, Bain DL (2007) Coactivator assembly at the promoter: efficient recruitment of SRC2 is coupled to cooperative DNA binding by the progesterone receptor. *Biochemistry*.46(39):11023-32.

115. Herault Y, Chatelain G, Brun G, Michel D (1993) The PUR element stimulates transcription and is a target for single strand-specific binding factors conserved among vertebrate classes. *Cell Mol Biol Res.*39(8):717-25.
116. Herring BP, Kriegel AM, Hoggatt AM (2001) Identification of Barx2b, a serum response factor-associated homeodomain protein. *J Biol Chem.*276(17):14482-9.
117. Holycross BJ, Blank RS, Thompson MM, Peach MJ, Owens GK (1992) Platelet-derived growth factor-BB-induced suppression of smooth muscle cell differentiation. *Circ Res.*71(6):1525-32.
118. Hoskins RA, Carlson JW, Kennedy C, Acevedo D, Evans-Holm M, Frise E, Wan KH, Park S, Mendez-Lago M, Rossi F, Villasante A, Dimitri P, Karpen GH, Celniker SE (2007) Sequence finishing and mapping of *Drosophila melanogaster* heterochromatin. *Science.*316(5831):1625-8.
119. Howell DC, Laurent GJ, Chambers RC (2002) Role of thrombin and its major cellular receptor, protease-activated receptor-1, in pulmonary fibrosis. *Biochemical Society transactions.*30(2):211-6.
120. Hu Y, Zhang Z, Torsney E, Afzal AR, Davison F, Metzler B, Xu Q (2004) Abundant progenitor cells in the adventitia contribute to atherosclerosis of vein grafts in ApoE-deficient mice. *J Clin Invest.*113:1258-65.
121. Hultman T, Stahl S, Hornes E, Uhlen M (1989) Direct solid phase sequencing of genomic and plasmid DNA using magnetic beads as solid support. *Nucleic Acids Res.*17(13):4937-46.

122. Hungerford JE, Little CD (1999) Developmental biology of the vascular smooth muscle cell: building a multilayered vessel wall. *Journal of vascular research*.36(1):2-27.
123. Ise T, Nagatani G, Imamura T, Kato K, Takano H, Nomoto M, Izumi H, Ohmori H, Okamoto T, Ohga T, Uchiumi T, Kuwano M, Kohno K (1999) Transcription factor Y-box binding protein 1 binds preferentially to cisplatin-modified DNA and interacts with proliferating cell nuclear antigen. *Cancer research*.59(2):342-6.
124. Ishida Y, Kondo T, Takayasu T, Iwakura Y, Mukaida N (2004) The essential involvement of cross-talk between IFN-gamma and TGF-beta in the skin wound-healing process. *J Immunol*.172(3):1848-55.
125. Itoh H, Wortman MJ, Kanovsky M, Uson RR, Gordon RE, Alfano N, Johnson EM (1998) Alterations in Pur $\alpha$  Levels and Intracellular Localization in the CV-1 Cell Cycle. *Cell Growth Differentiation*.9:651-65.
126. Itoh S, Umemoto S, Hiromoto M, Toma Y, Tomachika Y, Aoyagi S, Tanaka M, Fujii T, Matsuzaki M (2002) Importance of NAD(P)H oxidase-mediated oxidative stress and contractile type smooth muscle myosin heavy chain SM2 at the early stage of atherosclerosis. *Circulation*.105:2288-95.
127. Iwasaki H, Isayama T, Ichiki T, Kikuchi M (1987) Intermediate filaments of myofibroblasts. Immunochemical and immunocytochemical analyses. *Pathology, research and practice*.182(2):248-54.
128. Jameson DM, Seifried SE (1999) Quantification of protein-protein interactions using fluorescence polarization. *Methods*.19(2):222-33.

129. Ji J, Tsika GL, Rindt H, Schreiber KL, McCarthy JJ, Kelm RJ, Jr., Tsika R (2007) Pur $\alpha$  and Pur $\beta$  Collaborate with Sp3 To Negatively Regulate  $\beta$ -Myosin Heavy Chain Gene Expression during Skeletal Muscle Inactivity. *Mol Cell Biol.*27(4):1531-43.
130. Johnson AC, Jinno Y, Merlino GT (1988) Modulation of epidermal growth factor receptor proto-oncogene transcription by a promoter site sensitive to S1 nuclease. *Mol Cell Biol.*8(10):4174-84.
131. Johnson AD, Owens GK (1999) Differential activation of the SM $\alpha$ A promoter in smooth vs. skeletal muscle cells by bHLH factors. *Am J Physiol.*276(6 Pt 1):C1420-31.
132. Johnson EM (2003) The Pur protein family: clues to function from recent studies on cancer and AIDS. *Anticancer Res.*23:2093-100.
133. Johnson EM, Chen PL, Krachmarov CP, Barr SM, Kanovsky M, Ma ZW, Lee WH (1995) Association of human Pur $\alpha$  with the retinoblastoma protein, Rb, regulates binding to the single-stranded DNA Pur $\alpha$  recognition element. *J Biol Chem.*270:24352-60.
134. Juliano RL, Dixit VR, Kang H, Kim TY, Miyamoto Y, Xu D (2005) Epigenetic manipulation of gene expression: a toolkit for cell biologists. *J Cell Biol.*169:847-57.
135. Jung F, Johnson AD, Kumar MS, Wei B, Hautmann M, Owens GK, McNamara C (1999) Characterization of an E-box-dependent cis element in the smooth muscle alpha-actin promoter. *Arterioscler Thromb Vasc Biol.*19(11):2591-9.

136. Jurk M, Weissinger F, Lottspeich F, Schwarz U, Winnacker EL (1996) Characterization of the single-strand-specific BPV-1 origin binding protein, SPSF I, as the HeLa Pur alpha factor. *Nucleic Acids Res.*24(14):2799-806.
137. Kahler CM, Schratzberger P, Wiedermann CJ (1997) Response of vascular smooth muscle cells to the neuropeptide secretoneurin. *Arterioscler Thromb Vasc Biol.*17:2029-35.
138. Kanai Y, N. Dohmae, N. Hirokawa (2004) Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. *Neuron.*43:513-25.
139. Kasuga T, Cheng J, Mitchelson KR (1995) Metastable single-strand DNA conformational polymorphism analysis results in enhanced polymorphism detection. *PCR methods and applications.*4(4):227-33.
140. Kawai-Kowase K, Owens GK (2007) Multiple repressor pathways contribute to phenotypic switching of vascular smooth muscle cells. *Am J Physiol Cell Physiol.*292(1):C59-69.
141. Kelm RJ, Jr., Cogan JJ, Elder PK, Strauch AR, Getz MJ (1999) Molecular interactions between single-stranded DNA-binding proteins associated with an essential MCAT element in the mouse smooth muscle  $\alpha$ -actin promoter. *J Biol Chem.*274(20):14238-45.
142. Kelm RJ, Jr., Elder PK, Getz MJ (1999) The single-stranded DNA-binding proteins Pur $\alpha$ , Pur $\beta$ , and MSY1 specifically interact with an exon derived mouse vascular smooth muscle  $\alpha$ -actin messenger RNA sequence. *J Biol Chem.*274(53):38268-75.

143. Kelm RJ, Jr., Elder PK, Strauch AR, Getz MJ (1997) Sequence of cDNAs encoding components of vascular actin single-stranded DNA-binding factor 2 establish identity to Pur $\alpha$  and Pur $\beta$ . *J Biol Chem.*272:26726-33.
144. Kelm RJ, Jr., Sun S, Strauch AR, Getz MJ (1996) Repression of transcriptional enhancer factor-1 and activator protein-1-dependent enhancer activity by vascular actin single-stranded DNA binding factor 2. *J Biol Chem.*271(39):24278-85.
145. Kelm RJ, Jr., Wang SX, Polikandriotis JA, Strauch AR (2003) Structure/function analysis of mouse Pur $\beta$ , a single-stranded DNA-binding repressor of vascular smooth muscle  $\alpha$ -actin gene transcription. *J Biol Chem.*278:38749-57.
146. Khalili K, Valle LD, Muralidharan V, Gault WJ, Darbinian N, Otte J, Meier E, Johnson EM, Daniel DC, Kinoshita Y, Amini S, Gordon J (2003) Pur $\alpha$  is essential for postnatal brain development and developmentally coupled cellular proliferation as revealed by genetic inactivation in the mouse. *Mol Cell Biol.*23(19):6857-75.
147. Kim JH, Bushel PR, Kumar CC (1993) Smooth muscle alpha-actin promoter activity is induced by serum stimulation of fibroblast cells. *Biochem Biophys Res Commun.*190(3):1115-21.
148. Knapp AM, Ramsey JE, Wang SX, Godburn KE, Strauch AR, Kelm RJ, Jr. (2006) Nucleoprotein interactions governing cell type-dependent repression of the mouse smooth muscle alpha-actin promoter by single-stranded DNA-binding proteins Pur alpha and Pur beta. *J Biol Chem.*281(12):7907-18.
149. Knapp AM, Ramsey JE, Wang SX, Strauch AR, Kelm RJ, Jr. (2007) Structure-function analysis of mouse Pur beta II. Conformation altering mutations disrupt single-

stranded DNA and protein interactions crucial to smooth muscle alpha-actin gene repression. *J Biol Chem.*282(49):35899-909.

150. Kocher O, Gabbiani G (1986) Cytoskeletal features of normal and atheromatous human arterial smooth muscle cells. *Human pathology.*17(9):875-80.

151. Kolodgie FD, Burke AP, Farb A, Gold HK, Yuan J, Narula J, Finn AV, Virmani R (2001) The thin-cap fibroatheroma: a type of vulnerable plaque: the major precursor lesion to acute coronary syndromes. *Current opinion in cardiology.*16(5):285-92.

152. Koppel DE (1972) Analysis of Macromolecular Polydispersity in Intensity Correlation Spectroscopy: The Method of Cumulants. *J Chem Phys.*57(11):4814-20.

153. Kouzine F, Liu J, Sanford S, Chung HJ, Levens D (2004) The dynamic response of upstream DNA to transcription-generated torsional stress. *Nature structural & molecular biology.*11(11):1092-100.

154. Kouzine F, Sanford S, Elisha-Feil Z, Levens D (2008) The functional response of upstream DNA to dynamic supercoiling in vivo. *Nature structural & molecular biology.*15(2):146-54.

155. Krachmarov CP, Chepenik LG, Barr-Vagell S, Khalili K, Johnson EM (1996) Activation of the JC virus Tat-responsive transcriptional element by association of the Tat protein of human immunodeficiency virus 1 with cellular protein Pur $\alpha$ . *Proc Natl Acad Sci.*93:14112-7.

156. Krueger A, Protozanova E, Frank-Kamenetskii MD (2006) Sequence-dependent base pair opening in DNA double helix. *Biophys J.*90(9):3091-9.

157. Kumar MS, Hendrix JA, Johnson D, Owens GK (2003) Smooth muscle  $\alpha$ -actin gene requires two E-boxes for proper expression in vivo and is a target of class I basic helix-loop-helix proteins. *Circ Res.*92:840-7.
158. Kumar MS, Owens GK (2003) Combinatorial control of smooth muscle-specific gene expression. *Arterioscler Thromb Vasc Biol.*23:737-47.
159. Kuntz ID (1971) Hydration of macromolecules. III. Hydration of polypeptides. *J Am Chem Soc.*93(2):514-6.
160. Kuntz ID (1971) Hydration of macromolecules. IV. Polypeptide conformation in frozen solutions. *J Am Chem Soc.*93(2):516-8.
161. Lancillotti F, Lopez MC, Arias P, Alonso C (1987) Z-DNA in transcriptionally active chromosomes. *Proc Natl Acad Sci U S A.*84(6):1560-4.
162. Lasham A, Lindridge E, Rudert F, Onrust R, Watson J (2000) Regulation of the human fas promoter by YB-1, Puralpha and AP-1 transcription factors. *Gene.*252(1-2):1-13.
163. Laue TM (1992) Short column sedimentation equilibrium analysis for rapid characterization of macromolecules. In: *Technical Information DS-835*, Spinco Business Unit: Palo Alto, CA.
164. Laue TM, Shah BD, Ridgeway TM, Pelletier S (1992) Computer-Aided Interpretation of Analytical Sedimentation Data for Proteins. In: *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, Harding SE, Rowe AJ, Horton JC, (eds.), pp. 90-125, The Royal Chemistry Society: Cambridge, UK.

165. Layne MD, Yet SF, Maemura K, Hsieh CM, Liu X, Ith B, Lee ME, Perrella MA (2002) Characterization of the mouse aortic carboxypeptidase-like protein promoter reveals activity in differentiated and dedifferentiated vascular smooth muscle cells. *Circ Res.*90(6):728-36.
166. Leavitt J, Gunning P, Kedes L, Jariwalla R (1985) Smooth muscle  $\alpha$ -actin is a transformation-sensitive marker for mouse NIH 3T3 and Rat-2 Cells. *Nature.*316:840-2.
167. Lei M, Baumann P, Cech TR (2002) Cooperative binding of single-stranded telomeric DNA by the Pot1 protein of *Schizosaccharomyces pombe*. *Biochemistry.*41(49):14560-8.
168. Lemaitre V, Soloway PD, D'Armiento J (2003) Increased medial degradation with pseudo-aneurysm formation in apolipoprotein E-knockout mice deficient in tissue inhibitor of metalloproteinases-1. *Circulation.*107(2):333-8.
169. Lezon-Geyda K, Najfeld V, Johnson EM (2001) Deletions of PURA, at 5q31, and PURB, at 7p13, in myelodysplastic syndrome and progression to acute myelogenous leukemia. *Leukemia.*15(6):954-62.
170. Li G, Chen YF, Kelpke SS, Oparil S, Thompson JA (2002) Estrogen attenuates integrin- $\beta_3$ -dependent adventitial fibroblast migration after inhibition of osteopontin production in vascular smooth muscle cells  
*Circulation.*101:2949-55.
171. Li G, Oparil S, Kelpke SS, Chen YF, Thompson JA (2002) Fibroblast growth factor receptor-1 signalling induces osteopontin expression and vascular smooth muscle cell-dependent adventitial fibroblast migration in vitro. *Circulation.*106:854-9.

172. Li S, Wang D-Z, Wang Z, Richardson JA, Olson EN (2003) The serum response factor coactivator myocardin is required for vascular smooth muscle development. *Proc Natl Acad Sci.*100:9366-70.
173. Li Y, Koike K, Ohashi S, Funakoshi T, Tadano M, Kobayashi S, Anzai K, Shibata N, Kobayashi M (2001) Pur alpha protein implicated in dendritic RNA transport interacts with ribosomes in neuronal cytoplasm. *Biological & pharmaceutical bulletin.*24(3):231-5.
174. Libby P, Ridker PM, Maseri A (2002) Inflammation and atherosclerosis. *Circulation.*105:1135-43.
175. Libby P, Theroux P (2005) Pathophysiology of coronary artery disease. *Circulation.* (111):3481-8.
176. Limesand SW, Jeckel KM, Anthony RV (2004) Puralpha, a single-stranded deoxyribonucleic acid binding protein, augments placental lactogen gene transcription. *Mol Endocrinol.*18(2):447-57.
177. Liu H, Barr SM, Chu C, Kohtz DS, Kinoshita Y, Johnson EM (2005) Functional interaction of Pur $\alpha$  with the Cdk2 moiety of cyclin A/Cdk2. *Biochem Biophys Res Commun.*328(4):851-7.
178. Liu J, He L, Collins I, Ge H, Libutti D, Li J, Egly JM, Levens D (2000) The FBP interacting repressor targets TFIID to inhibit activated transcription. *Molecular cell.*5(2):331-41.
179. Liu LF, Wang JC (1987) Supercoiling of the DNA template during transcription. *Proc Natl Acad Sci U S A.*84(20):7024-7.

180. Lundblad JR, Laurance M, Goodman RH (1996) Fluorescence polarization analysis of protein-DNA and protein-protein interactions. *Mol Endocrinol*.10(6):607-12.
181. Ma C, Chegini N (1999) Regulation of matrix metalloproteinases (MMPs) and their tissue inhibitors in human myometrial smooth muscle cells by TGF-beta1. *Molecular human reproduction*.5(10):950-4.
182. Ma ZW, Bergemann AD, Johnson EM (1994) Conservation in human and mouse Pur alpha of a motif common to several proteins involved in initiation of DNA replication. *Gene*.149(2):311-4.
183. MacDonald GH, Itoh-Lindstrom Y, Ting JP (1995) The transcriptional regulatory protein, YB-1, promotes single-stranded regions in the DRA promoter. *J Biol Chem*.270(8):3527-33.
184. Mack CP, Owens GK (1999) Regulation of smooth muscle alpha-actin expression in vivo is dependent on CArG elements within the 5' and first intron promoter regions. *Circ Res*.84(7):852-61.
185. Mallat Z, Gojova A, Marchiol-Fournigault C, Esposito B, Kamate C, Merval R, Fradelizi D, Tedgui A (2001) Inhibition of transforming growth factor-beta signaling accelerates atherosclerosis and induces an unstable plaque phenotype in mice. *Circ Res*.89(10):930-4.
186. Manabe I, Owens GK (2001) Recruitment of serum response factor and hyperacetylation of histones at smooth muscle-specific regulatory regions during differentiation of a novel P19-derived in vitro smooth muscle differentiation system. *Circ Res*.88:1127-34.

187. Marmur J, Doty P (1962) Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J Mol Biol.*5:109-18.
188. Matsumoto K, Wolffe AP (1998) Gene regulation by Y-box proteins: coupling control of transcription and translation. *Trends Cell Biol.*8:318-23.
189. McDonald OG, Wamhoff BR, Hoofnagle MH, Owens GK (2006) Control of SRF binding to CArG box chromatin regulates smooth muscle gene expression in vivo. *J Clin Invest.*116(1):36-48.
190. Melnikova IN, Yang Y, Gardner PD (2000) Interactions between regulatory proteins that bind to the nicotinic receptor beta4 subunit gene promoter. *European journal of pharmacology.*393(1-3):75-83.
191. Michelotti EF, Tomonaga T, Krutzsch H, Levens D (1995) Cellular nucleic acid binding protein regulates the CT element of the human c-myc protooncogene. *J Biol Chem.*270(16):9494-9.
192. Michelotti GA, Michelotti EF, Pullner A, Duncan RC, Eick D, Levens D (1996) Multiple single-stranded cis elements are associated with activated chromatin of the human c-myc gene in vivo. *Mol Cell Biol.*16(6):2656-69.
193. Mikawa T, Fischman DA (1992) Retroviral analysis of cardiac morphogenesis: discontinuous formation of coronary vessels. *Proc Natl Acad Sci U S A.*89(20):9504-8.
194. Min BH, Foster DN, Strauch AR (1990) The 5'-flanking region of the mouse vascular smooth muscle alpha-actin gene contains evolutionarily conserved sequence motifs within a functional promoter. *J Biol Chem.*265(27):16667-75.

195. Minty A, Kedes L (1986) Upstream regions of the human cardiac actin gene that modulate its transcription in muscle cells: presence of an evolutionarily conserved repeated motif. *Mol Cell Biol.*6(6):2125-36.
196. Miwa T, Boxer LM, Kedes L (1987) CArG boxes in the human cardiac alpha-actin gene are core binding sites for positive trans-acting regulatory factors. *Proc Natl Acad Sci U S A.*84(19):6702-6.
197. Miwa T, Kedes L (1987) Duplicated CArG box domains have positive and mutually dependent regulatory roles in expression of the human alpha-cardiac actin gene. *Mol Cell Biol.*7(8):2803-13.
198. Monaghan A, Webster A, Hay RT (1994) Adenovirus DNA binding protein: helix destabilising properties. *Nucleic Acids Res.*22(5):742-8.
199. Mosse PR, Campbell GR, Campbell JH (1986) Smooth muscle phenotypic expression in human carotid arteries. II. Atherosclerosis-free diffuse intimal thickenings compared with the media. *Arteriosclerosis (Dallas, Tex.)*6(6):664-9.
200. Muralidharan V, Sweet T, Nadruga Y, Amini S, Khalili K (2001) Regulation of Puralpha gene transcription: evidence for autoregulation of Puralpha promoter. *J Cell Physiol.*186(3):406-13.
201. Muralidharan V, Tretiakova A, Steplewski A, Haas S, Amini S, Johnson E, Khalili K (1997) Evidence for inhibition of MyEF-2 binding to MBP promoter by MEF-1/Pur alpha. *J Cell Biochem.*66(4):524-31.

202. Nair SK, Burley SK (2003) X-ray structures of Myc-Max and Mad-Max recognizing DNA. Molecular bases of regulation by proto-oncogenic transcription factors. *Cell*.112(2):193-205.
203. Negishi Y, Nishita Y, Saegusa Y, Kakizaki I, Galli I, Kihara F, Tamai K, Miyajima N, Iguchi-Ariga SM, Ariga H (1994) Identification and cDNA cloning of single-stranded DNA binding proteins that interact with the region upstream of the human c-myc gene. *Oncogene*.9(4):1133-43.
204. Nordheim A, Rich A (1983) Negatively supercoiled simian virus 40 DNA contains Z-DNA segments within transcriptional enhancer sequences. *Nature*.303(5919):674-9.
205. Notredame C, Higgins DG, Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol*.302(1):205-17.
206. Ohashi S, Koike K, Omori A, Ichinose S, Ohara S, Kobayashi S, Sato TA, Anzai K (2002) Identification of mRNA/protein (mRNP) complexes containing Puralpha, mStaufen, fragile X protein, and myosin Va and their association with rough endoplasmic reticulum equipped with a kinesin motor. *J Biol Chem*.277(40):37804-10.
207. Osugi T, Taniura H, Ikemoto M, Miki N (1991) Effects of chronic exposure of NG108-15 cells to morphine or ethanol on binding of nuclear factors to cAMP-response element. *Biochem Biophys Res Commun*.174(1):25-31.
208. Owczarzy R, Tataurov AV, Wu Y, Manthey JA, McQuisten KA, Almabrazi HG, Pedersen KF, Lin Y, Garretson J, McEntaggart NO, Sailor CA, Dawson RB, Peek AS

- (2008) IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Res.*
209. Owens GK (1995) Regulation of differentiation of vascular smooth muscle cells. *Physiol Rev.*75:487-517.
210. Owens GK, Kumar MS, Wamhoff BR (2003) Molecular regulation of vascular smooth muscle cell differentiation markers in development and disease. *Physiol Rev.*84:767-801.
211. Pace CN, Vajdos F, Fee L, Grimsley G, Gray T (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.*4(11):2411-23.
212. Paddison PJ, Caudy AA, Hannon GJ, Conklin DS (2002) Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Devel.*16:948-58.
213. Pecora R (1972) Quasi-elastic light scattering from macromolecules. *Annu Rev Biophys Bioeng.*1:257-76.
214. Penberthy WT, Zhao C, Zhang Y, Jessen JR, Yang Z, Bricaud O, Collazo A, Meng A, Lin S (2004) Pur alpha and Sp8 as opposing regulators of neural gata2 expression. *Developmental biology.*275(1):225-34.
215. Petruska J, Goodman MF, Boosalis MS, Sowers LC, Cheong C, Tinoco I, Jr. (1988) Comparison between DNA melting thermodynamics and DNA polymerase fidelity. *Proc Natl Acad Sci U S A.*85(17):6252-6.
216. Philo JS (1997) An improved function for fitting sedimentation velocity data for low-molecular-weight solutes. *Biophys J.*72(1):435-44.

217. Philo JS (2000) A method for directly fitting the time derivative of sedimentation velocity data and an alternative algorithm for calculating sedimentation coefficient distribution functions. *Analytical Biochemistry*.279:151-63.
218. Pidkovka NA, Cherepanova OA, Yoshida T, Alexander MR, Deaton RA, Thomas JA, Leitinger N, Owens GK (2007) Oxidized phospholipids induce phenotypic switching of vascular smooth muscle cells in vivo and in vitro. *Circ Res*.101(8):792-801.
219. Powell DW, Mifflin RC, Valentich JC, Crowe SE, Saada JI, West AB (1999) Myofibroblasts. I. Paracrine cells important in health and disease. *Am J Physiol*.277:C1-C9.
220. Rahmouni AR, Wells RD (1989) Stabilization of Z DNA in vivo by localized supercoiling. *Science*.246(4928):358-63.
221. Rajavashisth TB, Taylor AK, Andalibi A, Svenson KL, Lusk AJ (1989) Identification of a zinc finger protein that binds to the sterol regulatory element. *Science*.245(4918):640-3.
222. Ramsey JE, Daugherty MA, Kelm RJ, Jr. (2007) Hydrodynamic studies on the quaternary structure of recombinant mouse Purobeta. *J Biol Chem*.282(3):1552-60.
223. Rasimas JJ, Kar SR, Pegg AE, Fried MG (2007) Interactions of human O6-alkylguanine-DNA alkyltransferase (AGT) with short single-stranded DNAs. *J Biol Chem*.282(5):3357-66.
224. Rasimas JJ, Pegg AE, Fried MG (2003) DNA-binding mechanism of O6-alkylguanine-DNA alkyltransferase. Effects of protein and DNA alkylation on complex stability. *J Biol Chem*.278(10):7973-80.

225. Reddy S, Ozgur K, Lu M, Chang W, Mohan SR, Kumar CC, Ruley HE (1990) Structure of the human smooth muscle alpha-actin gene. Analysis of a cDNA and 5' upstream region. *J Biol Chem.*265(3):1683-7.
226. Regan CP, Adams PJ, Madsen CS, Owens GK (2000) Molecular mechanisms of decreased smooth muscle differentiation marker expression after vascular injury. *J Clin Invest.*106:1139-47.
227. Rich A, Nordheim A, Wang AH (1984) The chemistry and biology of left-handed Z-DNA. *Annual review of biochemistry.*53:791-846.
228. Rippe K (1997) Analysis of protein-DNA binding at equilibrium. *BIF Futura.*12:20-6.
229. Ritchie S, Boyd FM, Wong J, Bonham K (2000) Transcription of the human c-*Src* promoter is dependent on Sp1, a novel pyrimidine binding factor SPy, and can be inhibited by triplex-forming oligonucleotides. *J Biol Chem.*275(2):847-54.
230. Robertson RM, Laib S, Smith DE (2006) Diffusion of isolated DNA molecules: dependence on length and topology. *Proc Natl Acad Sci U S A.*103(19):7310-4.
231. Ronnov-Jessen L, Petersen OW (1996) A function for filamentous  $\alpha$ -smooth muscle actin: retardation of motility in fibroblasts. *J Cell Biol.*134:67-80.
232. Ross R (1999) Atherosclerosis - an inflammatory disease. *New England J Med.*340:115-25.
233. Ross R (1999) Atherosclerosis is an inflammatory disease. *American heart journal.*138(5 Pt 2):S419-20.

234. Russell AP, Holleman DS (1974) The thermal denaturation of DNA: average length and composition of denatured areas. *Nucleic Acids Res.*1(8):959-78.
235. Rustighi A, Tessari MA, Vascotto F, Sgarra R, Giancotti V, Manfioletti G (2002) A polypyrimidine/polypurine tract within the Hmga2 minimal promoter: a common feature of many growth-related genes. *Biochemistry.*41(4):1229-40.
236. Ruyechan WT, Wetmur JG (1975) Studies on the cooperative binding of the Escherichia coli DNA unwinding protein to single-stranded DNA. *Biochemistry.*14(25):5529-34.
237. Sadakata T, Kuo C, Ichikawa H, Nishikawa E, Niu SY, Kumamaru E, Miki N (2000) Puralpha, a single-stranded DNA binding protein, suppresses the enhancer activity of cAMP response element (CRE). *Brain research.*77(1):47-54.
238. Safak M, Gallia GL, Khalili K (1999) Reciprocal interaction between two cellular proteins, Puralpha and YB-1, modulates transcriptional activity of JCVCY in glial cells. *Mol Cell Biol.*19(4):2712-23.
239. SantaLucia J, Jr. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci U S A.*95(4):1460-5.
240. Sappino AP, Masouye I, Saurat JH, Gabbiani G (1990) Smooth muscle differentiation in scleroderma fibroblastic cells. *The American journal of pathology.*137(3):585-91.
241. Sappino AP, Schurch W, Gabbiani G (1990) Differentiation repertoire of fibroblastic cells: expression of cytoskeletal proteins as marker of phenotypic

modulations. *Laboratory investigation; a journal of technical methods and pathology*.63(2):144-61.

242. Sartore S, Chiavegato A, Faggin E, Franch R, Puato M, Ausoni S, Pauletto P (2001) Contribution of adventitial fibroblasts to neointima formation and vascular remodeling. *Circ Res*.89:1111-21.

243. Schildmeyer LA, Braun R, Taffet G, Debiase M, Burns AE, Bradley A, Schwartz RJ (2000) Impaired vascular contractility and blood pressure homeostasis in the smooth muscle  $\alpha$ -actin null mouse. *FASEB J*.14:2213-20.

244. Schneider DJ, Absher PM, Ricci MA (1997) Dependence of augmentation of arterial endothelial cell expression of plasminogen activator inhibitor type 1 by insulin on soluble factors released from vascular smooth muscle cells. *Circulation*.96(9):2868-76.

245. Schneider TD (2002) Consensus sequence Zen. *Applied bioinformatics*.1(3):111-9.

246. Schneider TD, Stephens RM (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res*.18(20):6097-100.

247. Schulick AH, Taylor AJ, Zuo W, Qiu CB, Dong G, Woodward RN, Agah R, Roberts AB, Virmani R, Dichek DA (1998) Overexpression of transforming growth factor beta1 in arterial endothelium causes hyperplasia, apoptosis, and cartilaginous metaplasia. *Proc Natl Acad Sci U S A*.95(12):6983-8.

248. Schwartz SM, Stemerman MB, Benditt EP (1975) The aortic intima. II. Repair of the aortic lining after mechanical denudation. *The American journal of pathology*.81(1):15-42.

249. Sen S, Lahiri A, Majumdar R (1992) Melting characteristics of highly supercoiled DNA. *Biophys Chem.*42(3):229-34.
250. Senear DF, Brenowitz M (1991) Determination of binding constants for cooperative site-specific protein-DNA interactions using the gel mobility-shift assay. *J Biol Chem.*266(21):13661-71.
251. Senear DF, Brenowitz M, Shea MA, Ackers GK (1986) Energetics of cooperative protein-DNA interactions: comparison between quantitative deoxyribonuclease footprint titration and filter binding. *Biochemistry.*25(23):7344-54.
252. Serini G, Bochaton-Piallat ML, Ropraz P, Geinoz A, Borsi L, Zardi L, Gabbiani G (1998) The fibronectin domain ED-A is crucial for myofibroblastic phenotype induction by transforming growth factor-beta1. *J Cell Biol.*142(3):873-81.
253. Serini G, Gabbiani G (1999) Mechanisms of myofibroblast activity and phenotypic modulation. *Experimental cell research.*250(2):273-83.
254. Shah NM, Groves AK, Anderson DJ (1996) Alternative neural crest cell fates are instructively promoted by TGF $\beta$  superfamily members. *Cell.*85:331-43.
255. Shelley CS, Da Silva N, Teodoridis JM (2001) During U937 monocytic differentiation repression of the CD43 gene promoter is mediated by the single-stranded DNA binding protein Pur alpha. *Br J Haematol.*115(1):159-66.
256. Shelley CS, Teodoridis JM, Park H, Farokhzad OC, Bottinger EP, Arnaout MA (2002) During differentiation of the monocytic cell line U937, Pur alpha mediates induction of the CD11c beta 2 integrin gene promoter. *J Immunol.*168(8):3887-93.
257. Sheterline P, Sparrow JC (1994) Actin. *Protein Profile.*1:1-121.

258. Shimizu RT, Blank RS, Jervis R, Lawrenz-Smith SC, Owens GK (1995) The smooth muscle alpha-actin gene promoter is differentially regulated in smooth muscle versus non-smooth muscle cells. *J Biol Chem.*270(13):7631-43.
259. Shimotai Y, Minami H, Saitoh Y, Onodera Y, Mishima Y, Kelm RJ, Jr., Tsutsumi K (2006) A binding site for Pur alpha and Pur beta is structurally unstable and is required for replication in vivo from the rat aldolase B origin. *Biochem Biophys Res Commun.*340(2):517-25.
260. Singleton CK, Klysik J, Stirdivant SM, Wells RD (1982) Left-handed Z-DNA is induced by supercoiling in physiological ionic conditions. *Nature.*299(5881):312-6.
261. Sinha S, Hoofnagle MH, Kingston PA, McCanna ME, Owens GK (2004) Transforming growth factor- $\beta$  signaling contributes to development of smooth muscle cells from embryonic stem cells. *Am J Physiol Cell Physiol.*287:C1560-C8.
262. Skowronski J, Furtak K, Klysik J, Panusz H, Plucienniczak A (1978) The 1360 bp long basic repeat unit of calf satellite I DNA contains GC rich nucleus of about 140 bp. *Nucleic Acids Res.*5(11):4077-85.
263. Smidt MP, Russchen B, Snippe L, Wijnholds J, Ab G (1995) Cloning and characterisation of a nuclear, site specific ssDNA binding protein. *Nucleic Acids Res.*23(13):2389-95.
264. Spencer JA, Baron MH, Olson EN (1999) Cooperative transcriptional activation by serum response factor and the high mobility group protein SSRP1. *J Biol Chem.*274(22):15686-93.

265. Stacey DW, Hitomi M, Kanovsky M, Gan L, Johnson EM (1999) Cell cycle arrest and morphological alterations following microinjection of NIH3T3 cells with Puro $\alpha$ . *Oncogene*.18:4254-61.
266. Stafford III WF (1992) Boundary analysis in sedimentation transport experiments: a procedure for obtaining sedimentation coefficient distributions using the time derivative of the concentration profile. *Anal Biochem*.203(2):295-301.
267. Stafford III WF (1994) Sedimentation Boundary Analysis of Interacting Systems: Use of the Apparent Sedimentation Coefficient Distribution Function. In: *Modern Analytical Ultracentrifugation*, Schuster T, Laue TM, (eds.), pp. 119-37, Birkhauser: Boston.
268. Stafford WF, Sherwood PJ (2004) Analysis of heterologous interacting systems by sedimentation velocity: curve fitting algorithms for estimation of sedimentation coefficients, equilibrium and kinetic constants. *Biophys Chem*.108(1-3):231-43.
269. Stenina OI, Poptic EJ, DiCorleto PE (2000) Thrombin activates a Y box-binding protein (DNA-binding protein B) in endothelial cells. *J Clin Invest*.106(4):579-87.
270. Stoflet ES, Schmidt LJ, Elder PK, Korf GM, Foster DN, Strauch AR, Getz MJ (1992) Activation of a muscle-specific actin gene promoter in serum-stimulated fibroblasts. *Mol Biol Cell*.3(10):1073-83.
271. Subramanian SV, Kelm RJ, Jr., Polikandriotis JA, Orosz CG, Strauch AR (2002) Reprogramming of vascular smooth muscle  $\alpha$ -actin gene expression as an early indicator of dysfunctional remodeling following heart transplant. *Cardiovasc Res*.54:539-48.

272. Subramanian SV, Polikandriotis JA, Kelm RJ, Jr., David JJ, Orosz CG, Strauch AR (2004) Induction of vascular smooth muscle  $\alpha$ -actin gene transcription in transforming growth factor  $\beta$ -1-activated myofibroblasts mediated by dynamic interplay between the Pur repressor proteins and Sp1/Smad coactivators. *Mol Biol Cell*.15(10):4532-43.
273. Suck D, Oefner C (1986) Structure of DNase I at 2.0 Å resolution suggests a mechanism for binding to and cutting DNA. *Nature*.321(6070):620-5.
274. Sun D, Guo K, Rusche JJ, Hurley LH (2005) Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents. *Nucleic Acids Res*.33(18):6070-80.
275. Sun S, Stoflet ES, Cogan JG, Strauch AR, Getz MJ (1995) Negative regulation of the vascular smooth muscle  $\alpha$ -actin gene in fibroblasts and myoblasts: disruption of enhancer function by sequence-specific single-stranded DNA-binding proteins. *Mol Cell Biol*.15:2429-36.
276. Sutton DH, Conn GL, Brown T, Lane AN (1997) The dependence of DNase I activity on the conformation of oligodeoxynucleotides. *Biochem J*.321 ( Pt 2):481-6.
277. Swamynathan SK, Nambiar A, Guntaka RV (1998) Role of single-stranded DNA regions and Y-box proteins in transcriptional regulation of viral and cellular genes. *Faseb J*.12(7):515-22.

278. Swartz EA, Johnson AD, Owens GK (1998) Two MCAT elements of the SM alpha-actin promoter function differentially in SM vs. non-SM cells. *Am J Physiol.*275(2 Pt 1):C608-18.
279. Takagi T (1990) Application of low-angle laser light scattering detection in the field of biochemistry. *J Chromatography.*506:409-16.
280. Takai T, Nishita Y, Iguchi-Ariga SM, Ariga H (1994) Molecular cloning of MSSP-2, a c-myc gene single-strand binding protein: characterization of binding specificity and DNA replication activity. *Nucleic Acids Res.*22(25):5576-81.
281. Takeji M, Moriyama T, Oseto S, Kawada N, Hori M, Imai E, Miwa T (2006) Smooth muscle alpha-actin deficiency in myofibroblasts leads to enhanced renal tissue fibrosis. *J Biol Chem.*281(52):40193-200.
282. Tamrazi A, Carlson KE, Daniels JR, Hurth KM, Katzenellenbogen JA (2002) Estrogen receptor dimerization: ligand binding regulates dimer affinity and dimer dissociation rate. *Mol Endocrinol.*16(12):2706-19.
283. Tang WW, Ulich TR, Lacey DL, Hill DC, Qi M, Kaufman SA, Van GY, Tarpley JE, Yee JS (1996) Platelet-derived growth factor-BB induces renal tubulointerstitial myofibroblast formation and tubulointerstitial fibrosis. *The American journal of pathology.*148(4):1169-80.
284. Tataurov AV, You Y, Owczarzy R (2008) Predicting ultraviolet spectrum of single stranded and double stranded deoxyribonucleic acids. *Biophys Chem.*133(1-3):66-70.

285. Tewari DS, Cook DM, Taub R (1989) Characterization of the promoter region and 3' end of the human insulin receptor gene. *J Biol Chem.*264(27):16238-45.
286. Thiele BJ, Doller A, Kahne T, Pregla R, Hetzer R, Regitz-Zagrosek V (2004) RNA-binding proteins heterogeneous nuclear ribonucleoprotein A1, E1, and K are involved in post-transcriptional control of collagen I and III synthesis. *Circ Res.*95(11):1058-66.
287. Tomasek JJ, Gabbiani G, Hinz B, Chaponnier C, Brown RA (2002) Myofibroblasts and mechano-regulation of connective tissue remodelling. *Nature reviews.*3(5):349-63.
288. Tomonaga T, Levens D (1996) Activating transcription from single stranded DNA. *Proc Natl Acad Sci U S A.*93(12):5830-5.
289. Tretiakova A, Gallia GL, Shcherbik N, Jameson B, Johnson EM, Amini S, Khalili K (1998) Association of Puralpha with RNAs homologous to 7 SL determines its binding ability to the myelin basic protein promoter DNA sequence. *J Biol Chem.*273(35):22241-7.
290. Tretiakova A, Otte J, Croul SE, Kim JH, Johnson EM, Amini S, Khalili K (1999) Association of JC virus large T antigen with myelin basic protein transcription factor (MEF-1/Puralpha) in hypomyelinated brains of mice transgenically expressing T antigen. *J Virol.*73(7):6076-84.
291. Tretiakova A, Stepkowski A, Johnson EM, Khalili K, Amini S (1999) Regulation of myelin basic protein gene transcription by Sp1 and Puralpha: evidence for association of Sp1 and Puralpha in brain. *J Cell Physiol.*181(1):160-8.

292. Tretyachenko-Ladokhina V, Ross JB, Senechal DF (2002) Thermodynamics of E. coli cytidine repressor interactions with DNA: distinct modes of binding to different operators suggests a role in differential gene regulation. *J Mol Biol.*316(3):531-46.
293. Viglasky V, Antalik M, Adamcik J, Podhradsky D (2000) Early melting of supercoiled DNA topoisomers observed by TGGE. *Nucleic Acids Res.*28(11):E51.
294. Voloshin ON, Shlyakhtenko LS, Lyubchenko Yu L (1989) Localization of melted regions in supercoiled DNA. *FEBS letters.*243(2):377-80.
295. Wang H, Peters GA, Zeng X, Tang M, Ip W, Khan SA (1995) Yeast two-hybrid system demonstrates that estrogen receptor dimerization is ligand-dependent in vivo. *J Biol Chem.*270(40):23322-9.
296. Wang J, Niu W, Nikiforov Y, Naito S, Chernausek S, Witte D, LeRoith D, Strauch A, Fagin JA (1997) Targeted overexpression of IGF-I evokes distinct patterns of organ remodeling in smooth muscle cell tissue beds of transgenic mice. *J Clin Invest.*100(6):1425-39.
297. Wang SX, Elder PK, Zheng Y, Strauch AR, Kelm RJ, Jr. (2005) Cell cycle-mediated regulation of smooth muscle  $\alpha$ -actin gene transcription in fibroblasts and vascular smooth muscle cells involves multiple adenovirus E1A-interacting cofactors. *J Biol Chem.*280:6204-14.
298. Wang Z, Lin XH, Qiu QQ, Deuel TF (1992) Modulation of transcription of the platelet-derived growth factor A-chain gene by a promoter region sensitive to S1 nuclease. *J Biol Chem.*267(24):17022-31.

299. Wang Z, Rao PJ, Castresana MR, Newman WH (2005) TNF- $\alpha$  induces proliferation or apoptosis in human saphenous vein smooth muscle cells depending on phenotype. *Am J Physiol Heart Circ Physiol*.288:H293-H301.
300. Wang Z, Wang DZ, Pipes GC, Olson EN (2003) Myocardin is a master regulator of smooth muscle gene expression. *Proc Natl Acad Sci U S A*.100(12):7129-34.
301. Wang ZY, Lin XH, Nobuyoshi M, Qui QQ, Deuel TF (1992) Binding of single-stranded oligonucleotides to a non-B-form DNA structure results in loss of promoter activity of the platelet-derived growth factor A-chain gene. *J Biol Chem*.267(19):13669-74.
302. Warnhoff BR, Bowles DK, McDonald OG, Sinha S, Somlyo AP, Somlyo AV, Owens GK (2004) L-type voltage-gated Ca<sup>2+</sup> channels modulate expression of smooth muscle differentiation marker genes via a Rho kinase/myocardin/Srf-dependent mechanism. *Circ Res*.95:406-14.
303. Wei Q, Miskimins WK, Miskimins R (2005) Stage-specific expressions of myelin basic protein in oligodendrocytes involves Nkx2.2-mediated repression that is relieved by the Sp1 transcription factor. *J Biol Chem*.280:16284-94.
304. Weintraub H, Davis R, Tapscott S, Thayer M, Krause M, Benezra R, Blackwell TK, Turner D, Rupp R, Hollenberg S, et al. (1991) The myoD gene family: nodal point during specification of the muscle cell lineage. *Science*.251(4995):761-6.
305. Weiss JN (1997) The Hill equation revisited: uses and misuses. *Faseb J*.11(11):835-41.

306. Wen J, Arakawa T, Philo JS (1996) Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. *Anal Biochem.*240(2):155-66.
307. Wilkins RC, Lis JT (1999) DNA distortion and multimerization: novel functions of the glutamine-rich domain of GAGA factor. *J Mol Biol.*285(2):515-25.
308. Wilson WW (2003) Light scattering as a diagnostic for protein crystal growth--a practical approach. *J Struct Biol.*142(1):56-65.
309. Woodcock-Mitchell J, Mitchell JJ, Low RB, Kieny M, Sengel P, Rubbia L, Skalli O, Jackson B, Gabbiani G (1988) Alpha-smooth muscle actin is transiently expressed in embryonic rat cardiac and skeletal muscles. *Differentiation; research in biological diversity.*39(3):161-6.
310. Wortman MJ, Johnson EM, Bergemann AD (2005) Mechanism of DNA binding and localized strand separation by Pur $\alpha$  and comparison with Pur family member, Pur $\beta$ . *Biochem Biophys Acta.*1743:64-78.
311. Wortman MJ, Krachmarov CP, Kim JH, Gordon RG, Chepenik LG, Brady JN, Gallia GL, Khalili K, Johnson EM (1999) Interaction of HIV-1 Tat with Pur $\alpha$  in nuclei of human glial cells: characterization of RNA-mediated protein-protein binding. *J Cell Biochem.*77:65-74.
312. Wu J, Lee C, Yokom D, Jiang H, Cheang MC, Yorida E, Turbin D, Berquin IM, Mertens PR, Iftner T, Gilks CB, Dunn SE (2006) Disruption of the Y-box binding protein-1 results in suppression of the epidermal growth factor receptor and HER-2. *Cancer research.*66(9):4872-9.

313. Yagil G (1991) Paranemic structures of DNA and their role in DNA unwinding. *Crit Rev Biochem Mol Biol.*26(5-6):475-559.
314. Yoshida T (2008) MCAT elements and the TEF-1 family of transcription factors in muscle development and disease. *Arterioscler Thromb Vasc Biol.*28(1):8-17.
315. Yoshida T, Owens GK (2005) Molecular determinants of vascular smooth muscle cell diversity. *Circ Res.*96:280-91.
316. Zambrano N, De Renzis S, Minopoli G, Faraonio R, Donini V, Scaloni A, Cimino F, Russo T (1997) DNA-binding protein Pur alpha and transcription factor YY1 function as transcription activators of the neuron-specific FE65 gene promoter. *Biochem J.*328 ( Pt 1):293-300.
317. Zeng L-H, Okamura K, Tanaka H, Miki N, Kuo C-H (2005) Concomitant translocation of Pur $\alpha$  with its binding proteins (PurBPs) from nuclei to cytoplasm during neuronal development. *Neuroscience Res.*51:105-9.
318. Zhang A, David JJ, Subramanian SV, Liu X, Fuerst MD, Zhao X, Leier CV, Orosz CG, Kelm RJ, Jr., Strauch AR (2008) Serum response factor neutralizes Pur alpha- and Pur beta-mediated repression of the fetal vascular smooth muscle alpha-actin gene in stressed adult cardiomyocytes. *Am J Physiol Cell Physiol.*294(3):C702-14.
319. Zhang A, Liu X, Cogan JG, Fuerst MD, Polikandriotis JA, Kelm RJ, Jr., Strauch AR (2005) YB-1 coordinates vascular smooth muscle alpha-actin gene activation by transforming growth factor beta1 and thrombin during differentiation of human pulmonary myofibroblasts. *Mol Biol Cell.*16(10):4931-40.

320. Zhang H, Facemire CS, Banes AJ, Faber JE (2002) Different  $\alpha$ -adrenoreceptors mediate migration of vascular smooth muscle cells and adventitial fibroblasts in vitro. *Am J Physiol Heart Circ Physiol*.282:H2364-H70.
321. Zhang M, Smith EP, Kuroda H, Banach W, Chernausek SD, Fagin JA (2002) Targeted expression of a protease-resistant IGFBP-4 mutant in smooth muscle of transgenic mice results in IGFBP-4 stabilization and smooth muscle hypotrophy. *J Biol Chem*.277(24):21285-90.
322. Zhang Q, Pedigo N, Shenoy S, Khalili K, Kaetzel DM (2005) Pur $\alpha$  activates PDGF-A gene transcription via interactions with a G-rich, single-stranded region of the promoter. *Gene*.348:25-32.
323. Zijderveld DC, van der Vliet PC (1994) Helix-destabilizing properties of the adenovirus DNA-binding protein. *J Virol*.68(2):1158-64.

This document was created with Win2PDF available at <http://www.win2pdf.com>.  
The unregistered version of Win2PDF is for evaluation or non-commercial use only.  
This page will not be added after purchasing Win2PDF.