Characterization of regulatory P sequences with the capacity for dual roles in directing expression of the human growth hormone chorionic somatomammotropin gene family.

by

Lisa Daly Norquay

A thesis
submitted to the Faculty of Graduate Studies
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Physiology
Faculty of Medicine
University of Manitoba
Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION PAGE

Characterization of Regulatory P Sequences with the Capacity for Dual Roles in Directing Expression of the Human Growth Hormone Chorionic Somatomammotropin Gene Family

BY

Lisa Daly Norquay

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Dedicated to my daughter Abigail Grace

ABSTRACT

In a multicellular organism, where almost every cell carries identical genetic information, different cell types are observed that display diverse structures and functions. A fundamental question in biology therefore centres around the mechanisms that regulate how and when the genetic information is used. As a result, regulation of differential gene expression patterns is one of the most basic and critical processes in the development and function of multicellular organisms.

The human growth hormone (GH) / chorionic somatomammotropin (CS) gene family represents a unique model system for examining mechanisms that regulate differential gene expression. The GH/CS gene family consists of five tandemly arranged genes contained within a single locus on chromosome 17. These genes include GH-N, which is expressed almost exclusively in the anterior pituitary somatotrophs, and GH-V and the CS genes (CS-L, CS-A, CS-B), which are expressed in placental syncytiotrophoblasts. The differential gene expression pattern of the GH/CS family is quite remarkable, given the fact that these genes are believed to have evolved from a single ancestral gene by duplication, and thus are 94-99% similar in both their coding and immediate flanking regulatory sequences.

There are several regulatory elements that direct expression of the GH/CS locus. Some of these elements are particularly intriguing, as they have the capacity to participate in both placental and pituitary expression mechanisms. The P sequences are located approximately 2 kilobases upstream of the placental members of the GH/CS family. These sequences have been implicated in dual roles; repression of gene activity in pituitary cells, and enhancement of gene expression in the placenta. This thesis seeks to characterize the P sequences and their role in regulation of the GH/CS locus, primarily through identification of the P sequence factors (PSFs).

Previous characterization of the P sequences localized pituitary repressor activity to a 263 base pair fragment (263P), and identified two regions of protein binding, termed P sequence element (PSE)-A and PSE-B. To identify candidate pituitary PSFs *in vitro*, structural assays (nuclease protection and electrophoretic mobility shift competition) to identify protein-DNA interactions, and functional assays (transient gene transfer) to determine effects on gene expression, were used with both the PSEs and 263P. Through the use of these approaches, PSE-A was found to be a composite element in pituitary cells, capable of mutually exclusive RFX1 and NF-1 binding, and the analysis of PSE-B implicated members of the NF-1 family as PSF-B. When the pituitary repressor complex was evaluated in terms of larger fragments that contained both PSE-A and PSE-B, a model was generated whereby at least two separate complexes have the potential to form on P sequence fragments; an NF-1 repressor complex, and a non-functional complex that contains RFX1. In addition, 263P repressor function in pituitary cells was linked to the transcription factor Pit-1, and the potential for novel interactions between Pit-1 and the PSFs was investigated.

Gene transfer experiments also provided the first evidence for significant *in vitro* P sequence function in placental cells, and revealed the capacity for functional cross-talk between 263P and the CS-B downstream enhancer element. Furthermore, *in vitro* structural experiments with nuclear proteins from placental cells implicated RFX1 and the NF-1 family as candidate placental PSFs.

To investigate the possible association of these candidate factors with P sequences in the nucleus, a chromatin immunoprecipitation technique was developed for use with human tissue samples. This confirmed the association of both NF-1 and RFX1 with P sequences in human pituitary, but not term placenta, samples. This technique was also used to extend the analysis of histone hyperacetylation patterns in the GH/CS locus. As a result of these studies, the human GH/CS locus is identified as a novel target of the NF-1 and RFX transcription factor families, and when taken together, these observations allow for an extension of the GH/CS locus regulatory model in both the pituitary and the placenta.

ACKNOWLEDGEMENTS

I would like to start out by thanking my exceptional mentor Dr. Peter A. Cattini. Peter, there really aren't adequate words to describe the enormous impact you have had on my life, except to say that I wouldn't have reached this point without you in my corner. Any accomplishment of mine is also yours. You continually challenge me to become a better scientist and a better person. I hope that during my time in your lab I have managed to do a little of both. Thank you so much for all of your time and effort, your frequently tested patience, and unending support. I think we communicate beautifully for people from two different planets.

I would like to thank my advisory committee, Dr. Mary Lynn Duckworth, Dr. Janice G. Dodd and Dr. R. Daniel Gietz. I am especially grateful to Mary Lynn and Janice for their guidance and encouragement. You have both been so accessible and supportive, time and time again. Thank you for everything. I'll miss our lunches. I would also like to thank my external examiner Dr. Robert J. Matusik for evaluating my thesis and braving a trip to Winnipeg in January for my examination. I look forward to many stimulating conversations at future endocrine meetings (and maybe a dance or two as well).

To my fellow members of the Cattini Lab, Margaret Bock (Marge, my friend and frequent saviour) Yan Jin (Yan Beautiful Yan), Rama Mohan Surabhi, Aristides Lytras, Farah Sheikh, David Sontag, Karen Detillieux, Sarah Jimenez, Xiaoyang Yang (My Yangyang), Kevin Oomah, Jamit Dhaliwal, Scott Gregoire, and Alina Sofronescu. When I look at this list I am flooded with so many memories. I feel like every one of these names should be followed by an exclamation mark. In some way or another you have all contributed to my thesis. I am grateful for your intellectual support and technical assistance. Most importantly though, I thank you for the hundreds of ways you have each contributed to my life and my experience here. In my heart I will always be a member of 'The Cattini Lab', who's proud enough to shout it out from a boat in the middle of the Red River.

I would like to thank Gail McIndless and Judy Olfert in the Physiology office for all of their help. A big thank you to Patricia Sheppard, a wonderful collaborator and friend. Thank you to Dr. James Davie and Dr. Virginia Spencer for their assistance in setting up the chromatin immunoprecipitation assays. Thank you to Dr. N. Tanese and Dr. M. Imagawa for NF-1 reagents, and Dr. W. Reith for RFX1 reagents. In addition I would like to acknowledge the many funding agencies that have granted me their support, the Canadian Institutes of Health Research, University of Manitoba Faculty of Graduate Studies, Manitoba Health Research Council, Women's Health Research Foundation of Canada, Deer Lodge Hospital Staff Association Memorial Fund, University of Manitoba Students' Union, Graduate Students' Association, and the University of Manitoba Alumni Association, as well as the University of Manitoba major student research awards committee for honouring me with the 2004 Apotex Fermentation Inc. Molecular Biology and E.L Drewry Memorial awards.

Last but certainly not least. Thank you to my loving family. You not only believe that I can do anything, you continually tell me so. Thank you for being so supportive and proud. To my husband Wade. Anyone who knows us also knows that you most certainly are my better half. You have been my calm refuge through the crazy ups and downs of my Ph.D. Your faith in me inspires me to be the person you believe I am capable of being. I love you. And this time I think I really did break a lead.

CONTENTS

PAGE

ABSTRACT ii
ACKNOWLEDGEMENTS iv
LIST OF FIGURES xvii
LIST OF TABLES xxi
ABBREVIATIONS xxii

CHAPTER 1: INTRODUCTION 1

1.1 **Regulation of differential gene expression patterns.** 1

1.1.1 Overcoming the inactive state of the genome: derepression. 1
1.1.1a Chromatin structure 1
1.1.1b Increasing DNA accessibility 4
1.1.1c Modifications of nucleosomes 5

1.1.2 The activation of gene expression. 6
1.1.2a DNA regulatory elements 7
1.1.2b Locus control regions are enhancer elements. 8
1.1.2c *Trans*-acting factors 9

1.1.3 Mechanisms of transcriptional repression. 10

1.2 **The human growth hormone (GH) / chorionic somatomammotropin (CS) family as a model system for examining differential gene expression.** 11

1.2.1 Transcriptional regulation of the GH/CS locus in the pituitary. 14

1.2.1a Activity of the GH/CS 5′-flanking sequences *in vitro.* 14

1.2.1b Expression of the GH/CS locus in the pituitary *in vivo.* 15

1.2.1c The lack of placental gene expression in the pituitary. 17

1.2.1d The P sequences 19

1.2.2 Transcriptional regulation of the GH/CS locus in the placenta. 20

1.2.2a Regulation of the placental promoters *in vitro.* 20

1.2.2b The CS genes have placental enhancers in their 3′-flanking DNA sequences. 20

1.2.2c Expression of the GH/CS locus in the placenta *in vivo.* 21

1.2.2d The P sequences 23

1.3 Research objectives 24

CHAPTER 2: MATERIALS AND METHODS 26

2.1 Materials 26

2.1.1 Antibodies 26

2.1.1a c-myc 26

2.1.1b Prostate dorsolateral protein (DLP) 26

2.1.1c Hyperacetylated histone H4 27
2.1.1d NF-1 27
2.1.1e Pit-1 27
2.1.1f RFX1 28
2.1.1g RFX2 28
2.1.1h RFX3 28

2.1.2 Cell Lines 28

2.1.3 Oligonucleotides 29

2.2 Methods 31

2.2.1 Chromatin immunoprecipitation 31
2.2.1a Isolation of intact nuclei and formaldehyde cross-linking. 31
2.2.1b Nuclear lysis and sonication of the chromatin. 32
2.2.1c Immunoprecipitation (IP) and DNA isolation. 33
2.2.1d Analysis of the ChIP DNA. 35

2.2.2 Gene Transfer 35
2.2.2a Transfection by calcium phosphate/DNA precipitation 35
2.2.2b Luciferase assays 36
2.2.2c Chloramphenicol acetyltransferase (CAT) assay 36

2.2.3 Identification of putative transcription factor binding sites. 37

2.2.4 Immunoprecipitation and Protein (western) blotting 37
2.2.4a Immunoprecipitation 37
2.2.4b Protein (western) blotting 39

2.2.5 In vitro binding assays 39

2.2.5a Preparation of nuclear extracts 39

2.2.5b Electrophoretic mobility shift assay (EMSA) 40

2.2.5c Nuclease protection 40

2.2.6 Plasmid DNA 41

2.2.6a Plasmid constructions 42

2.2.6b Preparation of competent bacterial cells and bacterial transformation 43

2.2.6c Isolation of plasmid DNA from bacteria 43

2.2.6d Sequencing of plasmid DNA 44

2.2.7 RNA 44

2.2.7a RNA isolation 44

2.2.7b PolyA$^+$-enrichment 45

2.2.7c RNA (northern) blotting 45

2.2.7d NF-1 specific probes 45

2.2.8 Statistical analysis 46

CHAPTER 3: IDENTIFICATION OF P SEQUENCE FACTORS (PSFs) *IN VITRO* 47

3.1 Functional characterization of 263P *in vitro*. 47

3.1.1 P sequences repress CS-A promoter activity in pituitary GC cells in the absence of the CS-B downstream enhancer region. 47

3.1.2 Detection of significant P sequence function in placental cell lines. 48

3.2 Identification of candidate P sequence factors (PSFs) using PSE-A and PSE-B. 50

3.2.1 PSE-A is a composite element with the capacity for mutually exclusive regulatory factor X-1 (RFX1) and nuclear factor-1 (NF-1) binding. 55

3.2.1a PSE-A contains putative binding sites for several transcription factors. 55

3.2.1b A truncated PSE-A oligonucleotide (PSE-A3) retains PSE-A EMSA complexes. 58

3.2.1c Scanning mutations define the core region within PSE-A3 for pituitary complex formation. 59

3.2.1d RFX1 binds to PSE-A3 in pituitary GC nuclear extracts. 62

3.2.1e Direct binding of NF-1 to PSE-A3 was detected in the absence of RFX1 association. 63

3.2.1f PSE-A3 can repress the activity of the CS-A promoter in GC cells *in vitro*. 66

3.2.1g RFX1 is a candidate PSE-A binding protein in placental as well as pituitary cells. 68

3.2.2 The PSE-B oligonucleotide is an NF-1 site. 71

3.2.2a NF-1 proteins can bind to PSE-B4 in pituitary GC nuclear extracts. 71

3.2.2b NF-1 binding is responsible for PSE-B4 repressor activity. 73

3.2.2c Minor variations outside of the NF-1 core binding region in PSE-B4 affect functional activity. 77

3.2.2d Detection of NF-1A, NF-1C and NF-1X transcripts in rat pituitary GC cells. 82

3.2.2e The NF-1 family has the capacity to interact with RFX1. 85

3.2.2f Direct binding of NF-1 to PSE-B4 in placental nuclear extracts. 89

3.3 The P sequence complex in pituitary GC cells. 92

3.3.1 An in vitro P sequence complex that contains RFX1 does not repress CS-A promoter activity in pituitary GC cells. 92

3.3.1a Direct binding of RFX1 to the 103P element. 92

3.3.1b The 103P element does not repress the activity of the CS-A promoter in pituitary GC cells. 94

3.3.2 NF-1, but not RFX1 binding is detected with 263P. 96

3.3.2a Competition with NF-1 elements disrupts protein binding to the 263P fragment. 96

3.3.2b The EF-C/MDBP RFX element does not compete for protein binding to 263P. 98

3.3.2c An NF-1 half-site in PSE-A3 contributes to its ability to compete 263P protein binding. 100

3.3.3 Functional evidence in support of mutually exclusive NF-1 and RFX1 263P complexes. 101

3.3.3a The ability of 263P to repress CS-A promoter activity involves sequences at both the PSE-A/RFX and PSE-B/NF-1 sites. 102

3.3.3b Overexpression of NF-1 in GC cells increases the degree of CS-A promoter repression by 263P. 104

3.3.3c Overexpression of RFX1 in GC cells enhances luciferase activity independent of 263P. 106

3.4 Summary of Chapter 3 results. 108

CHAPTER 4: THE P SEQUENCE COMPLEX *IN SITU* 111

4.1 Development of the ChIP technique for analysis of human tissue samples. 111

4.1.1 Determination of the cross-linking time. 115

4.1.2 Optimization of the sonication conditions. 115

4.1.3 Determining the amount of chromatin input for the immunoprecipitations. 117

4.1.4 Specificity of the NF-1 and RFX1 antibodies for use in ChIP. 118

4.1.5 Analysis of the ChIP DNA 119

4.2 Analysis of histone H4 hyperacetylation of the GH/CS locus in human pituitary and term placenta tissue. 123

4.2.1 Histone H4 hyperacetylation of the GH/CS locus in human term placenta. 123

4.2.2 Histone H4 hyperacetylation of the GH/CS locus in human pituitary tissue. 124

4.3 NF-1 and RFX1 associate with 263P in human pituitary tissue. 127

4.4 The association of NF-1 and RFX1 with P sequences was not detected in human term placenta tissue. 129

4.5 Summary of Chapter 4 results 131

CHAPTER 5: CHARACTERIZATION OF A POTENTIAL PITUITARY REPRESSOR MECHANISM INVOLVING THE TRANSCRIPTION FACTOR PIT-1 134

5.1 263P repressor function is linked to Pit-1 activation of the CS-A promoter in pituitary GC cells. 134

5.1.1 263P enhances the activity of a minimal thymidine kinase promoter (Tkp). 134

5.1.2 263P represess Pit-1-mediated activity of the CS-A promoter in pituitary GC cells. 136

5.2 Potential interactions between Pit-1 and the PSFs. 139

5.2.1 EMSA and nuclease protection competitions suggest that Pit-1 has the capacity for interactions with the PSFs. 139

5.2.1a Pit-1 oligonucleotides compete for protein binding to P sequence elements. 140

5.2.1b Competition of Pit-1 binding to the CS-A promoter with an RFX element. 144

5.2.2 The capacity for PSF and Pit-1 interactions was not detected in the absence of DNA. 146

5.2.2a Pit-1 was not detected in immunoprecipitations of pituitary GC nuclear extracts with NF-1 or RFX1 antibodies. 146

5.2.2b Co-immunoprecipitation of Pit-1 with NF-1 or RFX1 was also not detected using overexpressed epitope-tagged proteins. 148

5.3 Summary of Chapter 5 results 151

CHAPTER 6: DISCUSSION 153

6.1 P sequence complexes 153

6.1.1 Interpretation of the in vitro and in vivo complexes. 153

6.1.1a The pituitary P sequence complex *in vitro*. 153

6.1.1b The pituitary P sequence complex *in situ*. 157

6.1.1c The placental P sequence complex. 158

6.1.2 The implications of NF-1 involvement. 161

6.1.2a Diversity of the NF-1 family. 162

6.1.2b The mechanisms of NF-1 activity. 164

6.1.2c NF-1 is a context dependant regulator. 165

6.1.3 The implications of RFX involvement in p sequence complexes. 167

6.1.3a The RFX family of transcription factors. 167

6.1.3b DNA binding in the RFX family. 167

6.1.3c RFX activity as an activator and repressor of transcription. 169

6.1.4 The possibility of additional PSFs. 169

6.1.4a Candidate PSFs 170

6.1.4b Future analysis of the candidate PSFs. 174

6.2 Potential mechanisms of pituitary repression through P sequences. 175

6.2.1 Repression of CS-A promoter activity in pituitary GC cells. 175

6.2.1a The 263P repressor complex in GC cells. 175

6.2.1b Potential mechanisms for *in vitro* repression. 177

6.2.2 Applying the in vitro analysis to potential mechanisms for in vivo repression. 179

6.2.2a Regulation of the GH/CS locus in the pituitary *in vivo.* 179

6.2.2b Potential mechanisms for repression of placental GH/CS gene expression in the pituitary *in vivo.* 180

6.3 Potential mechanisms for the involvement of P sequences in activated expression of the GH/CS locus in the placenta. 185

6.3.1 The in vitro system did not model P sequence enhancer activity in the placenta. 185

6.3.2 Potential mechanisms for P sequence enhancer activity in vivo. 186

6.3.2a Regulation of the GH/CS locus in the placenta *in vivo.* 186

6.3.2b Proposed P sequence mechanisms. 188

6.4 Future Directions 192

6.4.1 The pituitary system. 192

6.4.1a The contribution of 263P to a lack of placental gene expression in the pituitary. 192

6.4.1b The role of Pit-1. 193

6.4.1c Identification of additional PSFs. 194

6.4.2 *The placental system.* 194

6.4.2a Regulation of the GH/CS locus in the placenta. 194

6.4.3b Identification of the placental PSFs. 195

CHAPTER 7: REFERENCES 196

LIST OF FIGURES

PAGE

CHAPTER 1: INTRODUCTION

1.1 The nucleosome is the basic repetitive unit of chromatin. 2

1.2 The compaction of chromatin in the nucleus. 4

1.3 The human GH/CS locus on chromosome 17. 12

CHAPTER 3: IDENTIFICATION OF P SEQUENCE FACTORS (PSFs) *IN VITRO*

3.1 P sequences repress CS-A promoter activity in pituitary GC cells in the absence of the CS-B downstream enhancer region. 48

3.2 Functional effects of 263P on CS-A promoter activity can be detected in placental cells in the absence of the CS-B downstream enhancer region. 50

3.3 Possible interpretations of the 263P complex based on competition of the nuclease protection pattern. 52

3.4 Comparison of nuclease protection patterns generated with placental JAR and pituitary GC nuclear extract. 54

3.5 PSE-A3 contains putative binding sites for the NF-1 and RFX transcription factor families. 57

3.6 EMSA patterns of PSE-A3 and PSE-A5 probes with GC nuclear extract. 59

3.7 Specific PSE-A3 EMSA complexes with pituitary GC nuclear extract. 60

3.8 Scanning mutations define the core region within PSE-A3 for pituitary complex formation. 61

3.9 RFX1 from pituitary GC nuclear extract binds to PSE-A3. 63

3.10 Direct binding of NF-1 was not detected with a PSE-A3 probe. 64

3.11 The direct binding of NF-1 to PSE-A was only detected in the absence of RFX1 association. 66

3.12 Scanning mutations do not reveal a core sequence in PSE-A3 for the formation of placental EMSA complexes. 68

3.13	RFX1 associates with PSE-A3 in placental nuclear extracts.	70
3.14	Direct binding of NF-1 from pituitary GC nuclear extract to PSE-B4.	72
3.15	NF-1 is responsible for PSE-B repressor activity.	74
3.16	PSE-B4m does not bind NF-1.	76
3.17	Minor variations outside of the NF-1 core binding region in PSE-B4 do not disrupt NF-1 binding.	80
3.18	Minor variations outside of the NF-1 core binding region in PSE-B4 affect functional activity.	82
3.19	NF-1 transcripts were detected in pituitary GC cells by RNA blotting.	84
3.20	PSE-B4 can compete for RFX complexes from a PSE-A3 probe.	85
3.21	The NF-1 family has the capacity to interact with RFX1.	87
3.22	Interactions between endogenous NF-1 and RFX1 proteins are detected through co-immunoprecipitations of pituitary GC nuclear extract.	89
3.23	NF-1 from placental nuclear extracts associates with PSE-B4.	91
3.24	P sequences upstream of the CS-A gene.	93
3.25	Direct binding of RFX1, but not the NF-1 family, to 103P.	94
3.26	103P does not affect CS-A promoter activity in transiently transfected pituitary GC cells.	95
3.27	Competition with NF-1 elements disrupts protein binding to the 263P fragment.	97
3.28	The EF-C/MDBP RFX element does not compete for protein binding to the 263P fragment.	99
3.29	An NF-1 half-site in PSE-A3 contributes to its ability to compete for 263P protein binding.	101
3.30	The ability of 263P to repress CS-A promoter activity involves sequences at both the PSE-A/RFX1 and PSE-B/NF-1 sites.	103
3.31	Overexpression of NF-1 in GC cells increases the degree of CS-A promoter repression by 263P.	105

3.32 Overexpression of RFX1 in GC cells enhances luciferase activity independent of 263P. 107

CHAPTER 4: THE P SEQUENCE COMPLEX *IN SITU*

4.1 Isolated human placenta and pituitary nuclei. 113

4.2 The ChIP protocol requires assessment of several key steps. 114

4.3 Optimization of the sonication conditions for placenta and pituitary chromatin. 116

4.4 Chromatin input above 2U into the immunoprecipitations decreases the difference between the specific and background signals. 118

4.5 Specificity of the NF-1 and RFX1 antibodies for use in ChIP. 119

4.6 The intra-experimental and inter-experimental variations between PCR reactions are not statistically significant. 122

4.7 Histone H4 hyperacetylation of the GH/CS locus in human term placenta chromatin. 124

4.8 Histone H4 hyperacetylation of the GH/CS locus in human pituitary chromatin. 126

4.9 NF-1 and RFX1 associate with 263P in human pituitary tissue. 128

4.10 The association of NF-1 and RFX1 with P sequences was not detected in human term placenta tissue. 130

CHAPTER 5: CHARACTERIZATION OF A POTENTIAL PITUITARY REPRESSOR MECHANISM INVOLVING THE TRANSCRIPTION FACTOR PIT-1

5.1 263P enhances Tk promoter activity. 136

5.2 263P does not repress non-Pit-1 mediated activity of the CS-A promoter in pituitary GC cells. 139

5.3 Pit-1 oligonucleotides compete for 263P protein binding. 141

5.4 A Pit-1 oligonucleotide competes for PSE-A3 and PSE-B4

EMSA complexes. 143

5.5 An RFX element competes for Pit-1 binding to the CS-A promoter. 145

5.6 Pit-1 was not detected in the immunoprecipitations of pituitary GC nuclear extracts with NF-1 or RFX1 antibodies. 147

5.7 Co-immunoprecipitation of Pit-1 with NF-1 or RFX1 was not detected using overexpressed epitope-tagged proteins. 149

CHAPTER 6: DISCUSSION

6.1 Potential PSF complexes based on nuclease protection data. 154

6.2 Model of a potential *in vitro* P sequence complex. 154

6.3 Model of a second potential in vitro P sequence complex. 155

6.4 The sequence of 263P upstream of the PSE-B nuclease protection region. 170

6.5 The sequence of 263P downstream of the PSE-A nuclease protection region. 172

6.6 The pituitary repressor mechanism appears to result from inhibition of the Pit-1 promoter complex by the NF-1 263P repressor complex. 176

6.7 Spacing of nucleotides in Pit-1 binding sites. 183

6.8 DNA binding domain structures. 190

LIST OF TABLES

PAGE

CHAPTER 1: INTRODUCTION

1.1 Expression patterns and phenotypes of human GH transgenic lines. 16

1.2 Summary of placental expression in human GH/CS transgenic mice. 22

CHAPTER 2: MATERIALS AND METHODS

2.1 Oligonucleotide sequences 29

2.2 PCR primers 30

CHAPTER 3: IDENTIFICATION OF P SEQUENCE FACTORS (PSFs) *IN VITRO*

3.1 Results of the database search for putative PSE-A binding proteins. 55

3.2 PSE-ACSp.Luc does not consistently repress CS-Ap. activity. 67

CHAPTER 6: DISCUSSION

6.1 Results of the database search for additional putative PSFs upstream of PSE-B. 171

6.2 Results of the database search for additional putative PSFs downstream of PSE-A. 172

ABBREVIATIONS

103P	103 base pair P sequence fragment
263P	263 base pair P sequence fragment
A_{260}	absorbance at 260 nanometers
Ab	antibodies
ANOVA	analysis of variance
AR	androgen receptor
ATP	adenosine triphosphate
B/I	bound/input
bp	base pair
CAT	chloramphenicol acetyltransferase
CBP	CREB binding protein
C/EBP	CCAAT enhancer binding protein
ChIP	chromatin immunoprecipitation
CS	chorionic somatomammotropin
CS-Ap	chorionic somatomammotropin-A promoter
CSp	chorionic somatomammotropin promoter
Csx	cardiac specific homeobox
CTD	COOH-terminal domain
DLP	dorsolateral protein
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylenediaminetetraaceticacid
EF-C	enhancer factor-C
EMSA	electrophoretic mobility shift assay
ENH	241 base pair CS enhancer
ER	estrogen receptor
FBS	fetal bovine serum
FGF	fibroblast growth factor
FP	footprint

GH	growth hormone
GHp	growth hormone promoter
GR	glucocorticoid receptor
GTF	general transcription factor
HA	hemagglutinin
h.aff.	high affinity
HAT	histone acetyltransferase
HDAC	histone deacetylase
HFH	hepatocyte nuclear factor homolog
HNF	hepatocyte nuclear factor
HS	hypersensitive
IP	immunoprecipitation
kb	kilobase
kDA	kiloDalton
LCR	locus control region
Luc	luciferase
MHC	major histocompatibility complex
mRNA	messenger ribonucleic acid
N-CoR	nuclear corepressor
NF-1	nuclear factor-1
NF-IL-6	nuclear factor-interleukin-6
NIB	nuclei isolation buffer
nm	nanometer
NR	not reported
NRS	normal rabbit serum
NT	non-transfected
PAGE	polyacrylamide gel electrophoresis
PBS-CMF	calcium- and magnesium-free phosphate buffered saline
PCR	polymerase chain reaction
PIC	pre-initiation complex
PMSF	phenylmethylsulfonyl flouride

PR	progesterone receptor
Prl	prolactin
PSE	P sequence element
PSF	P sequence factor
RFX	regulatory factor X
RNA	ribonucleic acid
RSVp	rous sarcoma virus promoter
RT-PCR	reverse transcriptase polymerase chain reaction
SDS	sodium dodecyl sulfate
SEM	standard error of the mean
SRC-1	steroid receptor coactivator-1
SSC	sodium chloride and sodium citrate buffer
TAF	TATA-binding protein associated factor
TBP	TATA-binding protein
TEF-1	transcriptional enhancer factor-1
TGF-β	transforming growth factor-β
Tkp	thymidine kinase promoter
TNIB	tissue nuclei isolation buffer
$TNIB^+$	tissue nuclei isolation buffer with additives
tRNA	transfer ribonucleic acid
TSH	thyroid stimulating hormone

CHAPTER 1

INTRODUCTION

1.1 Regulation of differential gene expression patterns.

The somatic cell genome contains a complete manual of information for the construction and function of an organism. Expression of particular genes from this genome will vary both temporally, that is at a certain stage of cell development or in response to environmental signals, as well as spatially where expression occurs in one cell type but not another. The result of this variability is differential gene expression patterns. Without this phenomenon, the ability to produce different cell types from the same genome, and thus permit the development and function of multi-cellular organisms, would not be possible. A fundamental question in biology therefore centres around the mechanisms that contribute to differential gene expression patterns.

Regulation of gene expression can be arbitrarily divided as occurring in two major stages. The first stage involves overcoming the structural and spatial constraints that are imposed on the genome within the nucleus. In this thesis, these processes are referred to as derepression, based on the perspective that the natural state of the genome is restrained or inactive. The second stage of regulation is the activation of transcription that occurs once genes are accessible.

1.1.1 Overcoming the inactive state of the genome: derepression.

1.1.1a Chromatin Structure

The accessibility of the nuclear genome is restricted through its compaction as chromatin. Chromatin is a complex of DNA and associated proteins; a necessary structure that enables the DNA to fit into the available space of the nucleus. The basic unit of chromatin, the nucleosome, is repetitive (Fig. 1.1).

The nucleosome consists of an octamer of core histone proteins (histone H3, histone H4, histone H2A, and histone H2B), around which is wrapped 146 base pairs (bp) of DNA (1-7). The core histone proteins have globular COOH-terminal domains containing protein structures known as histone-fold motifs (1). The histone-fold motifs are responsible for both the intra-nucleosomal interactions and contacts with the phosphodiester backbone of the DNA (5). When the DNA 'wraps' around the core histone structure, interactions with the histone-fold motifs 'pull' the DNA into the nucleosome, inducing significant structural changes to the DNA helix (8). The core histone proteins also consist of NH_2-terminal tails that project outward from the nucleosomal core (5). The exposed histone tails do not appear to be required for the integrity of the nucleosomal structure (1, 5), however, they are subject to several modifications. The potential for these modifications to be involved in the regulation of gene accessibility will be discussed in further detail later in this section.

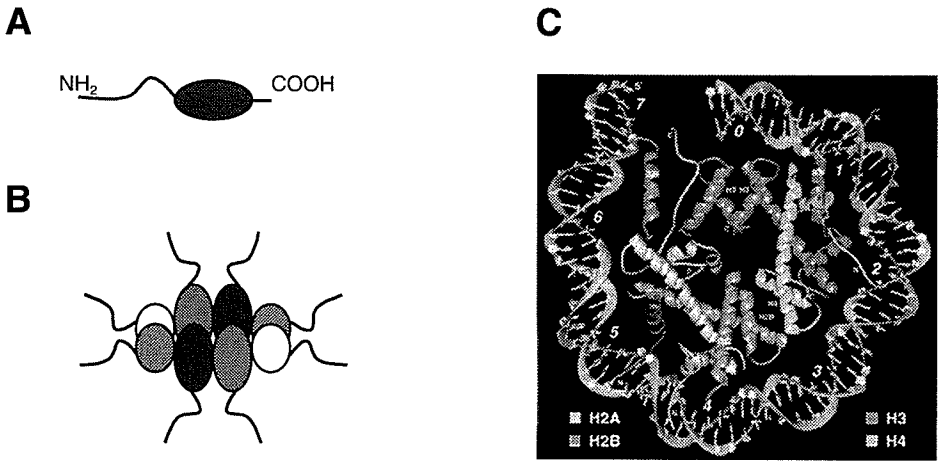


Figure 1.1: The nucleosome is the basic repetitive unit of chromatin.
(A) The core histone proteins consist of globular COOH-terminal domains and NH_2-terminal tails. **(B)** The nucleosome contains an octamer of core histones proteins. The octamer is composed of a histone H3 (blue) – histone H4 (green) tetramer, that is flanked by histone H2A (yellow) – histone H2B (red) dimers. The NH_2-terminal tails project outwards from the globular core. **(C)** In the nucleosome structure, 146 base pairs of DNA is wrapped around the histone octamer core. Figure 1.1C is taken from (5).

The repetitive nucleosome units are visualized as "beads on a string" at lower than physiological salt concentrations (9-11). This structure has been termed the 10 nm fiber. Within the 10 nm fiber, nucleosomes are separated by 'linker' DNA, which can vary in its length between cell types as well as within a single nucleus (12). The capacity for the length of linker DNA to be dynamic rather than static is an important characteristic in the accessibility of DNA sequences.

Compaction of DNA continues beyond the 10 nm fiber (Fig. 1.2). At closer to physiological salt concentrations, a structure referred to as the 30 nm fiber is visualized (13). There have been several proposed models for the structure of the 30 nm fiber. Perhaps the best known of these models is the 'classic' symmetrical solenoid structure, often used in textbook schematics (14, 15). A drawback to many of these model structures, however, is the use of an invariable length for the linker DNA. The relatively more recent Woodcock and Horowitz model accounts for variable linker DNA, and therefore, more accurately resembles the structures that are visualized in the nucleus (10, 11, 16, 17). Stabilization of the 30 nm fiber structure occurs through the presence of histone H1, or 'linker histone' (13, 15), and the NH_2-terminal tails of the core histone proteins (5). Beyond the 30 nm fiber, chromatin is further compacted through attachment to the nuclear matrix and the formation of loop structures (18-22), which can vary in length from 5,000-100,000 bp per loop (23). Variation in the organization and length of loops represents another possible level for regulating gene expression patterns (23). The chromosomal fibers that are readily visualized during mitosis, are the most compacted chromatin state.

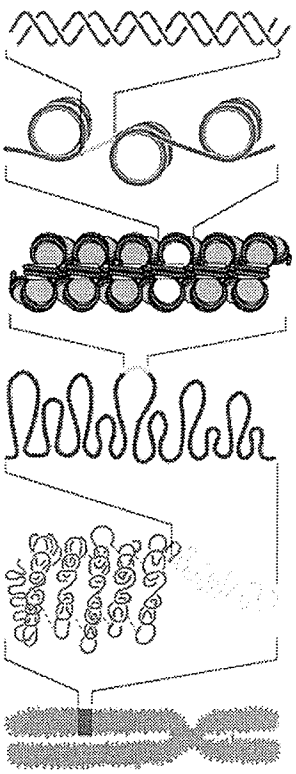


Figure 1.2: The compaction of chromatin in the nucleus.
The compaction of DNA into chromatin permits it to fit within the available space of the nucleus. This schematic demonstrates the different levels of chromatin compaction from 'naked' DNA to the condensed mitotic chromosome. The 30 nm structure shown is the solenoid model. This schematic is modified from (14).

Compaction of DNA into chromatin structures permits it to fit within the available space of the nucleus. A consequence of this compaction, however, is that the accessibility of the DNA for processes such as replication and transcription is restricted. Evidence has shown that the mere presence of nucleosomes on DNA sequences can restrict the association of many transcription factors (24) as well as TATA-binding protein from the RNA polymerase initiation complex (25, 26). The structure of chromatin must therefore be flexible enough to allow for an increase in accessibility when the DNA template is required.

1.1.1b Increasing DNA accessibility

The structure of chromatin can be broadly divided into two classes; heterochromatin, which is densely packaged and transcriptionally inert, and euchromatin, which is

relatively less condensed and can contain transcriptionally active genes. Transitions between these two states are evident through variations in the sensitivity of the DNA to nuclease digestion. The nuclease DNaseI digests both linker and nucleosomal DNA, and the efficiency of digestion is restricted by compacted chromatin structures. Through the use of DNaseI experiments it has been observed that the chromatin structure of transcriptionally active regions of the genome is less compact than non-transcriptionally active regions, reflected by an increased sensitivity to DNaseI digestion (27, 28). A second type of nuclease also demonstrates that variability in chromatin structure exists between regions of the DNA that are transcriptionally active and those that are not. Micrococcal nuclease preferentially digests linker DNA over nucleosomal DNA. This preference results in a characteristic 'laddering' pattern of mono- and poly-nucleosomal repeats when micrococcal nuclease digested chromatin is run in agarose gels (6). Southern (DNA) blotting of chromatin, digested with micrococcal nuclease, results in a 'smear' when an active gene is used as a probe, as opposed to the characteristic repetitive nucleosomal pattern (29). From these pieces of information, it is interpreted that not only does the level of chromatin compaction undergo transitions (as evidenced by the DNaseI experiments), but that the structure of nucleosomes themselves may be modified in transcribed regions (as evidenced by the micrococcal nuclease experiments). Nucleosome modifications are generally believed to be at the root of chromatin structure transitions. These modifications are classified as those that involve ATP-dependant remodeling machines, and those that employ post-translational modifications of histone proteins.

1.1.1c Modifications of nucleosomes.

ATP-dependant chromatin modifying complexes modify the location and/or structure of nucleosomes (reviewed in (30). As their name indicates, these complexes contain multiple protein subunits and rely on the energy of ATP hydrolysis for their activity. The purpose of these complexes is to accelerate transitions in nucleosomal structure. In terms of gene activation, this function has the capacity to destabilize chromatin structures that are intrinsically inhibitory to transcription, thus facilitating an increased DNA sequence accessibility for transcription factors and RNA polymerase. There are several known ATP-dependant chromatin modifying complexes, including the SWI/SNF and ISWI

families (30). Several mechanisms that account for the ability of the ATP-dependant complexes to modify chromatin structure have been documented. These involve actual displacement of nucleosomes (31), movement (sliding) of nucleosomes (32-36), and modifications to nucleosomal structure (37, 38).

Chromatin accessibility can also be regulated through reversible post-translational modifications of histones. Although the NH_2-terminal tails of the core histone proteins are not required for the integrity of nucleosomal structure (5), they are subject to several types of post-translational modifications that have been linked to gene expression. These modifications include acetylation (39-41), methylation (42, 43), and phosphorylation (44, 45). Through these modifications, it is believed that the accessibility of the DNA sequence is altered. As a specific example, core histone acetylation disrupts the formation of condensed chromatin structures (46, 47), and several transcriptional co-activators have been demonstrated to possess histone acetyltransferase (HAT) activity, such as CREB binding protein (CBP)/p300 (48) and steroid receptor coactivator-1 (SRC-1) (49). Post-translational modifications of core histone NH_2-terminal tails has been proposed to represent a 'code' for expression that is 'read' by the transcription machinery in the cell nucleus. This theory is known as the histone code hypothesis (50).

1.1.2 The activation of gene expression.

Transcription is an activated process and does not occur simply because the chromatin structure of a gene is in an accessible state. As previously mentioned, not all genes in euchromatic regions are active. Following the processes involved in derepression, the initiation of transcription requires recruitment of the pre-initiation complex (PIC), which consists of RNA polymerase II and the general transcription factors (GTFs) (reviewed in (51, 52). This is a large complex composed of multiple subunits, as the GTFs themselves contain multiple proteins. For example, TFIID consists of TATA-binding protein (TBP) and TBP-associated factors (TAFs). The rate at which the PIC is recruited and re-recruited to the transcription initiation site are some of the primary factors for

determining the level at which a gene will be expressed. Recruitment of the PIC involves *cis*-acting DNA regulatory elements and *trans*-acting proteins. Although for the purpose of this chapter, regulatory regions and associated factors are discussed at the level of gene activation, it should be noted that they are also involved in regulating gene accessibility as described in the preceding section.

1.1.2a DNA regulatory elements

The vast majority of the genome does not contain sequences that code for proteins (53). DNA regulatory elements are sequences that are involved in determining whether a gene is or is not transcribed. There are several categories of DNA regulatory elements, among them are promoters, enhancers, repressors, and locus control regions. A precise definition of each of these categories can be quite arbitrary and depend on the context in which they are being defined. They will, however, all contain information essential for the formation of protein complexes. These complexes will serve to promote or limit transcription, depending on the function of the regulatory element. For the purpose of this introduction, repressor elements and locus control regions will be discussed separately in later sections.

The role of a promoter element is to direct accurate initiation of transcription for a linked gene (14). This function requires recruitment of the RNA polymerase PIC, and therefore, promoters can be distinguished from other categories of regulatory elements in that they are found in the immediate 5′-flanking DNA of a gene. The sequences that are considered to constitute the promoter may be arbitrarily defined either by the absence or presence of gene activity, or alternatively, by a certain level of gene activity.

Enhancer elements function to increase the rate of transcription or the probability that transcription initiation will occur (54-58) and reviewed in (59, 60). To achieve this, enhancers contain multiple sites for *trans*-acting factors. Unlike promoter elements, enhancers are not limited to the immediate 5′-flanking sequences, and often can be found at significant distances from the transcription initiation site. They have classically been defined as orientation and distance independent (59, 60), but these features do not hold for every enhancer element (61-63).

While the sensitivity to nuclease of transcriptionally active genes can be in the range of 2-3 times greater compared to non-active genes, some regions of the genome are referred to as hypersensitive. Hypersensitive (HS) regions are so sensitive to DNaseI nuclease digestion that they appear to be 'nucleosome free' (64, 65). These regions of hypersensitivity tend not to map to genes themselves, but rather indicate those regions of the DNA that are involved in regulating the transcriptional activity of a gene, that is, functional elements such as enhancers, promoters, and locus control regions (LCRs) (64, 65).

1.1.2b Locus control regions are enhancer elements.

LCRs are a class of regulatory sequences that are often defined by an ability to confer a distinct chromatin environment in transgenic animals or stable cell lines. Through this ability, the inclusion of LCR elements in a construct results in transgene expression that is not influenced by the site of integration, so that from line to line, the expression level can be correlated to the number of transgenes that have been inserted. This capacity has previously been interpreted to be due to a dominant role of the LCR in 'opening' of a chromatin locus. That is, the LCR was believed to be responsible for chromatin modifications that regulated the accessibility of a locus for transcription. Following this derepression function, the LCR was believed to co-operate with more proximal regulatory elements to result in transcriptional activation of a gene or genes within this open and accessible domain.

This interpretation of LCR function was largely based on studies of the β-globin multigene cluster, which is specifically expressed in erythroid cells. Deletion of distal upstream HS sites, contained within the LCR, resulted in loss of both nuclease sensitivity and expression of the β-globin genes (66). It was also observed that efficient, site-of integration independent expression in transgenic mouse lines required that these regions be included in the transgene construct (67). Based on this and additional studies, it is not generally disputed that the β-globin LCR has the capacity to contribute to the formation of an independent chromatin domain at ectopic loci. Additional information, however,

has been obtained from analysis of the endogenous β-globin locus in both human and mouse genomes. These more recent findings, suggest that the role of the LCR as 'dominant' in the derepression process may be overstated. From the evidence outlined below, it is now argued that 'opening' the chromatin domain is not a function exclusive to the LCR, and that more proximal regulatory elements have the capacity to derepress the locus in the absence of the LCR.

The endogenous locus models rely on the use of targeted deletions of the β-globin LCR. In these experiments, deletion of the LCR significantly affects the level of transcription, without major changes in nuclease sensitivity, histone acetylation, transcription factor or PIC recruitment to the promoter (55, 68-72). From these observations, the function of the LCR is seen to be as a potent enhancer of gene transcription, subsequent to changes in the accessibility of the locus. This enhancer capacity may involve mechanisms that assist in the transition of RNA polymerase II from an initiation to an elongation complex (55). It should not be interpreted, however, that enhancer activity in the β-globin locus is restricted to the LCR, as more proximal elements have also demonstrated a contribution to gene expression levels (73). The hypothesis that the mechanism of LCR enhancer activity involves processes occurring at the promoter is further supported through separate observations that the LCR and promoter elements function in close spatial proximity, despite their linear distance in the DNA sequence (74, 75), and the detection of RNA polymerase association with HS sites of the β-globin LCR (76).

1.1.2c *Trans*-acting factors

The *cis*-acting DNA sequences do not activate transcription themselves, but require the association of *trans*-acting factors. *Trans*-acting factors associate with specific DNA sequences through non-covalent interactions between their DNA binding domains and acidic DNA sequences (14). Factors with similar DNA binding domains are grouped together as a family or class of transcription factors. These families include homeodomain proteins, zinc finger proteins, and leucine zippers (14).

In addition to DNA binding domains, *trans*-acting factors also contain activation or repression domains that are responsible for their activity. Three major classes of activation domain have been well characterized as acidic, proline-rich, and glutamine rich (14). Some *trans*-activating factors do not contain one of these types of activation domains. They may contain a less characterized motif that is responsible for activation, or alternatively, they may function through the recruitment of other factors. In addition, there are classes of *trans*-acting factors that do not contact the DNA sequences directly, and thus are termed co-activators or co-repressors.

1.1.3 Mechanisms of transcriptional repression.

Repressing gene expression involves opposing the mechanisms that lead to gene activation. That is, if turning a gene 'on' involves both depression, to increase gene accessibility, and recruitment of the transcriptional machinery for transactivation, then turning a gene 'off' involves limiting the accessibility of a gene and/or preventing the processes that lead to transactivation. Therefore, whereas many steps are coordinated for a gene to be expressed, repression can involve interference with just one or more of these stages.

Repressor elements are DNA sequences that contain binding sites for *trans*-acting factors. Repressor mechanisms involve binding of a repressor protein that inhibits either the association or function of an activating factor (reviewed in (77-81). This may involve competition for overlapping binding sites, destabilizing an activator complex, or blocking the recruitment of other activators and co-activators. Alternatively, the binding of the repressor protein may aid in the recruitment of additional factors, such as other repressor proteins and/or co-repressors, for the formation of a repressor complex. A repressor complex has the potential to be actively involved in processes such as chromatin modifications, or interference with RNA polymerase activity (77-81).

1.2 The human growth hormone (GH) / chorionic somatomammotropin (CS) family as a model system for examining differential gene expression.

The development and maintenance of differential expression patterns is an extremely complex process. While individual components, such as promoter and enhancer elements, have been studied for a number of genes, it is of considerable interest to have model systems in which the separate components can be studied both on an individual basis as well as in relation to one another. The human growth hormone (GH) / chorionic somatomammotropin (CS) locus is one such model system. The human GH/CS locus on chromosome 17 contains a family of five highly related genes. From 5′ to 3′, the genes within this locus are GH-N, CS-L, CS-A, GH-V, and CS-B (Fig. 1.3).

Figure 1.3: The human GH/CS locus on chromosome 17.

The genes of the GH/CS locus are believed to have evolved through processes of gene duplication (82-84), resulting in similar gene structures and flanking regulatory sequences (83). The primary cells that express the GH and CS genes are the somatotrophs of the anterior pituitary (GH-N) and placental syncytiotrophoblasts (GH-V and the CS genes) (83). For the purpose of this thesis, two levels of transcription will be arbitrarily described. High-level expression will be defined as activated expression. Lower expression levels, that will usually not involve cell-specific or environment-specific induction, will be defined as basal expression. The expression of the GH-N and CS-A genes accounts for greater than 3% of the total mRNA in the somatotrophs and syncytiotrophoblasts respectively (83), and thus, expression in these cells is defined as activated. Transcripts for the human GH/CS family have also been observed in several ectopic tissues. These include the brain, colon, kidney, lung, mammary glands, muscle, ovaries, prostate cancer cell lines, skin, and testis, (85-91) as well as several lymphoid tissues and cell types (92-100). Due to the lower levels of transcript in these tissues, relative to pituitary or placental expression, this is defined as basal gene expression.

There are several features of the GH/CS locus that make it an interesting model system for the study of differential gene expression. The activated pituitary/placenta and basal ectopic expression suggests that more than one regulatory mechanism may be involved for expression of genes from this locus. As will be discussed in greater detail within the following sections, separate mechanisms to achieve activated expression are certainly present for regulating pituitary and placental expression. The mutually exclusive expression of GH-N in the pituitary somatotrophs and the remainder of the family in placental syncytiotrophoblasts, occurs despite the extensive sequence homologies in their flanking regulatory sequences. There are several lines of experimental evidence which suggest that the mechanisms employed in each of these two tissues for activated expression of the GH/CS locus are distinct.

The GH/CS model system is also distinct from another well-studied gene family, the β-globin model system. Like the GH/CS family, the β-globin genes are part of a single locus. They are specifically expressed in erythroid cells, with different members of the

family being expressed at different stages of erythroid development (reviewed in (101, 102). The β-globin family therefore provides an excellent model system for the study of expression patterns that vary temporally. What makes the GH/CS system distinct as a model of differential gene expression patterns is that the expression does not only vary temporally, but also spatially, as expression is separately activated in both somatotrophs as well as syncytiotrophoblasts. The differential spatial expression, especially given the extensive homology within the locus, is a particularly intriguing feature of this model.

1.2.1 Transcriptional regulation of the GH/CS locus in the pituitary.

1.2.1a Activity of the GH/CS 5′-flanking sequences *in vitro*.

In transfected pituitary cell cultures, activated expression from GH-N promoter sequences is attributed to two regions of protein binding within the first 150 base pairs of 5′-flanking DNA (103, 104). These two sites are recognized by the same *trans*-activating factor (104), which was purified (105) and subsequently cloned by two groups as the pituitary-specific POU-homeodomain protein Pit-1 (106, 107). Pit-1 plays an integral role in development of the anterior pituitary, as it is required for the presence of three anterior pituitary cell types, somatotrophs, lactotrophs, and thyrotrophs (108, 109). In addition to contributing to activated expression of GH-N, Pit-1 is involved in auto-regulation of its own gene (110), as well as regulation of the prolactin gene in lactotrophs (111, 112) and thyroid stimulating hormone (TSH) in thyrotrophs (113).

Despite mutually exclusive expression of GH-N and the placental GH/CS genes *in vivo*, the CS-A promoter is as efficient as the GH-N promoter at driving reporter gene expression *in vitro* (114). Likewise, stable integration of the CS-A gene into a rat pituitary cell line, with 496 bp of 5′-flanking and 630 bp of 3′-flanking sequences, results in activity that is comparable to a stably integrated GH-N gene (496 bp of 5′-flanking and 628 bp of 3′-flanking sequences) (115, 116). Examination of GH-N and CS-A promoter sequences helps to explain this apparent *in vitro*/*in vivo* paradox. The gene duplications that gave rise to the human GH locus, resulted in extensive similarity in the sequences flanking the GH/CS genes (82, 83). Both the proximal and distal Pit-1 sites of the GH-N

promoter are highly conserved in the placental gene promoters. Disruption of the proximal Pit-1 site decreases activity of stably integrated GH-N and CS-A promoters in pituitary cells by greater than 90% (117). The proximal site of the CS-A promoter is 100% identical to GH-N, and association of Pit-1 with the CS promoters has been demonstrated *in vitro* (118, 119). Therefore, like the GH-N promoter, Pit-1 binding to the CS-A promoter drives activity in transfected pituitary cells. These observations suggest that the mechanism for selectively activating GH-N and not CS-A in pituitary somatotrophs *in vivo* may lie outside of the highly homologous promoter regions.

1.2.1b Expression of the GH/CS locus in the pituitary *in vivo*.

Appropriate pituitary expression of GH-N *in vivo* was in fact found to rely on sequences not contained within the proximal promoter region. Several lines of transgenic mice were generated with various segments of GH-N 5′-flanking DNA, but were unsuccessful in producing specific and efficient expression of the human GH-N gene in the pituitary (Table 1.1). In 1995, Cooke and colleagues reported the detection of DNaseI hypersensitive (HS) sites in the distal 5′-flanking sequences of the GH/CS locus in human GH-secreting pituitary adenoma chromatin (120). Two of these sites, HS I and HS II, located at -14.6 kb and -15.4 kb from the GH transcription start site respectively, are pituitary-specific (120). The remaining two HS sites, HS III at -27.5 kb and HS V at -32.5 kb, were observed in both pituitary adenoma and human term placenta chromatin (120). These four HS sites were not detected in K562 human erythroid cell nuclei (120), and thus their formation appeared to be restricted to tissues in which genes from the GH/CS locus were expressed. On the basis of these observations, five transgenic mouse lines were created using a genomic fragment of human DNA that contained the GH-N gene and 40 kb of 5′-flanking DNA (-40 GH-N) (120). This construct included the complete set of pituitary HS sites. Human GH-N was consistently and efficiently expressed in the pituitaries of all five transgenic lines. The expression was specific to pituitary somatotrophs, as there was no evidence of ectopic GH-N expression through northern blotting, RT-PCR analysis, or a comparison of serum GH to pituitary GH mRNA ratios. Additionally, the level of GH-N expression correlated to transgene copy number (120, 121). From these observations, it was concluded that the GH/CS distal HS sites permitted

tightly regulated transgene expression that was independent of the site of integration. That is, the -40 GH-N construct appeared to generate its own chromatin environment that was not influenced by surrounding sequences. The distal regulatory region that encompasses the HS sites has thus been termed the locus control region (LCR) (120).

Table 1.1
Expression patterns and phenotypes of human GH transgenic lines.

Construct	Pituitary expression of GH-N	Ectopic expression	Copy number dependance	Gigantism	Reference
-0.5 kb GH-N	3/5 lines	NR	No	No	(120)
-5.0 kb GH-N	0/1 line	NR	No	No	(120)
-7.5 kb GH-N	3/5 lines	Yes	No	No	(120)
-22.5 kb GH-N	4/4 lines	Spleen	Yes	Yes	(120)
-40 kb GH-N	5/5 lines	No	Yes	No	(120)
-45.8 kb GH-N (P1 clone)	5/5 lines	Brain, ovary, testis, spleen	Yes	No	(122)
1.6 kb HS I/II (F) 0.5 kb GH-N	3/3 lines	Brain	No	Yes	(120)
1.6 kb HS I/II (R) 0.5 kb GH-N	3/3 lines	Brain	No	Yes	(120)
HS III-HS V 0.5 kb GH-N	5/5 lines	Kidney	Yes	No	(120)
-45.8 kb GH-N ΔHS I	6/6 lines	Brain, kidney	Yes	No	(123)
-45.8 kb GH-N ΔHS III	3/3 lines	NR	Yes	NR	(123)

NR: not reported

The most highly regulated expression of GH-N in transgenic mice has only been reported when the entire LCR is present (120), however, several lines have been made in an attempt to analyze the contributions of the individual HS site regions to GH/CS LCR function (See Table 1.1). Due to their close proximity and pituitary-specificity, HS I and HS II are often grouped together as the HS I/II region. In 17 separate transgenic mouse lines that contained the HS I/II region in its native position relative to the locus, GH-N was consistently expressed at high levels in the pituitary. The absence or specific deletion of the HS I/II region resulted in drastically reduced levels of pituitary transgene expression. Isolated HS I/II fragments were also shown to be potent enhancers of pituitary promoter activity in both transgenic mice and cell culture (120, 121, 124). For example, the expression of GH-N in the 1.6 kb HS I/II(F) 0.5 kb GH-N transgenic lines was approximately 1000-fold higher than in the 0.5 kb GH-N transgenic lines (120). The enhancer activity of HS I/II localizes to three Pit-1 binding sites (124-126). Therefore, not only does Pit-1 drive the activity of the GH-N promoter, but it is also involved in the LCR through its association with HS I/II.

The remainder of the LCR is also required for tightly regulated expression of GH-N *in vivo*. Despite the detection of GH-N expression in the pituitaries of five of five transgenic lines carrying the HS III-HS V 0.5 kb GH-N construct, this expression was approximately 100-fold lower than the -40 kb GH-N construct (120). Thus, the HS III-V region appears to confer a function which is distinct from the pituitary-specific enhancer activity of HS I/II. In constructs that include the HS I/II enhancer, but lack the HS III-V region, gigantism is observed (See Table 1.1). This region may therefore be involved in the conferring appropriate physiological control on GH-N expression.

1.2.1c The lack of placental gene expression in the pituitary.

While studies of the GH/CS LCR provide some insight into pituitary activation of GH-N, they have contributed very little to understanding the lack of placental gene expression in this tissue. The expression of GH-N and the placental genes is mutually exclusive in pituitary somatotrophs, to the extent that not even basal levels of placental gene

expression are detected (83). In assessments of GH/CS locus accessibility in the pituitary, obvious differences between GH-N and the placental genes were not observed, through either analysis of nuclease sensitivity (127) or methylation patterns (128). Additionally, the HS I/II region is not a specific activator of the GH-N gene in the pituitary. It is also able to enhance the activity of the minimal thymidine kinase promoter in both transgenic mice and cell culture (124). The question therefore arises, as to why the highly homologous placental promoters, which like the GH-N promoter can also be driven by Pit-1, are not activated in the pituitary by the same mechanism that activates GH-N.

Analysis of human pituitary chromatin revealed hyperacetylation of core histones in the LCR and the GH-N 5′-flanking DNA (129). The hyperacetylation pattern does not appear to extend to the placental genes of the locus (129), and thus correlates with the expression pattern of these genes. Both hyperacetylation and GH-N expression are disrupted by deletion of two Pit-1 sites from HS I in transgenic mouse lines (123). From this and other observations, it was concluded that Pit-1 binding at these sites was the key event in activating the GH locus and expression of GH-N in the pituitary (123).

The pituitary enhancer activity of HS I/II, and the activity of the GH-N and CS-A promoters *in vitro*, is mediated through the transcription factor Pit-1. Interference with Pit-1 binding to the placental promoters *in vivo* is therefore a potential explanation for their lack of activity in pituitary somatotrophs. It is unknown at this point if differences exist between Pit-1 binding at these locations *in vivo*, as until recently distinctions between these homologous regions have not been reported (130). However, even if Pit-1 binding were to be demonstrated with both the GH-N and the placental promoters *in vivo*, an alternative possibility for the blockage of placental promoter activity would be that the ability of associated Pit-1 to activate transcription is compromised. Pit-1 was previously shown to alternately recruit the histone acetyltransferase/transcriptional activator CBP and the histone deacetylase/repressor complex N-CoR (131, 132). Based on these observations, the hypothesis in the case of Pit-1 association with the placental promoters, would be that binding of Pit-1 to the placental promoters nucleates a repressor complex, as opposed to the activation complex that is formed at the GH-N promoter. The ability of

Pit-1 to recruit opposing complexes was demonstrated as being based on differences in the spacing of nucleotides within Pit-1 binding sites (131). As previously noted, the proximal and distal sites in the GH-N and placental promoters are highly homologous, and the spacing of the nucleotides in the Pit-1 recognition sites does not differ between GH-N and the placental promoters. This does not rule out the possibility, however, that instead of differences in the sequence information contained within the Pit-1 sites themselves, surrounding sequences and/or factors could alter the manner in which Pit-1 interacts with the placental GH/CS promoters.

1.2.1d The P sequences

Despite the extensive homology in the GH/CS locus, there are surrounding sequences that are found upstream of the placental GH/CS promoters, which are absent from the GH-N upstream region. Repeat elements, termed P sequences, are located approximately 2 kb upstream of each of the placental GH/CS genes (83). The capacity for these sequences to block pituitary expression of GH/CS genes has been demonstrated through insertion of a 2.4 kb fragment of P sequence DNA upstream of the both the GH-N and CS-A promoters in transiently transfected pituitary cells (133). Repressor activity was further localized to a 263 bp fragment (263P) (133). Within 263P, two regions of protein binding were detected and the corresponding DNA elements are P sequence element (PSE) –A and PSE-B (133). The identity of the factors associating with PSE-A and PSE-B was undetermined, although the ability to see binding to these regions in several cell types (133) indicated that the factors were not specific to the pituitary. Based on their location within the locus, and function *in vitro*, P sequences and their associating factors were hypothesized to participate in a mechanism that selectively repressed the placental GH/CS genes in the pituitary. From what is known about GH-N activation, one would expect that this mechanism may involve interference with Pit-1 (as outlined above). A direct link between Pit-1 and the P sequences factors (PSFs) was not made, as the identity of the PSFs was unknown. A putative interaction was inferred, however, as Pit-1 oligonucleotides competed for PSF binding in structural *in vitro* assays (133).

1.2.2 Transcriptional regulation of the GH/CS locus in the placenta.

1.2.2a Regulation of the placental promoters *in vitro*.

Whereas the pituitary-restricted transcription factor Pit-1 drives expression of both the GH-N and CS-A promoters in pituitary cells *in vitro* (103-107), placental promoter activity is attributed to factors with broader patterns of expression. Five regions of protein binding within the first 250 bps of the CS-A promoter have been observed by nuclease protection experiments with choriocarcinoma (BeWo) and human placenta nuclear proteins (134). These regions contain putative binding sites for Sp1 (FP1), a glucocorticoid response element (FP2), upstream-regulatory factor, MyoD, and an estrogen response element (FP3), NF-1 and AP-2 (FP4), and AP-2 (FP5) (134). The association of Sp1 (to homologous CS-B sequences), NF-1, and AP-2 was confirmed through *in vitro* binding assays, however, only Sp1 and AP-2 have been demonstrated to be positive contributors to *in vitro* promoter activity (134, 135). Thus, a placenta-restricted regulatory factor, analogous to the role of Pit-1 in the pituitary, has not been demonstrated for placental promoter activity.

1.2.2b The CS genes have placental enhancers in their 3′-flanking DNA sequences.

The expression of CS in the placenta is approximated to represent 3.5% of the total mRNA of syncytiotrophoblasts, similar to the 3% of total mRNA that GH-N represents in somatotrophs (83). Promoter activity alone may not explain this high level of activated expression, as in comparison, the homologous GH-V gene accounts for less than 0.001% of the total syncytiotrophoblast mRNA (83).

Roughly 2 kb downstream of the CS-B gene, a 1022 bp enhancer element was identified (136). Enhancer activity can be localized within a 241 bp fragment of the 1022 bp element, which has the capacity to increase CS-A promoter activity in transiently transfected placental cell lines by as much as 300-fold (137-139). Like many of the GH/CS locus regulatory regions, the CS enhancer is a repeat sequence that is also found downstream of the CS-A and CS-L genes. The CS-A and CS-L enhancers, however, are not as potent as the CS-B enhancer fragment in transiently transfected choriocarcinoma

cells lines (136, 137, 140), although activity of the CS-A enhancer region does increase in placental primary cell cultures undergoing differentiation (141). From these observations, it has been suggested that the CS enhancers may contribute to activated expression of the CS genes in the human placenta *in vivo*. Detailed analyses of CS-B enhancer function implicate multiple regulatory regions within the 241 bp element (137, 138, 140, 142, 143), as well as the involvement of several transcription factors, including a TEF-1-like factor (137, 142, 144, 145) reported to be TEF-5 (146, 147), and nuclear factor-interleukin-6 (NF-IL-6) (148).

1.2.2c Expression of the GH/CS locus in the placenta *in vivo*.

Based on data generated through *in vitro* models, activated expression of the GH/CS genes in the placenta may involve the CS downstream enhancer elements. This enhancer activity is not promoter-specific (137, 139), however, the capacity to co-operate with the CS promoter complex for increasing expression levels has been demonstrated (149). Thus, based on these observations, both the CS downstream enhancers and CS/GH promoter regions are candidates for involvement in activated expression of the GH/CS locus *in vivo*.

As was previously mentioned, the distal HS III and HS V regions were identified in human term placenta as well as pituitary adenoma chromatin (120). In this study, the presence of a placental-specific HS site, called HS IV, was also observed 30.5 kb 5′ to the GH-N transcription start site (120). Using chromatin immunoprecipitation (ChIP) assays, hyperacetylation of histones in the HS III and HS V regions has also been observed (150). This study included analysis of proximal locus regulatory elements, and demonstrated that the upstream P sequences were also hyperacetylated. These chromatin modifications can indicate sequences that are be involved in transcriptional regulation, and so these regions are candidates for activated expression of the GH/CS genes in the placenta. As in the pituitary, transgenic mouse models have provided some insight into how the expression of these genes is regulated (Table 1.2).

Table 1.2

Summary of placental expression in human GH/CS transgenic mice.

Construct	Placenta expression	Pituitary expression	Copy-number dependence	Reference
15 kb CS-A (5.4kb 5′ - 7.2 kb 3′)	Variable levels 4/6 lines (N)		No	(120)
HS III-HS V 0.5 kb GH-N	Not detected (RT)	5/5 lines (RT)	Yes	(120)
-40 kb GH-N	Not detected (RT)	5/5 lines (RT and N)	Yes	(120)
-45.8 kb GH-N (P1 clone)	5/5 lines (RT and N)	5/5 lines	4/5 lines	(122)
HS I/II 263P 0.5 kb GH-N 263P	Variable levels 2/3 lines (RT)	3/3 lines	No	(150)
HS I/II 0.5 kb GH-N	Not detected (RT)	3/3 lines	No	(150)

Method of detection for transgene expression in the placenta and pituitary is indicated in brackets: RT - RT-PCR, N - northern (RNA) blotting.

Perhaps the clearest result that can be taken from the studies summarized in Table 1.2, is that independently, none of the regions tested are sufficient to confer efficient and consistent expression in the placenta. The 15 kb CS-A transgene contained not only the proximal promoter sequences, but enough flanking DNA to contain the upstream P sequences and downstream CS enhancer (120). Expression of this construct in the placenta, however, was variable and detected in only four of six lines (120). Independent use of distal sequences also did not recapitulate placental expression in the transgenic mouse models. Although the data from other tissues demonstrates that the HS III-V region can confer copy-number dependant expression of a transgene (in both the HS III-HS V 0.5 kb GH-N and -40 kb GH-N lines), the level of expression in the placentas of

transgenic mice was either extremely low or non-existent, as it was not detected in placental RNA by RT-PCR. The single construct that has been reported to confer consistent and efficient expression of the locus in the placenta of transgenic mice is -45.8 kb GH-N (P1 clone), which includes the entire locus (with the exception of the CS-B gene) and the complete set of placental hypersensitive sites. Taken together, these results imply that neither proximal nor distal elements alone are sufficient for appropriate expression of the GH/CS gene family in the placenta, and instead suggest that some degree of co-operativity between these regions may be occurring.

1.2.2d The P sequences.

A second result that can be taken from the data in Table 1.2, is that P sequences have the capacity to enhance gene activity in the placenta of transgenic mouse models. Placental transgene expression was evident in two of three HS I/II 263P GH-N 263P lines, compared to a similar construct HS I/II GH-N, which lacked the 263P fragment and resulted in no detectable placental expression (150). The 15 kb CS-A transgene, which included the P sequences, also resulted in placental gene expression in four of six lines. In each of these constructs, the level of transgene expression was variable, and this would suggest that they are susceptible to site-of integration effects. This interpretation is consistent with the fact that the distal HS III-V region was not contained in these constructs, which has the capacity to produce expression levels that relate to transgene copy number (120).

It would, therefore, appear that in the placenta, P sequences act as local enhancer elements involved in activated expression. This function would be analogous to the enhancer activity of the HS I/II region in the pituitary, and is consistent with the observation of P sequence hyperacetylation in human term placenta (150). This enhancer function may be restricted to the P sequences, however, more likely is the possibility that they operate in conjunction with other local elements such as the placental promoters and downstream CS enhancer regions.

1.3 Research objectives

Regulation of the human growth hormone locus involves multiple elements and factors. The mechanisms and factors appear to vary, depending on whether expression is occurring in the pituitary or the placenta. For example, the HS I/II region is specifically involved in activated expression of GH-N in the pituitary, requiring the transcription factor Pit-1. In contrast, some elements have the potential to be involved in more than one of the two expression mechanisms. The P sequences for example have the potential to function as a dual element, with more than a single specific function, as discussed in sections 1.2.1d and 1.2.2d. In the pituitary, where the locus is accessible and there is no clear basis for differentiation between GH-N and placental promoter activity, the P sequences have been implicated as local repressor elements. In the placenta, data have also suggested that P sequences play a role in achieving activated expression of the placental genes.

The aim of this thesis is to further characterize P sequences, primarily through the identification of P sequence factors in pituitary and placental cells. The research objectives for this work are therefore as follows:

1. To identify candidate P sequence factors (PSFs) in pituitary GC cells.

2. To determine if the placental (choriocarcinoma) cell lines available are a suitable *in vitro* model for the study of placental enhancer function, and if so, to use this system to identify candidate PSFs.

3. To assess the potential for candidate PSFs identified *in vitro* to associate with P sequences in the context of pituitary and placenta chromatin

4. To examine the possibility that the P sequence repressor mechanism in pituitary cells involves the transcription factor Pit-1.

To meet these objectives, the results are presented in three chapters. Chapter 3 combines the use of functional (transient gene transfer) and structural (nuclease protection and electrophoretic mobility shift) assays to address objectives 1 and 2. The majority of the data within this chapter have been previously reported in (151) and (130). In Chapter 4, the development of a chromatin immunoprecipitation technique for use with human tissue samples is described, and used to address objective 3. A portion of the results presented in Chapter 4 has been reported in (130). Finally, objective 4 is examined in Chapter 5. Each of these chapters contains a limited discussion of the approaches and results, followed by a summary of the findings. A more detailed discussion can be found in Chapter 6.

CHAPTER 2

MATERIALS AND METHODS

2.1 Materials

For the experiments reported, materials and equipment were purchased or obtained from Amersham Biosciences, Baie d'Urfé, PQ, Canada, Baker company, Sanford, ME, USA, Beckman Instruments Canada Inc., Mississauga ON, Canada, Bio-Rad Laboratories Inc., Mississauga, ON, Canada, Calbiochem, La Jolla CA, USA, Clontech, Palo Alto, CA, USA, Corning Life Sciences Inc., Acton, MA, USA, Fisher Scientific, Nepean, ON, Canada, Health Sciences Centre, Winnipeg, Manitoba, Canada, Human Pituitary Repository in the Protein and Polypeptide Laboratory at the University of Manitoba, Winnipeg, Manitoba, Canada, Invitrogen, Burlington, ON, Canada, MJS Biolynx Inc., Bockville, ON, Canada, Promega Corp., Madison, WI, USA, QIAGEN, Mississauga, ON, Canada, Roche Diagnostics, Laval PQ, Canada, Santa Cruz Biotechnology Inc., Santa Cruz, CA, USA, Sigma-aldrich Canada Ltd., Oakville ON, Canada, Tropix Inc., Bedford, MA, U.S.A, University Core DNA Services, University of Calgary, Calgary, AB, Canada, Upstate Biotechnology, Lake Placid, NY, USA, Vibra Cell™, Sonics and Materials Inc., Danbury, CT, USA

2.1.1 Antibodies

2.1.1a c-myc

The c-myc antibody is a mouse monoclonal antibody that specifically recognizes the c-myc epitope tag. It was raised against a synthetic peptide that corresponds to residues 408-439 of the human p62-c-myc protein (Clontech).

2.1.1b Dorsolateral prostate protein (DLP)

The DLP antibody is a rabbit polyclonal antibody, specific for rodent prostate dorsolateral protein (152). It was used as an unrelated control antibody in

immunoprecipitations. This antibody was a generous gift from Dr. Janice G. Dodd (Department of Physiology, University of Manitoba, Winnipeg, Canada).

2.1.1c Hyperacetylated histone H4

A rabbit antiserum was purchased from Upstate Biotechnology that specifically recognizes the hyperacetylated forms of histone H4. Human histone H4 is sequentially modified by acetylation at lysines 5, 8, 12, and 16. This antibody is reported to detect the tri- and tetra-acetylated forms, but not the mono- or di-acetyl forms, of histone H4 (Upstate Biotechnology).

2.1.1d NF-1

Two NF-1 antibodies were used. A commercially available NF-1 antibody (NF-1 H-300) was purchased from Santa Cruz Biotechnology. It is a rabbit polyclonal antibody raised against a recombinant protein, corresponding to amino acids 1-300 at the NH_2-terminus of human NF-1. This antibody is reported to react with all NF-1 variants of human, rat, and mouse (Santa Cruz).

A rabbit polyclonal antibody to NF-1 (2899) was generously provided by Dr. N. Tanese (Department of Microbiology, NY School of Medicine, New York, USA). This antibody was raised against bacterially expressed recombinant NF-1C of human origin (153).

Unless otherwise stated, the commercial NF-1 antibody was used for immunoprecipitation and protein blotting experiments and the NF-1 2899 antibody was used for electrophoretic mobility shift or 'supershift' assays.

2.1.1e Pit-1

An affinity purified rabbit polyclonal Pit-1 (Pit-1 X-7) antibody was purchased from Santa Cruz Biotechnology. This antibody was raised against a polyhistidine fusion protein containing the full length Pit-1 protein of rat origin. It is reported to recognize Pit-1 of human, rat, and mouse (Santa Cruz).

2.1.1f RFX1

Two RFX1 antibodies were used. Both antibodies were used for immunoprecipitation, protein blotting, and supershift assays. A commercially available goat polyclonal RFX1 antibody (RFX1 D-19) was purchased from Santa Cruz biotechnology. This antibody was raised against a peptide mapping to an internal region of RFX1 of human origin and is reported to recognize RFX1 of human, rat, and mouse (Santa Cruz).

A rabbit polyclonal antibody to RFX1 was generously provided by Dr. W. Reith (Déptartment de Génétique et Microbiologie, Centre Médical Universitaire, Geneva, Switzerland). This antibody was raised against a bacterially produced recombinant RFX1 protein of human origin (154).

2.1.1g RFX2

A rabbit polyclonal RFX2 antibody was a generous gift from Dr. W. Reith.

2.1.1h RFX3

The RFX3 antibody was a generous gift from Dr. W. Reith. It is a rabbit polyclonal antibody that was raised against a bacterially expressed recombinant polyhistidine RFX3 peptide, corresponding to amino acids 183 to 707 (155).

2.1.2 Cell Lines

Cells were cultured in a humidified atmosphere at 37°C, 5% CO_2 and 95% air, in Dulbecco's modified Eagle's Medium supplemented with 8% (v/v) fetal bovine serum (FBS), 500 μM glutamine and antibiotics (50 U/mL penicillin, 50 μg/mL strptomycin) in 100 mm culture dishes. Cell stocks were maintained at -80 °C in 90% FBS, 10% dimethyl sulfoxide. Trypsin-EDTA was used to lift the cells following a single wash in calcium- and magnesium- free phosphate buffered saline (PBS-CMF).

The cell lines used were: rat pituitary adenoma (GC), human cervical carcinoma (HeLa), human placental choriocarcinoma (JAR and JEG-3), and human embryonic kidney (293).

The rat pituitary GC cells (156) are a well established model for examining the regulation of human growth hormone (114, 115). They express rat growth hormone, but not prolactin (157). The placental JAR (158) and JEG-3 (159) cell lines express human chorionic somatomammotropin, but at reduced levels compared to human term placenta syncytiotrophoblasts (160).

2.1.3 Oligonucleotides

Single-stranded oligonucleotides were obtained from University of Calgary core DNA services, Invitrogen, QIAGEN, and Santa Cruz Biotechnology. The sequences for the high affinity NF-1 site (h.aff.NF-1), EF-C/MDBP (RFX), RF, and HS I/II (F) and (R) primers, have been reported previously by others (124, 138, 161, 162)

Double-stranded oligonucleotides were generated by mixing equal masses of sense and antisense oligonucleotides, and incubating for 10 minutes in a boiling water bath. The heat source was then removed and the oligonucleotides were left in the water bath over night to anneal. The sense strand for each oligonucleotide, from 5′ to 3′, is given in Table 2.1. Primers for polymerase chain reaction (PCR) are listed in Table 2.2.

Table 2.1

Oligonucleotide sequences

Name	Sequence (5′ - 3′)
Egr-1	GGATCCAGCGGGGGCGAGCGGGGGCGA
h.aff.NF-1	CTAGCTATTTTGGCATCATGCCAATATG
NF-1	TTTTGGATTGAAGCCAATATGATAA
NF-1m	TTTTGGATTGAATAAAATATGATAA
Pit-1(GH)	CACCTGTTTCTGTGTACATTTATGCATGGGGCCACTGACG
Pit-1(Prl)	GGATCCACCTGATTATATATATATTCATGAAGGTGAGA
PSE-A3	TGTTGGTTGCCAACACCACTGCCAACCA
PSE-A3m1	CTAGGGTTGCCAACACCACTGCCAACCA

PSE-A3m2	TGTTTCGAGCCAACACCACTGCCAACCA
PSE-A3m3	TGTTGGTTTAACACACCACTGCCAACCA
PSE-A3m4	TGTTGGTTGCCAGTCTCACTGCCAACCA
PSE-A3m5	TGTTGGTTGCCAACACATACGCCAACCA
PSE-A3m6	TGTTGGTTGCCAACACCACTGCGTTACA
PSE-A4	TGTTGGTTGCCAACACCAC
PSE-A5	TGTTCAATGTTGGTTGCCAACACCACTGCCAACCACTTCT
PSE-B4	GATGGCAGGGCCCCAGCA
PSE-B4Cm	GATGGCCGGGCCCCAGCA
PSE-B4Gm	GATGGCAGGGCGCCAGCA
PSE-B4CGm	GATGGCCGGGCGCCAGCA
PSE-B4m	GATCGCAGGGTACCAGCA
RF	CTCATCAACTTGGTGTGGACGGC
RFX (EF-C/MDBP)	GGCCAGTTGCCTAGCAACTAATT
RFXm	GGCCATCGTCCTAGACTATAATT

Mutated sequences are indicated by underlining.

Table 2.2

PCR primers

Name	Sequence (5′ - 3′)
263P (F)	ATGTCTGGATCCTCCTACTGGC
263P (R)	AGCTCGGATCCCACTCTGTAGAAAC
103P (F)	TTCTGGGATCCAAGTGGGGAGATGGCAGGGC
103P (R)	AACAGGATCCCGCTTCCAGAAGTGGTTGGC
HS V (F)	CTAAACTCGAGTAGAGGATAAGTGTGAGGAC
HS V (R)	CCATCCTCGAGCGAGCTGGACCACCTTAACTT
HS IV (F)	TCCCAGGGATCCGGGAAGAAGTGGTGGAC
HS IV (R)	CAAAAGAAGCTTCCCACGTAATAAGGGAGGCAC
HS III (F)	TGCGTCAAGCTTGGGCACTGTCCCGATTCGAG
HS III (R)	AGGCTCGAGCTCGGGTGCAGGCACCTTGTTTC
HS I/II (F)	GATATCAAGCTTCCCGGGTCAGTCTCTCTCCAG
HS I/II (R)	TAAGGTGAGCTCCGAGGAACAGCCCGTTCC
GHp (F)	GCTATGCTCGGCGCCCATCGTCTGC
GHp (R)	GAAGGACCGCCCACCAAGGTCT
CSp (F)	CACAGAAACAGGTGGGGTCA
CSp (R)	GAAGGACCCCCCACCAAGAAGGAC
ENH (F)	GTCTACATTTCAGCTCATCAACTTGG
ENH (R)	GCTGTGAACACATGGGGTCTCATCTTTGCGG
Prlp (F)	CCTCCAAACCAATCTAGTCTCAGATCTCACC
Prlp (R)	GGAAGTCTCACGGTTTTCTCTTTCCC

hFGF-16 exon3 (F)	GAAGAAACTCACACGTGAATGTG
hFGF-16 exon3 (R)	TTACCTATAGTGAAAGAGGTCTC

(F) forward, (R) reverse

To amplify 263P by PCR, two different sets of primers were used. The set listed in table 2.2 as 263P(F) and 263P(R) was used to generate fragments for insertion into plasmid constructs, and therefore, these primers carry *BamHI* restriction endonuclease sites (underlined). In the PCR analysis of chromatin immunoprecipitation DNA, the 263P(F) primer (Table 2.2) was used with the reverse strand of PSE-A5 (Table 2.1). The 103P primers were also used to generate fragments for insertion into plasmid sites, and the *BamHI* restriction endonuclease sites are underlined.

2.2 Methods

2.2.1 Chromatin immunoprecipitation (ChIP)

ChIP is an assay used for the detection of protein DNA interactions *in situ*. A schematic of the protocol is presented in Figure 4.2.

2.2.1a Isolation of intact nuclei and formaldehyde cross-linking.

Post-mortem human pituitary tissue was obtained from the Human Pituitary Repository in the Protein and Polypeptide Laboratory at the University of Manitoba, and placentas from term deliveries at the Health Sciences Centre, Winnipeg, Manitoba. For each ChIP experiment, nuclei were isolated from 10 grams of placenta tissue or 3-6 pooled human pituitary samples (approximately 3.5 grams) according to a previously described protocol at 4 °C (127).

To isolate nuclei, tissue samples were scissor minced at 4 °C in approximately 40 mL (placenta) or 20 mL (pituitary) of tissue nuclei isolation buffer (TNIB: 10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$) containing freshly added sodium butyrate (9.1 mM,

n-butyric acid sodium salt, $C_4H_7O_2Na$), PMSF (1 mM), NP-40 (0.1%), and protease inhibitors (1 Roche Complete Mini™ protease inhibitor tablet per 50 mL of TNIB). TNIB containing these additives is referred to as TNIB$^+$. Samples were homogenized (Brinkmann mechanical homogenizer, set at #7) in approximately 2.5 mL aliquots for 15 passes. The homogenates were pooled in a sterile beaker and filtered through 2 layers of sterile cheesecloth. The cheesecloth was rinsed with 25 mL (placenta) or 20 mL (pituitary) of TNIB$^+$. The filtered homogenate and TNIB$^+$ rinse was combined, divided into two equal volume aliquots, and centrifuged at 2000 rpm in a swinging bucket HN-S centrifuge for 10 minutes at 4 °C. Supernatants were discarded and each pellet was resuspended in 10 mL of TNIB$^+$. Samples were homogenized a second time in approximately 2.5 mL aliquots for 10 passes, pooled, centifuged at 2000 rpm in the HN-S centrifuge for 10 minutes at 4 °C, and supernatant was discarded. For the isolation of placental nuclei, this pellet was resuspended in 20 mL of nuclei isolation buffer (NIB: 50 mM Tris-HCl pH 7.5, 25 mM KCl, 2 mM $MgCl_2$, 0.25 M sucrose) containing freshly added sodium butyrate (27 mM), PMSF (0.5 mM), and protease inhibitors (1 Roche Complete Mini™ protease inhibitor tablet was dissolved in 2 mL of sdH2O, and 800 μL was added to the NIB). The nuclei were further processed in a Wheaton 15 mL Dounce tissue grinder using the tight pestle for 10-20 passes, centrifuged at 2000 rpm in the HN-S centrifuge for 10 minutes at 4 °C, and the supernatant was discarded.

For both pituitary and placenta preparations, the isolated nuclei pellet was resuspended in HEPES buffer (10 mM HEPES pH 7.5, 3 mM $MgCl_2$, 10 mM NaCl, 1 mM PMSF) to approximately 20 A_{260} units/mL. Nuclei were equilibrated to room temperature and formaldehyde was added as a cross-linker at a final concentration of 1% for 5 minutes. The cross-linking reaction was quenched by the addition of glycine (final concentration 0.125 M).

2.2.1b Nuclear lysis and sonication of the chromatin.

Cross-linked nuclei were centifuged at 2000 rpm in the HN-S centrifuge for 10 minutes at 4 °C, and the supernatant was discarded. Nuclei were resuspended in RSB buffer (10 mM Tris-HCl pH 7.5, 3 mM $MgCl_2$, 10 mM NaCl, 1 mM PMSF, pH 7.5), centifuged at 2000

rpm in the HN-S centrifuge for 10 minutes, the supernatant was discarded, and nuclei were resuspended in lysis buffer (50 mM Tris-HCl pH 8.0, 1% SDS, 10 mM EDTA, 1 mM PMSF) with protease inhibitors (40 µL of a solution containing one Roche Complete Mini™ protease inhibitor tablet dissolved in 2 mL of sdH_2O) for 10 minutes at 4 °C. To shear the chromatin, the nuclear lysate was sonicated for a total time of 90 seconds (pituitary) or 120 seconds (placental) at 40% output (Vibra Cell™, Sonics and Materials Inc., Danbury, CT, USA) in 30-second pulses. To measure the total DNA content, the A_{260} of a sample was measured in 2 M NaCl / 5 M Urea. Sonicated material was then aliquoted into microfuge tubes and centrifuged at 10,000 rpm for 10 minutes at 4 °C to remove insoluble material. The supernatants were pooled and this soluble material represented the chromatin input for immunoprecipitation. To measure DNA content of the chromatin input, the A_{260} of a sample was assessed in 2 M NaCl / 5 M Urea. The % of DNA released from the nuclei was calculated as the (soluble A_{260} / total A_{260}) X 100%. The recovery of soluble chromatin ranged between 70 and 93% for all ChIP experiments.

2.2.1c Immunoprecipitation (IP) and DNA isolation.

IP conditions were based on a protocol obtained from the laboratory of Dr. James Davie (University of Manitoba, Winnipeg, Canada). A decision was made to increase the scale of the protocol by five times (5X). This was based on the observation that with the human tissue samples as a source of chromatin, the 1X protocol did not result in an amount of final DNA that was adequate for analysis.

For IP, 5 mL samples were prepared in dilution buffer (16.7 mM Tris-HCl pH 8, 167 mM NaCl, 1.2 mM EDTA, 1.1% Triton X-100, 0.01% SDS, 1 mM PMSF) with protease inhibitors (200 µL of a solution containing one Roche Complete Mini™ protease inhibitor tablet dissolved in 2 mL of sdH_2O) at an A_{260} of 2 units/mL. A 2 unit sample was put aside to represent the immunoprecipitation input. Samples were precleared for 3 hours at 4 °C using 300 μL of pre-treated (see below) protein A sepharose (Amersham Biosciences) and 25 μg of sheared salmon sperm DNA on a rotating platform. Centrifugation was done at 4000 rpm in a HNS centrifuge at 4 °C for 2 minutes. Specific antibodies were added for overnight incubation on a rotating platform as follows: 25 μL

anti-hyperacetylated (penta) histone H4 (Upstate biotechnology), 50 μL RFX1 (D-19) (Santa Cruz Biotechnology Inc.), and 50 μL NF-1 (H-300) (Santa Cruz Biotechnology Inc.). The following day, 300 μL of pre-treated protein A sepharose and 50 μg of sheared salmon sperm DNA was added for 1 hour on a rotating platform. Samples were washed for 10 minutes at 4 °C on a rotating platform in 5 mL of each wash buffer as follows: low salt buffer (20 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), high salt buffer (20 mM Tris-HCl pH 8, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), LiCl buffer (10 mM Tris-HCl pH 8, 0.25 M LiCl, 1 mM EDTA, 1% deoxycholic acid Na salt, 1% NP-40), and twice in TE buffer (10 mM Tris-HCl pH 8, 1 mM EDTA). The centrifugation for each wash was done at 4000 rpm in the HN-S centrifuge at 4 °C for 2 minutes. Samples were eluted twice in 1.25 mL of freshly prepared elution buffer (1% SDS, 0.1 M $NaHCO_3$) for 15 minutes at room temperature, and cross-links were reversed for 6 hours at 68 °C. DNA was isolated using QIAquick PCR purification kit (QIAGEN) according to manufacturers instructions. For each immunoprecipitated sample (bound), two columns were used, and for the input sample one column was used. Input DNA was eluted with 50 µL of EB buffer (QIAGEN), and each bound column was eluted with 30 µL of EB buffer and pooled (average final volume of bound DNA was 55 µL).

For pre-treatment of the protein A sepharose, 0.1 g of Protein A sepharose was suspended in 10 mL sdH2O and incubated at room temperature with for 30 minutes on a rotating platform. The sepharose was pelleted for 1 minute at 4000 rpm in the HN-S centrifuge, and resuspended in 10 mL of dilution buffer (16.7 mM Tris-HCl pH 8, 167 mM NaCl, 1.2 mM EDTA, 1.1% Triton X-100, 0.01% SDS, 1 mM PMSF) with protease inhibitors (200 µL of a solution containing one Roche Complete Mini™ protease inhibitor tablet dissolved in 2 mL of sdH_2O). Following a 10 minute incubation on the rotating platform at room temperature, the sepharose was again pelletted. The sepharose pellet was resuspended in a final volume of 1 mL dilution buffer containing 1 g BSA and 0.1 g sheared salmon sperm DNA. It was incubated overnight at 4 °C or for 1 hour at room temperature on a rotating platform. Sepharose was washed twice with dilution buffer and

resuspended in a final volume of 1 mL with dilution buffer (approximately 50% sepharose slurry) before use in the immunoprecipitations.

2.2.1d Analysis of the ChIP DNA.

ChIP PCR was done with 10 ng of input DNA or 5 μL bound DNA as template. PCR reactions (Taq DNA polymerase; QIAGEN) were at an annealing temperature of 55 °C for 28 cycles. PCR primer pairs are listed in Table 2.2. From a 50 µL PCR reaction, 20 µL of final product was run in 2% agarose gels. Images of the gels were analyzed by densitometry using the NIH Image 1.6 program. ChIP results are expressed as a ratio of bound/input to correct for possible differences between PCR primer pairs. The FGF-16 exon 3 primer pair was chosen as an internal standard that represented the background of non-specific sequences in a given immunoprecipitation (discussed further in chapter 4.1.5).

2.2.2 Gene Transfer

2.2.2a Transfection by calcium phosphate/DNA precipitation

For gene transfer, the calcium phosphate/DNA precipitation method was used, essentially as previously described (114).

Cells were plated at a density of $1x10^6$ cells per 100 mm plate 24 hours prior to the addition of DNA. For DNA precipitation, a 750 mL calcium solution (250 mM $CaCl_2$), containing 6 µg of epitope-tagged plasmid DNA, or alternatively, 30 µg of test (luciferase) plasmid, and either 3 µg of RSVp.CAT or 75 ng of pRL-Tkp.Luc (renilla luciferase) plasmid (Promega Corp.) as a control for DNA uptake, was 'bubbled' drop by drop into 750 mL of 2XHBSS solution (280 mM NaCl, 50 mM HEPES (FW 238.3), 1.5 mM dibasic Na_2HPO_4). The resulting solution was left to precipitate at room temperature for 30-45 minutes. A single precipitation was divided among three plates (480 µL per plate) to obtain a triplicate. In the reported results, n=1 is the activity from a single plate of a triplicate set. An experiment was defined by a minimum of one precipitation (n=3).

Cells were washed thoroughly with calcium- and magnesium-free phosphate buffered saline (PBS-CMF) 20-24 hours after the addition of DNA, and refed with growth media. Cells were harvested 48 hours after the addition of DNA. Each plate was harvested in 4 mL of PBS-CMF containing EDTA (1 mM), following a 4 mL PBS-CMF rinse. Cells were pelleted by centrifugation for 3 minutes at 3000 rpm at room temperature. Pellets were resuspended and lysed in 200-500 μL of Tris-triton buffer (100 mM Tris-HCL pH 7.8, 0.1% Triton X-100) for 15 minutes on ice, and centrifuged at 13,000 rpm for 15 minutes at 4 °C to remove insoluble material.

2.2.2b Luciferase assays

For luciferase assays, cell lysates were assessed immediately following harvest. The luciferase activity from 20 μL of cell lysate was determined using the Luciferase Assay System (Promega Corp.) or the Dual-luciferase assay system (Promega Corp.), according to manufacturers instructions, with a photon counting luminometer (EG&G Berthold, LUMAT LB9507 or ILA911 Luminometer, Tropix Inc., Bedford, MA, U.S.A.). For placental (JEG-3 and JAR) experiments, the cell lysate was sometimes diluted as 5 μL in a total volume of 20 μL with Tris-triton buffer.

2.2.2c Chloramphenicol acetyltransferase (CAT) assay

CAT activity was measured using a two-phase fluor diffusion assay (163), modified as previously described (164). Briefly, 50-200 μg of cell lysate protein was incubated in a 200 μL final volume with Tris-triton buffer at 70 °C for 10 minutes. The lysate was cooled to room temperature and 75 μl of ^{3}H-acetyl coenzyme A cocktail was added (0.5 μCi ^{3}H-acetyl CoA, 3.3 mM chloramphenicol, 0.1 M Tris-HCL pH 7.8). ScintiLene (3 mL, Fisher Scientific) was layered on top of the solution and the tubes were incubated for 30 minutes at 37 °C, before the activity was read in a scintilation counter. A minimum of five cycles (1 minute per tube) was evaluated. Values for CAT activity were determined by regression analysis to give cpm per minute per μg protein.

2.2.3 Identification of putative transcription factor binding sites.

To identify putative transcription factor binding sites in a given DNA element the MatInspector 2.2 program that is based on the Transfac 4.0 transcription factor binding site database was used (http://transfac.gbf.de/cgi-bin/matSearch/matsearch.pl). For the searches a matrix similarity ≥ 0.85 and a core similarity ≥ 0.85 were used as parameters (165, 166).

2.2.4 Immunoprecipitation and Protein (western) blotting

2.2.4a Immunoprecipitation

More than one IP protocol was used for the experiments that are reported. For the ChIP assays, the IP protocol was described in section 2.2.1c. A scaled down (1X) version of this protocol was also used for the experiments reported in Figure 4.5. In Figure 4.5 the immunoprecipitation input was 0.5 mg of total cellular lysate from JAR cells transiently transfected with cmyc-NF-1 and HA-RFX1 expression plasmids. Transiently transfected cells were harvested 48 hours after the addition of DNA and lysed in modified RIPA buffer (40 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% NP-40, 0.25% Na-deoxycholate, 1 mM EDTA, 1 mM PMSF, 1 mM Na_3VO_4, 1 mM NaF, 1 tablet/10 mL Complete™ Mini Protease inhibitor (Roche Diagnostics), and 1 μg/ml aprotinin) for 15 minutes at 4 °C. The total cellular lysates were pelleted to clear insoluble material, and protein concentrations were assessed using the Bio-Rad protein assay (Bio-Rad Laboratories Inc.). The final sepharose pellets from the IP were resuspended in 40 μL of Laemli sample buffer (2% SDS, 10% glycerol, 100 mM DTT, 60 mM Tris-HCl pH 6.8, and 0.001% bromophenol blue), boiled for 5 minutes, and 20 μL was resolved by SDS-PAGE in an 8% gel.

For Figure 3.21 the following protocol was used. Transiently transfected cells were harvested 48 hours after the addition of DNA and lysed in modified RIPA buffer for 15 minutes at 4 °C. The total cellular lysates were pelleted to clear insoluble material, and protein concentrations were assessed using the Bio-Rad protein assay. IP was carried out

with 0.5 mg of total cellular lysate. Lysates were precleared for 30 minutes at 4 °C using non-specific monoclonal antibodies or rabbit immunoglobulins, followed by a one-hour incubation with Pansorbin Cells (Calbiochem) prepared according to manufacturers instructions. Specific antibodies to RFX1 (1 μL, Reith RFX1 antibody (154), c-myc (2 μg) or, as an unrelated control, rat prostate DLP (1 μL), were incubated with precleared supernatants at 4 °C for 2 hours, and prepared Pansorbin cells were included for overnight incubation. Washes were performed first with supplemented NET Buffer (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 0.1% NP-40, 1 mM EDTA), followed by NET Buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% NP-40, 1 mM EDTA), and finally 10 mM Tris-HCl/0.1% NP-40. Pansorbin cell pellets were resuspended in 40 μL of Laemli sample buffer, boiled for 5 minutes, and 20 μL was resolved by SDS-PAGE in an 8% gel.

For the experiments reported in Figure 3.22, 100 μg of GC nuclear extract was incubated in a final volume of 1 mL with dilution buffer (16.7 mM Tris-HCl pH 8, 167 mM NaCl, 1.2 mM EDTA, 1.1% Triton X-100, 0.01% SDS, 1 mM PMSF) and protease inhibitors (40 μL of a solution containing one Roche Complete Mini™ protease inhibitor tablet dissolved in 2 mL of sdH_2O), and precleared with 1 μL of normal rabbit serum (NF-1 IPs) or 2 μL of normal goat serum (RFX1 IPs) for 30 minutes on a rotating platform at 4 °C. Prepared sepharose (50 μL, see section 2.2.1c) was added for 1 hour at room temperature, and 2 μg of NF-1 or RFX1 antibodies were added to the precleared supernatant overnight at 4 °C on a rotating platform. The following day, 100 μL of prepared sepharose was added for 3 hours at 4 °C on a rotating platform. Washes were performed first with supplemented NET Buffer, followed by NET Buffer, and finally 10 mM Tris-HCl/0.1% NP-40, for 5 minutes at 4 °C on a rotating platform. Final pellets were resuspended in 60 μL of Laemli sample buffer, boiled for 5 minutes, and 20 μL was resolved by SDS-PAGE in an 8% gel for RFX1 detection or a 12% gel for NF-1 detection.

2.2.4b Protein (western) blotting

For protein blotting, gels were transferred and immobilized on PVDF membranes (Roche Diagnostics) and blocked. Immunodetection was performed for one hour or overnight (c-myc antibodies) at room temperature (or 4 °C for Figure 3.21). Detection of antibody-antigen complexes was performed using BM chemiluminescence blotting substrate (POD, Roche Diagnostics) for Figure 3.21 or SuperSignal® West Pico chemiluminescent substrate (MJS Biolynx Inc.) (Figures 3.22 and 4.5) according to manufacturers instructions. Complexes were visualized on Kodak Biomax film (Amersham Biosciences).

Antibody dilutions for immunodetection were as follows, goat RFX1 antibodies, 1:10,000, mouse c-myc antibodies, 2 μg/mL, rabbit RFX1 antibodies, 1:1,000, rabbit NF-1 antibodies, 1:1,000, and rabbit Pit-1 antibodies, 1:1,000.

2.2.5 *In vitro* binding assays

2.2.5a Preparation of nuclear extracts

Nuclear protein extracts from GC, JAR, and JEG cells were made according to a published protocol (167) and dialysed as previously described (140). Briefly, cultured cells were harvested in PBS-CMF containing EDTA (1 mM), following a 4 mL PBS-CMF rinse, and pooled in a single 50 mL polystyrene tube. Cell pellets were resuspended in ice cold hypotonic buffer "A" (10 mM HEPES-KOH pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM dithiothreitol (DTT), 1 mM PMSF) at 5X the packed cell volume, and incubated for 10 minutes on ice. Cells were pelleted at 2000 rpm in the swinging bucket HN-S centrifuge at 4°C. The pellet was resuspended in ice cold hypotonic buffer "A" at 2X the packed cell volume, and homogenized in a 15 mL Dounce tissue grinder with the loose pestle until free nuclei were observed under the microscope. Nuclei were pelleted at 2000 rpm in the swinging bucket HN-S centrifuge at 4°C, and then 14,500 rpm in the JA-20 rotor of the Beckman J2-21 centrifuge for 20 minutes at 4 °C. The nuclear pellet was resuspended in high salt buffer "C" (20 mM HEPES-KOH pH 7.9, 25% (v/v) glycerol,

0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM DTT, 1 mM PMSF) at approximately 3 mL per 10^9 cells and homogenized with the Dounce tissue grinder. The nuclei were then incubated on a rotating platform at 4°C for 30 minutes before centrifugation in the JA-20 rotor of the Beckman J2-21 at 15,000 rpm for 30 minutes at 4 °C. The supernatant was collected and dialysed in Spectrapor dialysis tubing (6,000 – 8,000 Dalton cut-off) against buffer "D" (20 mM HEPES-KOH pH 7.9, 20% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM DTT, 1 mM PMSF) for 3-5 hours. Insoluble material was removed by centrifigation in the JA-20 rotor of the Beckman J2-21 at 15,000 rpm for 30 minutes at 4 °C. The nuclear protein was aliquoted in 50-500 µL volumes, quick frozen in a dry ice ethanol bath, and stored at -70°C. Protein concentration of the extracts was assessed using the Bio-Rad Protein Assay (Bio-Rad Laboratories Inc.) with bovine serum albumin as a standard.

2.2.5b Electrophoretic mobility shift assay (EMSA)

For EMSA, competitors were preincubated with 2-5 µg of nuclear protein and 1-2 µg poly dIdC in reaction buffer (10 mM Tris-HCl pH7.5, 50 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA, 5% glycerol) or (20 mM HEPES-KOH pH 7.9, 20% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM DTT, 1 mM PMSF) for 10 minutes at room temperature or 20 minutes on ice in a 20 µl final volume. Double-stranded competitor oligonucleotides were used in mass excess of the probe amount unless otherwise indicated. Antibody and normal rabbit serum competitors were used at 1 µL. Following preincubation, 1 ng of double-stranded radiolabelled oligonucleotide probe was added, and the reactions were incubated a further 10 minutes at room temperature or 20 minutes on ice before electrophoresis in a 4-5% acrylamide gel.

2.2.5c Nuclease protection

For nuclease protection, competitors were preincubated with 15-27 µg of nuclear protein (diluted with buffer "D" to 12.5 µL final volume), 1 µg poly dIdC , and 6.25 mM $MgCl_2$, in a 19 µL final volume for 15 minutes on ice. Double-stranded competitor oligonucleotides were used in pmole excess of the probe instead of mass excess because

of the size difference between the competitors and the probes. Following preincubation, 0.5-1.0 ng of probe was added to each reaction and incubated for an additional 15 minutes on ice. For the 263P fragment, it was radiolabelled at the 3´ end and 0.5 ng equaled 0.0028 pmoles. The CS-A promoter probe was a *PstI*/*BamHI* fragment of the CS-A 5′-flanking DNA radiolabelled at the *BamHI* site (1.0 ng equaled 0.006 pmoles). DNase digestion was carried out at 26 °C for 1 minute by adding 24 μL of DNase solution (7.8 mM $MgCl_2$, 3.9 mM $CaCl_2$, 4.3 mM HEPES-KOH pH 7.9, 4.3% glycerol, 2.2 mM KCl, 0.04 mM EDTA, 1 unit DNase). Reactions were stopped by the addition of 160 μL of STOP buffer (0.1 M Tris-HCl pH 7.5, 12.5 mM EDTA, 0.15 M NaCl, 1% (w/v) SDS, 1250 μg proteinase K, 100 μg tRNA), followed by incubation at 37 °C for 30 minutes to digest the proteins. Solutions were phenol extracted and ethanol precipitated. DNA pellets were resuspended in 5 μL of formamide loading buffer (80% (v/v) formamide, 1 mM EDTA, 0.1% (w/v) xylene cyanol, 0.1% (w/v) bromophenol blue) and electrophoresed in a 6% acrylamide gel.

2.2.6 Plasmid DNA

The plasmids pxp1 and pT81luc (Tkp.Luc) were obtained from Dr. Steve Nordeen (Molecular Biology Program, Department of Pathology, University of Colorado Health Sciences Center, Denver, Colorado, USA) (168). The plasmid RSVp.CAT contains the chloramphenicol acetyl transferase gene directed by the Rous sarcoma virus promoter and was described elsewhere (169). The NF-1 cDNA plasmids (pBSNF-1 set) were a generous gift from Dr. Masayoshi Imagawa (Laboratory of Environmental Biochemistry, Graduate School of Pharmaceutical Sciences, Osaka University, Osaka, Japan) (170). The RFX1 cDNA was a generous gift of Dr. Walter Reith in the plasmid pSG5RFX1(171). The plasmids pRLTkp.Luc (Promega), pCMV-Myc (Clontech), and pCMV-HA (Clontech) were purchased.

2.2.6a Plasmid contructions

Restriction endonuclease digested plasmid vectors and inserts were isolated from agarose gels with a GFX PCR DNA and gel band purification kit (Amersham). Ligations were incubated at room temperature for 6 hours to overnight.

263PTkp.Luc

For 263PTkp.Luc, the 263P fragment had been previously subcloned into pUC19 (133). To construct 263P/TKp.Luc, the 263P fragment was released from 263P/pUC19 using *Sac*I/*Hin*dIII and subcloned into pT81luc.

CSp.Luc plasmids

The immediate 5′ flanking region of the CS-A gene (-492/+6) was isolated as a *Eco*R1/*Bam*H1 fragment (blunted) and inserted into the *Hin*dIII site (blunted) upstream of the firefly luciferase (Luc) gene in pxp1.

For PSE-A3CSp.Luc, PSE-B4CSp.Luc, NF-1CSp.Luc or NF1CSp.Luc, PSE-B4mCSp.Luc, PSE-B4/CmCSp.Luc, PSE-B4/GmCSp.Luc, PSE-B4/CGmCSp.Luc, the appropriate double-stranded oligonucleotides were phosphorylated and inserted into the *Sma*I site of CSp.Luc.

For 103PCSp.Luc, the 103P fragment was generated by PCR using 103P(F) and 103P(R) as primers. The PCR product was ethanol precipitated, digested with *Bam*HI, isolated, and inserted into the *Bam*HI site of CSp.Luc.

For 263PCSp.Luc, 263P was released form 263P/TKp.Luc as a *Bam*HI/*Xho*I fragment and subcloned into CSp.Luc.

To generate 263Am2CSp.Luc, 263BmCSp.Luc, 263Am2/BmCSp.Luc, a two-step PCR approach with 263P primers, appropriate mutant oligonucleotides, and 263PCSp.Luc as an initial template, was utilized to introduce mutations into 263P (Taq DNA polymerase;

QIAGEN). The final PCR products were inserted as *Bam*H1 fragments upstream of the CS-A promoter in CSp.Luc.

To generate CSpPit-1m.Luc and 263CSpPit-1m.Luc, the CSp.Luc and 263PCSp.Luc plasmids were digested with *Nsi*I, staggered 3′ ends were removed with Klenow fragment, and the plasmids were religated.

Epitope-tagged plasmids

Hemagglutinin (HA) or c-myc (myc) tagged transcription factors (HA-RFX1, myc-RFX1, myc-NF-1A, myc-NF-1B, myc-NF-1C, and myc-NF-1X) were generated using the epitope-tagged mammalian expression vector set (Clontech). NF-1 cDNAs were isolated from pBSNF-1 vectors and subcloned into *Xho*I/*Not*I of pCMV-Myc (Clontech). The RFX1 cDNA was isolated from pSG5RFX1, blunted with Klenow fragment and subcloned into the (blunted) *Sal*1 site in pCMV-Myc and pCMV-HA (Clontech).

2.2.6b Preparation of competent bacterial cells and bacterial transformation

The DH5α strain of *Escherichia. coli* was used for the propagation of all plasmid DNA. To introduce plasmid DNA into the cells, the CaCl method was used. Briefly, bacteria was grown overnight in L broth at 37 °C, recultured the following day, and grown to log phase (approximately 2 1/2 hours, A_{260} of 400-600). Cells were pelleted, resuspended in cold 50 mM CaCl, and incubated on ice for 20 minutes. Cells were again pelleted, resuspended in CaCl buffer (50 mM CaCl, 10 mM Tris-HCl pH 8.0, 15% glycerol), and 100 µl aliquots were stored at -70 °C until needed. To introduce plasmid DNA, an aliquot of preprepared bacteria was thawed on ice for 10 minutes, DNA was added, and cells were incubated for 45-60 minutes on ice. Bacteria was heat shocked for 1 minute at 42 °C, transferred to 1 mL of L broth, and incubated at 37 °C for 1 hour, before plating onto LB-Agar plates containing antibiotics.

2.2.6c Isolation of plasmid DNA from bacteria

Plasmid DNA was isolated from bacteria by akaline lysis. For small scale plasmid isolation, single colonies were inoculated into 2 mL of L broth with antibiotics and grown

at 37 °C for 6 hours to overnight. Cells were pelleted, resuspended (25 mM Tris-HCl pH 8.0, 10 mM EDTA, 50 mM glucose), lysed for 5 minute on ice (0.2 M NaOH, 1% SDS), and proteins were precipitated for 5 minutes at room temperature (3 M Na acetate pH 5.2). Following centrifugation (5-10 minutes, 13,000 rpm 4 °C), the supernatant was phenol extracted and ethanol precipitated. For large scale plasmid isolation, QIAGEN maxi filters were used according to manufacturers instructions. Final DNA pellets from both techniques were resuspended in TE buffer (0.1 M Tris-HCl pH 8.0, 0.01 M EDTA).

2.2.6d Sequencing of plasmid DNA

All plasmid constructs were sequenced either by the fmol PCR sequencing kit (Promega) according to manufacturers instructions, or through the University Core DNA Services, University of Calgary.

2.2.7 RNA

2.2.7a RNA isolation

RNA was isolated from GC cells using guanidinium thiocyanate. Cells, cultured in 100 mm dishes, were rinsed with cold PBS and harvested in 1.8 mL of guanidium SCN (4 M guanidium isothiocyanate, 25 mM Na citrate pH 7.0, 0.5% (v/v) sarcosyl, 100 mM β-mercaptoethanol). To the cell lysate, 180 μL of 2 M Na acetate pH 4.0 and 1.8 mL sdH_2O-saturated phenol were added and vigorously mixed. To this, 360 μL of IAC (24:1 chloroform and isoamyl alcohol) was added and the mixture was incubated on ice for 20 minutes before being centrifuged in the JA20 rotor of the Beckmann J2-21 centrifuge at 10,000 rpm for 15 minutes at 4 °C. The supernatant was collected and RNA precipitated with an equal volume of isopropanol. RNA was pelleted in the JA20 rotor of the Beckmann J2-21 centrifuge at 10,000 rpm for 15 minutes at 4 °C, and resuspended in 400 μL of TE buffer. The resuspended RNA was phenol extracted with an equal volume of sdH_2O-saturated phenol, 40 μL of 2 M Na acetate pH 4.0, and 80 μL of IAC, before precipitation in isopropanol. The final RNA pellet was resuspended in sdH_2O at 65-100 °C.

2.2.7b PolyA$^+$-enrichment

A cellulose oligo-dT slurry was used to enrich for poly A$^+$ RNA. The oligo-dT was hydrated for 1 hour in 1X binding buffer (20 mM Tris HCl pH 7.5, 0.5 M NaCl, 1 mM EDTA, 0.1% SDS), washed twice and resuspended in this same buffer. An equal volume of 2X binding buffer was added to the total RNA. In general, 3-5 mg of total RNA was used, and 20 mg of dry oligo-dT was required for every mg of total RNA. Excess 1X binding buffer was removed from the hydrated oligo-dT, and the RNA was added. The slurry was incubated for 1 hour at room temperature on a rotating platform. Oligo-dT was then pelleted, washed twice in 10 mL of 1X binding buffer, and resuspended in an approximately equal volume of 1X binding buffer. To isolate the PolyA$^+$-enriched RNA, spin columns were used, and the RNA isolated in elution buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.05% SDS). RNA was stored under ethanol at -70 °C.

2.2.7c RNA (northern) blotting

PolyA$^+$-enriched RNA, 25 μg per lane, was run in a 1.5% formaldehyde-agarose gel and transferred to nitrocellulose. Following a 16-18 hour transfer, nitrocellulose was air dried and baked at 80 °C for 45 minutes – 1 hour. Blots were incubated with prehybridization buffer (50% formamide, 5X SSC, 5X Denhardts solution (50X - 1% (w/v) Ficoll, 1% polyvinyl pyrollidone, 1% (w/v) BSA), 50 mM $NaPO_4$ pH 6.5, 0.1% SDS, 100 µg/mL yeast tRNA, 25 µg/mL salmon sperm DNA) at 42 °C for 16-24 hours. Radiolabelled probes were prepared in probe buffer (1 mg/mL salmon sperm DNA, 1 mg/mL tRNA) and incubated with the blot in hybridization buffer (50% formamide, 5X SSC, 5X Denhardts solution, 50 mM $NaPO_4$ pH 6.5, 10 µg/mL yeast tRNA, 25 µg/mL salmon sperm DNA, 10% (w/v) dextran sulphate) for 24 hours. Blots were washed at 65 °C, 2-3 times for 15 minutes with a solution of 0.1X SSC / 0.1% SDS.

2.2.7d NF-1 specific probes

Rat cDNA's were kindly provided by Dr. M. Imagawa (170). Inserts were released from the pBSNF-1 plasmid set as *Eco*RI/*Xho*I fragments and gel purified. Specific probes for

each of the rat NF-1 genes were made by restriction digests of the cDNA's and isolation of carboxyl-terminal fragments (NF-1A/*Xcm*I 918 bp, NF-1B/*Pvu*II 599 bp, NF-1C/*Dde*I 207 bp, NF-1X/*Bgl*II 590 bp). Probes were radiolabelled by random-priming (Promega Prime-A-Gene). Each probe was assessed for cross-reactivity/specificity against 1 and 0.1 ng of rat NF-1 cDNAs (A, B, C and X) immobilized on nitrocellulose using a slot blot apparatus.

2.2.8 Statistical analysis

Statistical analysis of the data was done using a two-tailed, unpaired Student t test. A value of $p<0.05$ is considered statistically significant. In Figures, *<0.05, **<0.01 and ***<0.001.

CHAPTER 3

IDENTIFICATION OF P SEQUENCE FACTORS (PSFs) *IN VITRO*

3.1 Functional characterization of 263P *in vitro*.

3.1.1 P sequences repress CS-A promoter activity in pituitary GC cells in the absence of the CS-B downstream enhancer region.

The original observation of 263P repressor activity in pituitary cells was demonstrated with a construct that contained not only 263P, but also the 241 bp CS-B downstream enhancer (133). A subsequent study of the CS-B downstream enhancer, indicated that this element also repressed CS promoter activity in pituitary cells (172). To determine if 263P was a functional repressor in the absence of the CS-B downstream enhancer, a plasmid was constructed with -492/+6 of CS-A gene 5´-flanking DNA (CS-Ap) driving expression of a luciferase reporter gene (CSp.Luc). The 263P fragment was inserted upstream of CS-Ap to generate 263PCSp.Luc, and luciferase activity of both constructs was assessed in transiently transfected GC cells. Presence of 263P upstream of CS-Ap resulted in significant repression of luciferase activity. When CSp.Luc activity is arbitrarily set to 100%, the luciferase activity of 263PCSp.Luc averaged at 44% through 13 separate experiments (range of 20-68%). The definition of a single experiment can be found in Chapter 2.2.2. In Figure 3.1, a representative experiment is shown where 263PCSp.Luc activity was 43% of CSp.Luc ($n=12$, $p<0.0005$). *These experiments established that the 263P fragment functions as a repressor of CS-A promoter activity in rat pituitary GC cells in the absence of the CS-B downstream enhancer region.*

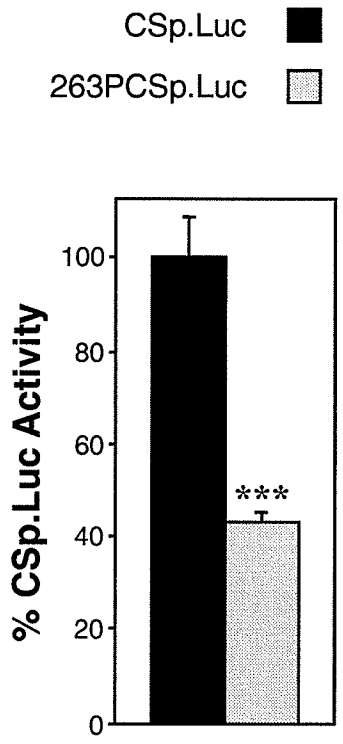


Figure 3.1: P sequences repress CS-A promoter activity in pituitary GC cells in the absence of the CS-B downstream enhancer region.
Hybrid luciferase (Luc) genes were used to assess the effect of 263P on CS-A promoter (-496/+6, CSp.Luc) activity in pituitary GC cells. A representative experiment is shown. To control for DNA uptake, cells were co-transfected with pRLTkp.Luc. Corrected values are expressed as percentages of CSp.Luc activity, which was arbitrarily set to 100%. The basal firefly Luc/Renilla Luc value for CSp.Luc was 3.22 ± 0.15 (n=12). Bars represent standard error of the mean (SEM). *** $p<0.0005$.

3.1.2 Detection of significant P sequence function in placental cell lines.

A recent analysis of human term placental chromatin demonstrated histone hyperacetylation associated with 263P (150). In this same study, 263P was also implicated as an enhancer of reporter gene expression in transgenic mouse placenta (150). These data suggest that P sequences may be involved in regulating expression of the GH/CS locus in the placenta. The potential for an *in vitro* model system of P sequence placental enhancer activity was assessed using human choriocarcinoma cell lines and transient transfection experiments.

The original characterization of 263P in placental choriocarcinoma JEG-3 cells observed reductions of CS-A promoter activity in the presence of 263P that were not statistically significant (133). The inclusion of the CS-B enhancer in these constructs, however, may have masked the ability to observe significant P sequence function. The CS-B enhancer element is a potent activator of CS promoter activity in JEG-3 cells (136, 139). To examine the effect of 263P on CS-Ap activity in the absence of the CS-B enhancer region, the hybrid luciferase constructs utilized in GC cells were introduced into human choriocarcinoma (JEG-3 and JAR) cell lines by transient transfection (Fig. 3.2). In both cell lines, the presence of 263P significantly repressed CS-A promoter activity. In JEG-3 cells, 263PCSp.Luc had 76% the luciferase activity of CSp.Luc ($n=15$, $p<0.0005$), while in JAR cells repression of CSp.Luc activity to 66% was observed ($n=12$, $p<0.0005$). *These experiments demonstrate significant effects of 263P on CS-Ap activity in placental cell culture lines.* The *in vitro* activity, however, does not reflect the reported enhancer activity that was observed in transgenic mouse placenta (150). Taken together, these observations also suggest that the P sequences and the CS enhancer region have the capacity to modify each others functional activity.

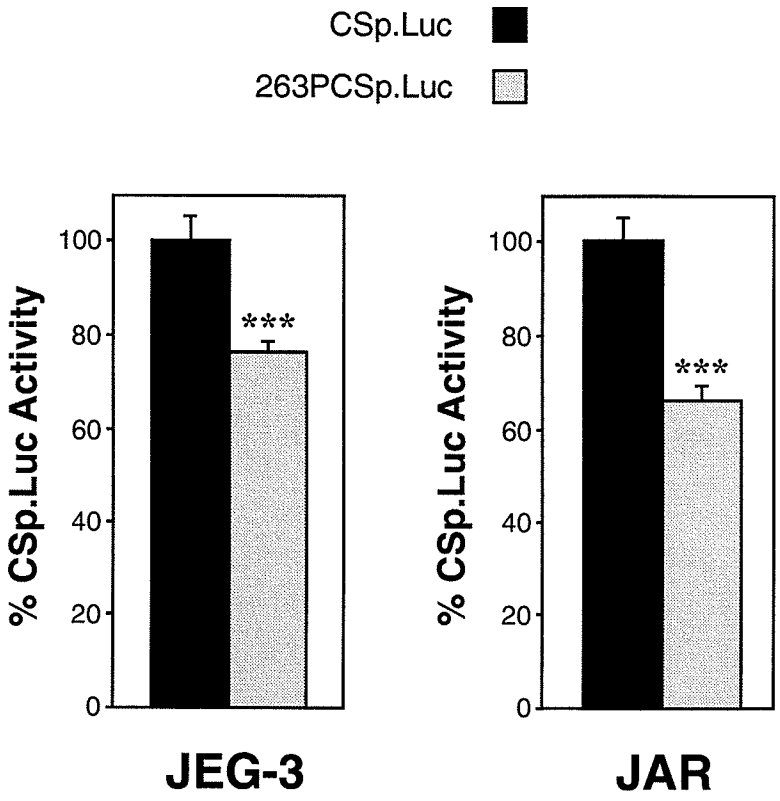


Figure 3.2: Functional effects of 263P on CS-A promoter activity can be detected in placental cells in the absence of the CS-B downstream enhancer region.
Hybrid luciferase (Luc) genes were used to assess the effect of 263P on CS-A promoter (-496/+6, CSp.Luc) activity in human placenta choriocarcinoma cell lines. To control for DNA uptake cells were co-transfected with pRLTkp.Luc (JAR and JEG-3) or RSVp.CAT (JEG-3). Corrected values are expressed as percentages of CSp.Luc activity, arbitrarily set to 100%. Bars represent SEM. ***p<0.0005.

3.2 Identification of candidate P sequence factors (PSFs) using PSE-A and PSE-B

Several pieces of both *in vitro* and *in vivo* data indicate that P sequences may play an important role in transcriptional regulation of the human GH/CS gene family (83, 130, 133, 150, 151). Identification of the factors (PSFs) that associate with this DNA was the first step taken to characterize P sequences. As previously indicated, PSE-A and PSE-B are the two regions where the direct interaction of proteins with 263P can be visualized by nuclease protection (Fig. 3.3). This nuclease protection pattern is competed with either

PSE-A or PSE-B oligonucleotides (Fig. 3.3A) and (130, 133, 151). Based on the nuclease protection pattern, there are several interpretations of the P sequence complex (Fig. 3.3B). It is possible that a single protein makes contact with the DNA at both PSE-A and PSE-B. Alternatively, it is possible that PSE-A and PSE-B represent occupation of the DNA by two separate proteins that are related (in that they recognize similar DNA elements). A third possibility is that the proteins that occupy PSE-A and PSE-B are different, but interact with each other. In this third model, the disruption of one binding event disrupts the second binding event due to these interactions.

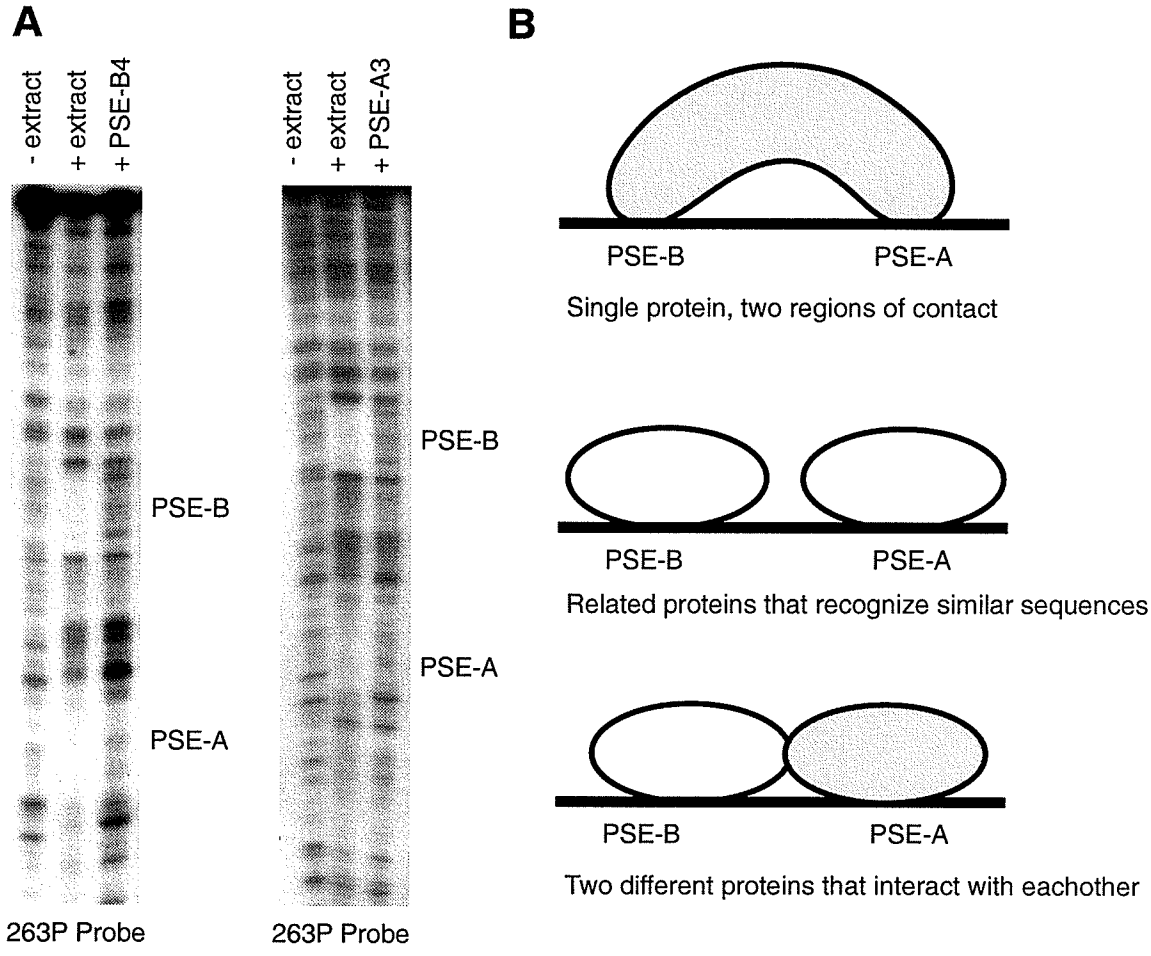


Figure 3.3: Possible interpretations of the 263P complex based on competition of the nuclease protection pattern.
(A) The PSE-A and PSE-B regions of protein-DNA interaction on 263P are visualized by nuclease protection assays. For these assays, 263P was radiolabelled and incubated with or without pituitary GC nuclear extract and competitor oligonucleotides, before digestion with DNaseI. The PSE-B4 and PSE-A3 competitors were included at 5000-fold pmole excess of 263P probe. Data reported in (130, 151) **(B)** Interpretations of the P sequence complex based on the nuclease protection pattern.

Nuclease protection of PSE-A and PSE-B was not originally reported with human term placenta or choriocarcinoma JEG-3 nuclear extracts (133). Nuclease protection patterns

with these extracts can be more difficult to observe, as the endogenous nuclease activity in placental extracts appears to be higher compared to other nuclear extracts (Norquay, unpublished observation). A nuclease protection pattern similar to that seen with GC and other nuclear extracts was reported for JEG-3 in the Ph.D. thesis of Dr. R.M. Surabhi (173). When 263P was used as a probe with a human choriocarcinoma JAR nuclear extract, a duplication of the GC nuclease protection pattern was not observed (Fig. 3.4). There were, however, important similarities between the JAR and GC patterns. The two hypersensitive sites that surround the PSE-B region with GC nuclear extract are evident with JAR nuclear extract (Fig. 3.4, closed arrowheads), indicating the possibility of protein binding near or at PSE-B. In addition, PSE-A I and PSE-A III bands are observed with both GC and JAR nuclear extracts, while the PSE-A II band that is observed without nuclear extract (free probe) is reduced with both extracts (Fig. 3.4, open arrowheads). *The similarities in the 263P nuclease protection patterns with GC and JAR extracts suggests that PSE-A and PSE-B regions may be useful in an attempt to identify P sequence binding factors in placental cells.*

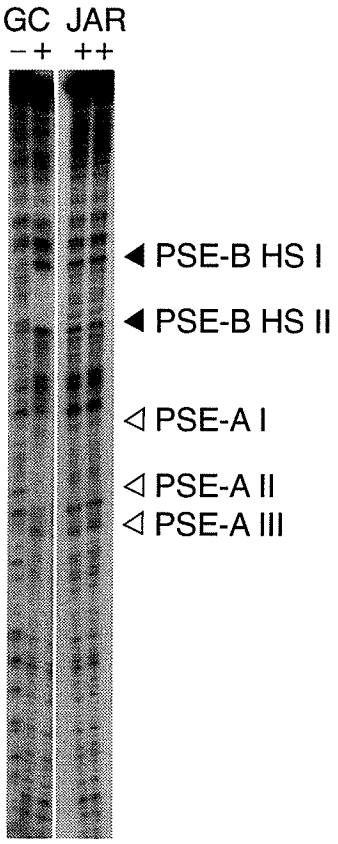


Figure 3.4: Comparison of the nuclease protection patterns generated with placental JAR and pituitary GC nuclear extract.
263P was radiolabelled and incubated with (+) or without (-) nuclear extract before digestion with DNaseI. GC nuclear extract 15 μg, JAR nuclear extract 20 and 27 μg. PSE-B hypersensitive (HS) sites are indicated by closed arrowheads, and the PSE-A bands of interest by open arrowheads.

Identification of the proteins that interact with PSE-A and PSE-B was the starting point for determining P sequence complex participants. To meet this aim, cell lines for standard *in vitro* structural (EMSA, nuclease protection) and functional (transient gene transfer) assays were used. Following the identification of candidate PSFs using PSE-A and PSE-B, the approach was expanded to assess the ability of these factors to participate in the 263P complex, *in vitro* as well as *in situ*. The pituitary GC cell line was used as the primary system for PSF identification, as a more consistent nuclease protection pattern,

and greater repression of CS-Ap activity, was observed in the GC cell line than in the placental JAR or JEG-3 cell lines.

3.2.1 PSE-A is a composite element with the capacity for mutually exclusive regulatory factor X-1 (RFX1) and nuclear factor-1 (NF-1) binding.

3.2.1a PSE-A contains putative binding sites for several transcription factors.

Putative PSE-A binding proteins were determined using two parameters: (i) presence of known transcription factor binding sites in the PSE-A sequence and (ii) the expression pattern of the putative factor. The 37 bp sequence of the PSE-A nuclease protection region was compared with known transcription factor binding sites using the MatInspector 2.2 program and TRANSFAC 4.0 database (165, 166). The parameters for this analysis were a matrix similarity of 0.85 and a core similarity of 0.85. Within PSE-A, similarities to the binding sites of HNF-3 forkhead homolog-2 (HFH-2), Nkx2.5, NF-1, and RFX were found (Table 3.1).

Table 3.1
Results of the database search for putative PSE-A binding proteins.

PSE-A

(+) 5′-TCAATGTTGGTTGCCAACACCACTGCCAACCACTTCT-3′
(-) 3′-AGTTACAACCAACGGTTGTGGTGACGGTTGGTGAAGA-5′

Factor	Core Similarity	Matrix similarity	Location
NF-1	1.0	0.947	(-) tgtTGGCaaccaacattg
NF-1	1.0	0.914	(-) ggtTGGCagtggtgttgg
RFX1	1.0	0.909	(-) cagtggtgttgGCAAcca
Nkx2.5	1.0	0.884	(-) agAAGTg
HFH-2	1.0	0.864	(+) caaTGTTggttg

The location of the match is designated (+) or (-) to identify the strand where the match occurs.

HFH-2 is a member of the winged helix family of transcription factors (also known as the hepatocyte nuclear factor-3 (HNF-3)/forkhead family of transcription factors) (174). HFH-2 is also known as Genesis, and its expression is limited to non-differentiated embryonic stem cells (175), embryonal carcinoma cells (175), and neural crest cells of the developing mouse embryo (176, 177). *Although PSE-A contains an HFH-2 consensus sequence that varies at only two nucleotides (165, 166), the limited expression pattern of this factor reduces its priority as a candidate PSE-A binding protein in the pituitary or placenta.*

Nkx2.5 is a homeodomain factor that is also referred to as cardiac specific homeobox (Csx) (178, 179). Nkx2.5 is the mammalian homolog of the *Drosophila* tinman gene that is required for heart development (180). It is expressed specifically in the heart, and Nkx2.5 transcripts were not detected in either the brain or placenta (178, 179). *Therefore, despite the presence of a putative Nkx2.5 binding site in PSE-A, based on its expression pattern, Nkx2.5 is not a high priority candidate PSE-A binding protein.*

The MatInspector search identified two putative NF-1 sites in PSE-A (Table 3.1). The expression pattern of NF-1 is assumed to be ubiquitous, based on the observation of NF-1 transcripts in several rat tissues (liver, kidney, lung, intestine, and brain) (170), and the role of NF-1 in the expression of more than 100 cellular and viral genes (181). Expression of NF-1 in pituitary and placental cells has also been previously demonstrated. NF-1 was shown to be involved in the regulation of the pituitary-specific transcription factor Pit-1 (182), as well tissue-specific expression of the human CYP11A1 gene (human placental cell lines included) (183). *Thus, the identification of NF-1 sites in PSE-A, and the expression of NF-1 in both pituitary and placenta cell lines, indicate that NF-1 is a high priority candidate PSE-A binding protein.*

NF-1 binds to a full-length consensus sequence (5′-TGGN$_6$GCCAAT-3′) (184), as well as to half-sites (5′-TTGGC-3′ or 5′-GCCAA-3′) with a reduced affinity (185). Within the consensus sequence, two regions were shown to be essential for NF-1 binding (the core regions) (186, 187), and these two core regions are shaded in the NF-1 consensus

that is shown in Figure 3.5A. *The MatInspector search identified two putative NF-1 half-sites in PSE-A (Table 3.1 and shown in Fig. 3.5A). Further analysis of the sequence also suggested the presence of a putative full-length NF-1 site in PSE-A that varied in one of the core nucleotides (Fig. 3.5A).*

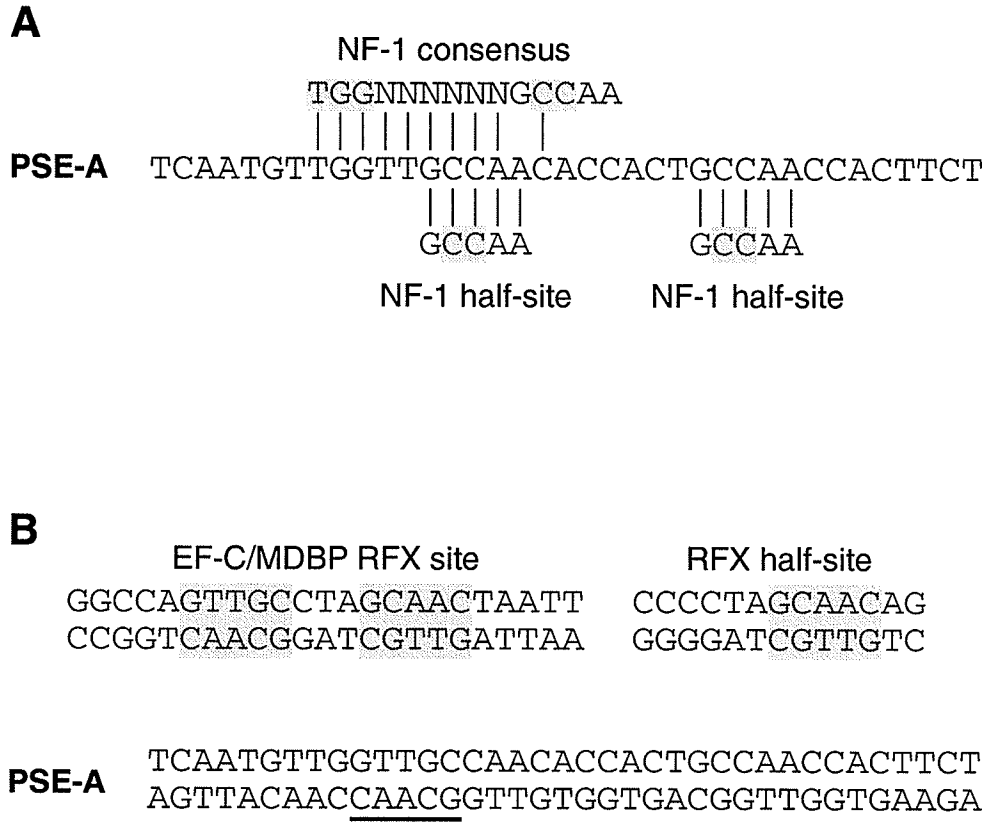


Figure 3.5: PSE-A contains putative binding sites for the NF-1 and RFX transcription factor families.
(A) The sequence of PSE-A is shown with the consensus NF-1 site aligned above. Core regions within the NF-1 consensus sequence, determined to be essential for the association of NF-1, are highlighted (186, 187). Two half-sites for NF-1 binding in PSE-A are also shown. **(B)** The 23 bp EF-C/MDBP RFX site is shown with the inverted repeat highlighted. Also shown is the 13 bp RFX half-site (161). A putative RFX half-site in PSE-A is underlined.

The RFX family in mammals consists of five members, RFX1-5 (188), three of which (RFX1, RFX2, and RFX3) are ubiquitously expressed (155). Analysis of RFX1, RFX2, and RFX3 expression, demonstrated transcripts of all three genes in both mouse placenta and brain (155). The RFX family shares a homologous winged-helix DNA binding domain, that recognize X-box motifs (5′-GTNRCC(0-3N)RGYAAC-3′) (189). The minimum sequence requirements for recognition of a target site by the RFX family were defined using a recombinant RFX1 DNA binding domain (161). Relative to the consensus site of NF-1, the X-box is more loosely defined and can show considerable variability (161, 166, 190). Two RFX1 molecules will associate with a 23 bp inverted repeat oligonucleotide (the EF-C/MDBP target site), while RFX1 monomers are also able to recognize and bind to half-sites, however, with less affinity (161). The MatInspector analysis identified one putative RFX binding site that corresponds to a half-site (Table 3.1 and shown in Fig. 3.5B). *Thus, the identification of a putative X-box half-site, and the ubiquitous expression pattern of RFX1, RFX2, and RFX3, indicates that these factors are also good candidates for PSE-A binding proteins.*

3.2.1b A truncated PSE-A oligonucleotide (PSE-A3) retains PSE-A EMSA complexes.

The putative NF-1 and RFX sites do not span the entire PSE-A region. EMSA was used to determine if a smaller PSE-A oligonucleotide could be used for analysis of candidate PSE-A binding proteins (Fig. 3.6). EMSA is a very sensitive assay, and the complexes that form on an EMSA probe represent the binding events that occur on that specific fragment of DNA with a particular nuclear extract. A 40 bp oligonucleotide consisting of the entire PSE-A nuclease protection region and 3 additional 5′-flanking nucleotides was generated (PSE-A5) and used as an EMSA probe with pituitary GC nuclear extracts. The PSE-A5 probe was compared to a 28 bp PSE-A oligonucleotide (PSE-A3) that contained all putative NF-1 and RFX sites. When PSE-A3 and PSE-A5 were used as EMSA probes under identical conditions, the same number of complexes were observed with GC nuclear extracts (Fig. 3.6). The basis for this analysis was the assumption, that if PSE-A5 contained sequence information that corresponded to the association of additional binding factors, this would be indicated by the observation of additional EMSA complexes with

the PSE-A5 probe when compared to the PSE-A3 probe. *Based on the comparison of PSE-A3 and PSE-A5 EMSA complexes, PSE-A3 was determined to be an adequate fragment for the identification of PSE-A binding proteins.*

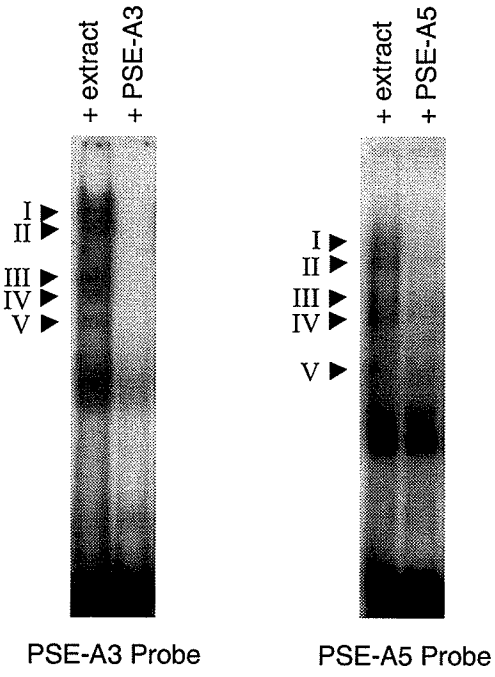


Figure 3.6: EMSA patterns of PSE-A3 and PSE-A5 probes with GC nuclear extract. PSE-A3 and PSE-A5 were radiolabelled and used as probes for EMSA with pituitary GC nuclear extract. Unlabelled oligonucleotides were used as competitors at 100-fold mass excess of probe.

3.2.1c Scanning mutations define the core region within PSE-A3 for pituitary complex formation.

The complexes seen when PSE-A3 was used as an EMSA probe with GC nuclear extract (I-V) were defined as specific complexes. This was due to the observation that complexes I-V were efficiently competed by unlabelled PSE-A3 (5 – 50-fold mass excess of probe), (Fig. 3.7, arrowheads) and not a 50-fold mass excess of an unrelated competitor oligonucleotide PSE-B4. It should be noted that complex V was only detected in longer exposures of both PSE-A3 and PSE-A5 EMSAs.

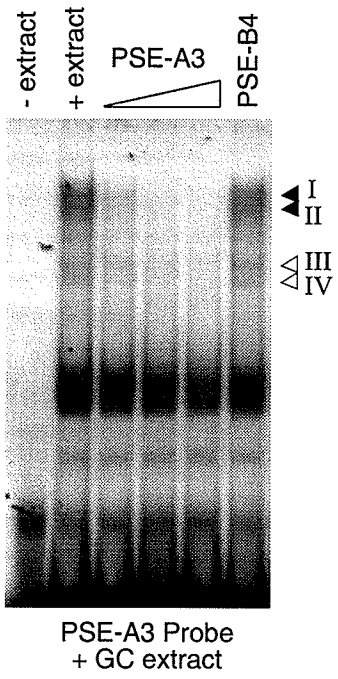


Figure 3.7: Specific PSE-A3 EMSA complexes with pituitary GC nuclear extract. PSE-A3 was radiolabelled and used as an EMSA probe with pituitary GC nuclear extract. Specific complexes I-IV are indicated by arrowheads and are defined by the ability to compete these complexes with unlabelled PSE-A3 (5-fold, 25-fold, and 50-fold mass excess of probe). As a control, PSE-B4 was also used as a competitor at 50-fold mass excess of probe. Data reported in (130).

A series of scanning mutations were created within PSE-A3 to determine the essential or core sequences required for the formation of these specific complexes (Fig. 3.8A). MatInspector analysis of the mutations did not detect the creation of additional known transcription factor binding sites (165, 166). The scanning mutations were then utilized as EMSA competitors with GC nuclear extract and a PSE-A3 probe (Fig. 3.8B). A core region in PSE-A3 was defined by analysis of complexes I and II. PSE-A3m1, PSE-A3m4, PSE-A3m5, and PSE-A3m6 were all efficient competitors of complex I, with incomplete competition of complex II. In contrast, PSE-A3m2 and PSE-A3m3 were unable to compete for either complex I or II. This region therefore appears to be essential for their formation. *The core region of PSE-A3 localizes to the putative X-box half site, as well as a putative NF-1 binding region consisting of a potential full length- and half-site (Fig. 3.5).*

A

```
tgttggttgccaacaccactgccaacca PSE-A3
CTAGggttgccaacaccactgccaacca PSE-A3m1
tgttTCGAgccaacaccactgccaacca PSE-A3m2
tgttggttTAACacaccactgccaacca PSE-A3m3
tgttggttgccaGTCTcactgccaacca PSE-A3m4
tgttggttgccaacacATACgccaacca PSE-A3m5
tgttggttgccaacaccactgcGTTAca PSE-A3m6
```

B

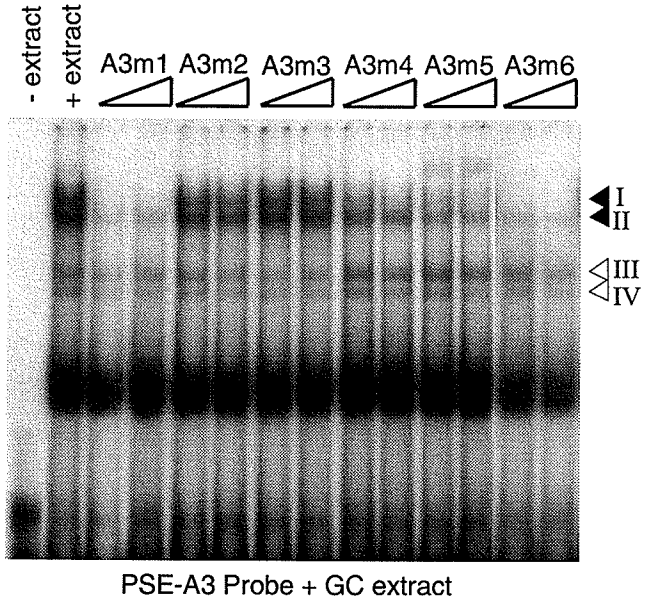


Figure 3.8: Scanning mutations define the core region within PSE-A3 for pituitary complex formation.

(A) The sequences for PSE-A3 and the scanning mutations are shown. Mutated sequences are indicated by upper case with the name of each mutant to the right of its corresponding sequence. The core region within PSE-A3 is highlighted. **(B)** Radiolabelled PSE-A3 was used as an EMSA probe with pituitary GC nuclear extract. The scanning mutations were utilized as competitors at 25-fold and 50-fold mass excess of probe. Specific PSE-A3 complexes are indicated by arrowheads. Data reported in (130)

3.2.1d RFX1 binds to PSE-A3 in pituitary GC nuclear extracts.

To assess the ability of the RFX family to bind directly to PSE-A3, EMSA with GC nuclear extract and specific competitors was used (Fig. 3.9). The EF-C/MDBP 23 bp target site (161) was utilized as a representative RFX binding site. Use of EF-C/MDBP (RFX) as an EMSA competitor resulted in competition of complexes I and II, and to a lesser extent, complexes III and IV from the PSE-A3 EMSA probe (Fig. 3.9A). As a negative control, a mutation was made within the 23 bp EF-C/MDBP RFX element (RFXm), and this mutant was unable to compete for PSE-A3 specific complexes. Complexes I and II were also competed by the addition of RFX1-specific antibodies. The 'supershifted' aggregate that is observed in the well (closed arrowhead) is characteristic of the RFX1 antibody (191).

In contrast to the confirmation of RFX1 binding by supershift analysis, neither the RFX2 nor RFX3 antibodies were efficient competitors of PSE-A3 specific complexes (Fig. 3.9B). Additionally, they provided no evidence of supershifted bands. An attempt to supershift the 23 bp EF-C/MDBP RFX element with these antibodies also produced a negative result. It is noted, however, that competition of complexes I and II from the PSE-A3 probe by the RFX1-specific antibody appears to be complete. That is, there is no evidence of additional RFX family member association in place of the competed RFX1. It is possible that RFX2 and RFX3 are either absent from these cells, which seems unlikely given their ubiquitous expression pattern (155), or present, but in some way unable to bind to the oligonucleotide probes under these conditions. *Thus, competition of PSE-A3-specific complexes by the EF-C/MDBP site indicates that PSE-A3 is an RFX element. Confirmation that RFX1 was capable of binding directly to PSE-A3 came from the observation that an RFX1-specific antibody competed complexes I and II, with evidence of supershifted aggregates in the wells. Direct binding of either RFX2 or RFX3 to PSE-A3 was not detected.*

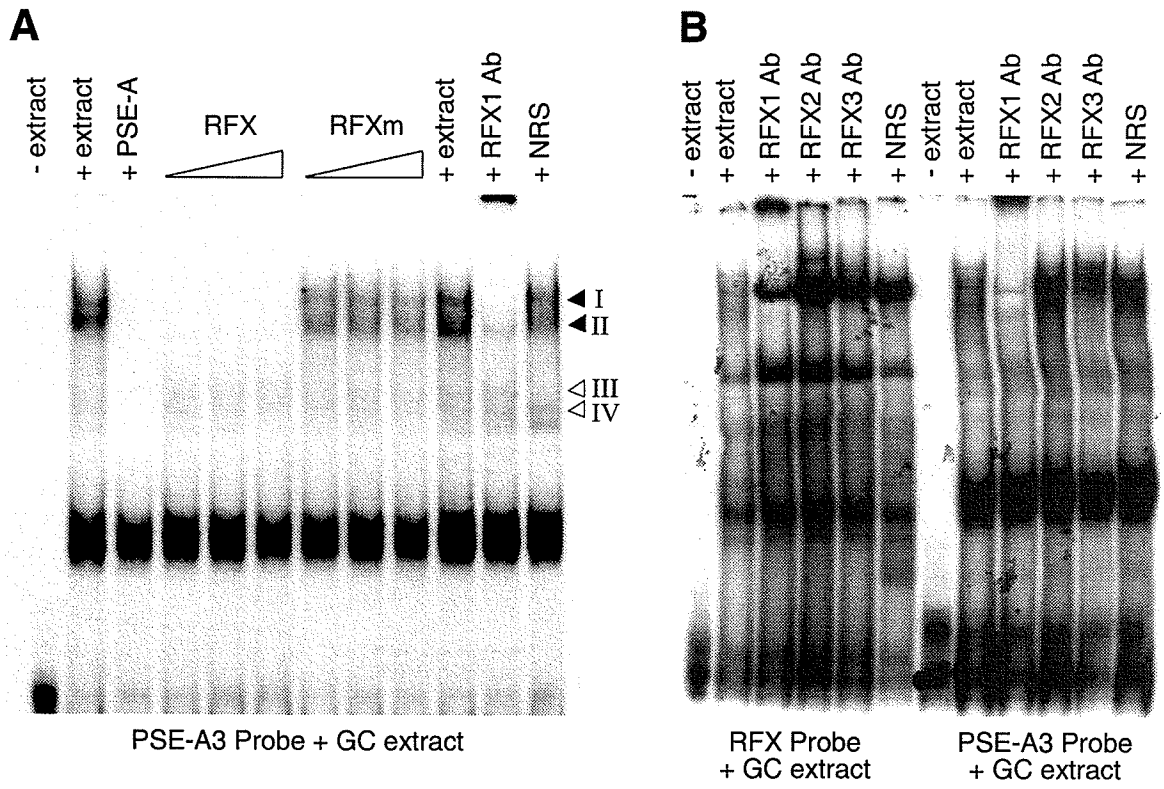


Figure 3.9: RFX1 from pituitary GC nuclear extract binds to PSE-A3.
(A) Radioloabelled PSE-A3 was used as an EMSA probe with pituitary GC nuclear extract. Competitor oligonucleotides were preincubated with nuclear extract at 100-fold mass excess of probe (PSE-A3) or 10-fold, 25-fold and 100-fold mass excess of probe (RFX and RFXm). Specific PSE-A3 complexes, defined as those that are competed by unlabelled PSE-A3, are indicated by arrowheads. The RFX complexes (I and II) are indicated by closed arrowheads. **(B)** Radioloabelled EF-C/MDBP RFX element or PSE-A3 were used as an EMSA probes with GC nuclear extract and antibodies as indicated. (Ab) antibodies (NRS) normal rabbit serum. Data reported in (130).

3.2.1e Direct binding of NF-1 to PSE-A3 was detected in the absence of RFX1 association.

An initial assessment of NF-1 binding to PSE-A3 was done using specific oligonucleotides and NF-1 antibodies as EMSA competitors from a PSE-A3 probe. When the PSE-A3 probe was incubated with pituitary GC cell nuclear extract, neither a commercially available NF-1 consensus oligonucleotide (Santa Cruz) nor the NF-1 antiserum were efficient competitors for specific PSE-A3 complexes (Fig. 3.10). *From this experiment, and the evidence of RFX1 binding (Fig. 3.9), it was clear that despite the*

presence of the putative NF-1 binding sites, RFX1 was the factor predominantly associating with the PSE-A3 oligonucleotide in GC nuclear extracts.

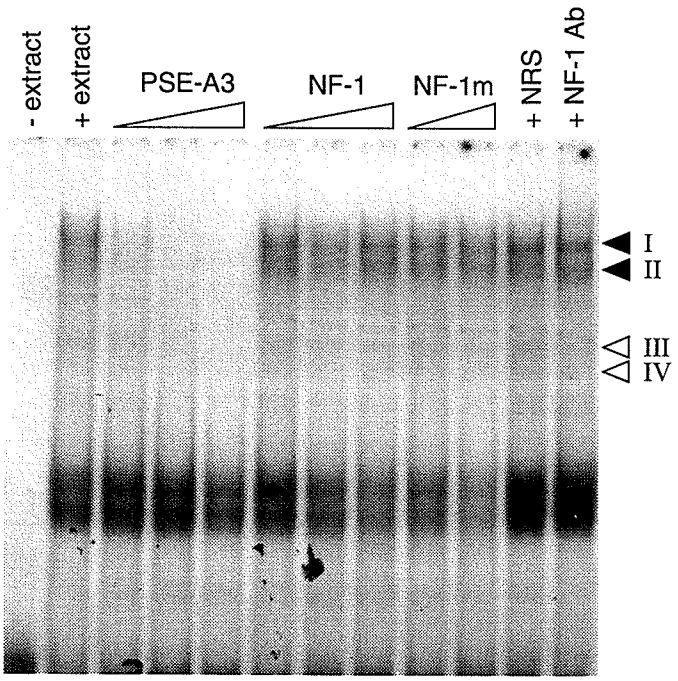


Figure 3.10: Direct binding of NF-1 was not detected with a PSE-A3 probe.
Radioloabelled PSE-A3 was used as an EMSA probe with pituitary GC nuclear extract. Competitor oligonucleotides were preincubated with nuclear extract at 5-fold, 25-fold and 50-fold mass excess of probe (PSE-A3 and NF-1) or 25-fold and 50-fold mass excess of probe (NF-1m). Specific PSE-A3 complexes, defined as those that are competed by unlabelled PSE-A3, are indicated by open arrowheads. Data reported in (151).

The existence of the putative NF-1 sites in PSE-A3, however, begged the question as to whether in the absence of RFX1 binding, NF-1 could associate with PSE-A3. To address this question two of the PSE-A3 scanning mutants, PSE-A3m2 and PSE-A3m6, were utilized as EMSA probes (Fig. 3.11). PSE-A3m2 contains a mutation in the PSE-A3 core region that disrupts the RFX binding site, but does not alter the overlapping NF-1 half-site (Figs. 3.5 and 3.8A). Evidence for disruption of the RFX site in PSE-A3m2 comes

from the fact that PSE-A3m2 was unable to compete for the RFX1 complexes (complexes I and II) from a PSE-A3 EMSA probe (Fig. 3.8B). The mutation in PSE-A3m6 corresponds to the 3´ NF-1 half-site, leaving the overlapping RFX/NF-1 regions intact (Figs. 3.5 and 3.8A). As one would predict, when PSE-A3m6 was used as an EMSA probe with GC nuclear extract, the pattern resembles that seen with the wild-type PSE-A3 probe, including the presence of RFX1 complexes I and II (Fig. 3.11B, closed arrowheads). No evidence of NF-1 association with PSE-A3m6 was observed, as neither the NF-1 consensus nor the NF-1 antibodies were efficient competitors of PSE-A3m6 complexes (Fig. 3.11B). In contrast, the PSE-A3m2 probe had a distinct EMSA pattern with GC nuclear extract, reflecting not only the absence of RFX complexes I and II, but the appearance of a novel complex (Fig. 3.11C, open arrowhead). This additional complex was competed by the NF-1 consensus oligonucleotide, but not a mutant NF-1 element. NF-1 binding to PSE-A3m2 was confirmed by competition of the PSE-A3m2 specific complex by NF-1 antiserum, resulting in a supershifted band of lower mobility (Fig. 3.11C, closed arrowhead). *The results of the experiments in Figure 3.11 suggest that PSE-A3 is a composite element capable of mutually exclusive RFX1 and NF-1 binding.*

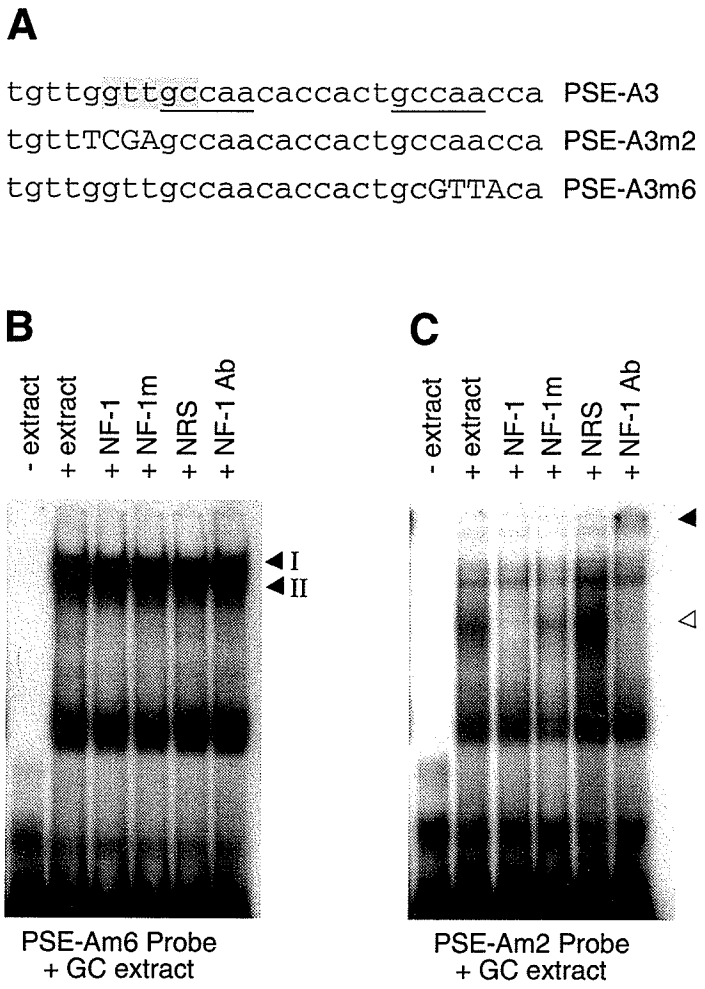


Figure 3.11: The direct binding of NF-1 to PSE-A was only detected in the absence of RFX1 association.

(A) The sequence of wild type PSE-A3 and the PSE-A3m2 and PSE-A3m6 mutations is shown. Mutated sequences are indicated by uppercase. The RFX1 half site is highlighted and the putative NF-1 half-sites are underlined. EMSA using pituitary GC nuclear extract and radiolabelled **(B)** PSE-Am6 and **(C)** PSE-Am2 probes. Competition was done with specific oligonucleotides at 50-fold mass excess of probe and NF-1 antibodies (Ab) as indicated. **(B)** RFX1 complexes I and II are indicated by open arrowheads. **(C)** NF-1 complexes and supershifted bands of lower mobility are indicated by open and closed arrowheads respectively. (NRS) normal rabbit serum. Data reported in (130).

3.2.1f PSE-A3 can repress the activity of the CS-A promoter in GC cells *in vitro*.

As with 263P, the original characterization of PSE-A demonstrated that this element repressed CS-A promoter activity in GC cells (133). PSE-A3 was inserted upstream of

CS-Ap in CSp.Luc (PSE-A3CSp.Luc) to assess functional activity in the absence of the CS-B downstream enhancer region. The luciferase activity of PSE-A3CSp.Luc was assessed in transiently transfected pituitary GC cells. The repression of CS-A promoter activity by PSE-A3 was variable (Table 3.2).

Table 3.2
PSE-ACSp.Luc does not consistently repress CS-Ap. Activity.

Experiment	CSp.Luc	263PCSp.Luc	PSE-ACSp.Luc	Significance
1	1.0 ± 0.24	0.38 ± 0.20	0.59 ± 0.08	p<0.05 (n=3)
2	1.0 ± 0.05	0.34 ± 0.04	0.75 ± 0.17	p<0.05 (n=3)
3	1.0 ± 0.10	0.52 ± 0.14	0.81 ± 0.06	p<0.05 (n=3)
4	1.0 ± 0.21	0.40 ± 0.08	0.77 ± 0.08	p<0.05 (n=6)
5	1.0 ± 0.06	0.43 ± 0.00	1.10 ± 0.14	p=0.12 (n=6)

The mean value and standard deviation (+/-) from each PSE-ACSp.Luc experiment is shown. 263PCSp.Luc was included as a control in each experiment. The p value is from an unpaired student t test comparison of PSE-ACSp.Luc and CSp.Luc.

In one of five separate experiments, no significant difference was observed between CSp.Luc and PSE-A3CSp.Luc activity (n=6), while significant repressor activity was observed in the remaining four experiments. This variability was significant when the means and standard deviations of the five experiments were compared using analysis of variance (ANOVA). Results from the four experiments where repressor function was observed were pooled, and the luciferase activity of PSE-A3CSp.Luc was 74% ± 1.1% of CSp.Luc 100% ± 1.5% ($p<0.0005$). *These experiments demonstrated that PSE-A3 can repress CS-A promoter activity in the absence of the CS-B downstream enhancer region in transiently transfected GC cells. The repressor function of PSE-A3, however, was variable, and the lack of repression in the one of five experiments was significant.*

The variability of the PSE-A3CSp.Luc construct may be understood in light of the ability to bind both RFX1 and NF-1. RFX1 is the predominant factor binding to the isolated PSE-A3 oligonucleotide, under the conditions used for EMSA. The context that the element is placed in, however, may affect the tendency to associate with one factor over

another. The 'context' could be the presence of additional sequences surrounding PSE-A3, as seen in 263P and PSE-A3CSp.Luc. Alternatively, the context of PSE-A3 could also be the state of the cells that nuclear extracts are made from, or transient transfections are performed in. That is, the cellular conditions may alter the balance, availability, and/or state of NF-1 and RFX1 proteins. Association of NF-1 or RFX1 with PSE-A3CSp.Luc could thus be affected and in turn result in an effect on functional activity of this construct.

3.2.1g RFX1 is a candidate PSE-A binding protein in placental as well as pituitary cells.

When placental JAR nuclear extract was used in place of pituitary GC nuclear extract with a PSE-A3 EMSA probe, a different pattern of interactions was observed. The sequences within PSE-A3 that were essential for the formation of this pattern were not determined as they were with pituitary extract, as none of the scanning mutations were efficient competitors of the PSE-A3 placental EMSA complex (Fig. 3.12).

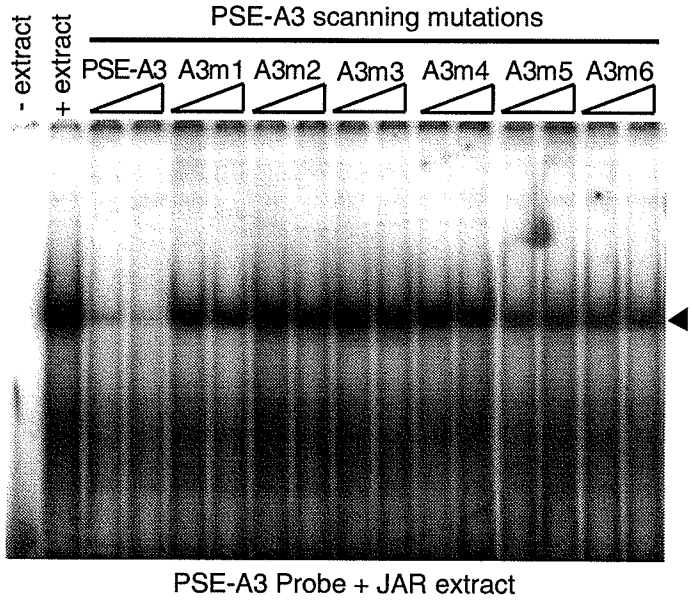


Figure 3.12: Scanning mutations do not reveal a core sequence in PSE-A3 for the formation of placental EMSA complexes.
EMSA with radiolabelled PSE-A3 probe and choriocarcinoma JAR nuclear extract. The specific complex, defined by competition with excess unlabelled PSE-A3, is indicated by

a closed arrowhead. Scanning mutations of PSE-A3 were used as competitors at 50-fold and 100-fold mass excess of probe.

As seen with pituitary GC nuclear extract, however, the binding of RFX1 to PSE-A3 was observed in both JAR and JEG-3 nuclear extracts (Fig. 3.13). The 23 bp EF-C/MDBP RFX element, but not the mutated RFX binding site were efficient competitors of the PSE-A3 pattern with placental JAR (Fig. 3.13A) and JEG-3 (Fig. 3.13B) nuclear extracts. Additional evidence of RFX1 binding was seen by the appearance of supershifted aggregates when RFX1 but not RFX2 and RFX3 antibodies were used. Unlike with the GC EMSAs, however, the RFX1 antibodies did not compete the RFX complex. This observation implies that in the absence of RFX1 binding, additional RFX family members may still associate with PSE-A3 in placental cells. RFX2 and RFX3 supershifts were also not observed with JAR extracts using an RFX probe as a positive control (data not shown). *Therefore, the lack of RFX2 and RFX3 supershifts with JAR nuclear extracts does not exclude the possibility that RFX2 or RFX3 may associate with PSE-A3. These experiments do, however, implicate RFX1 as a candidate PSE-A binding protein in placental cells.*

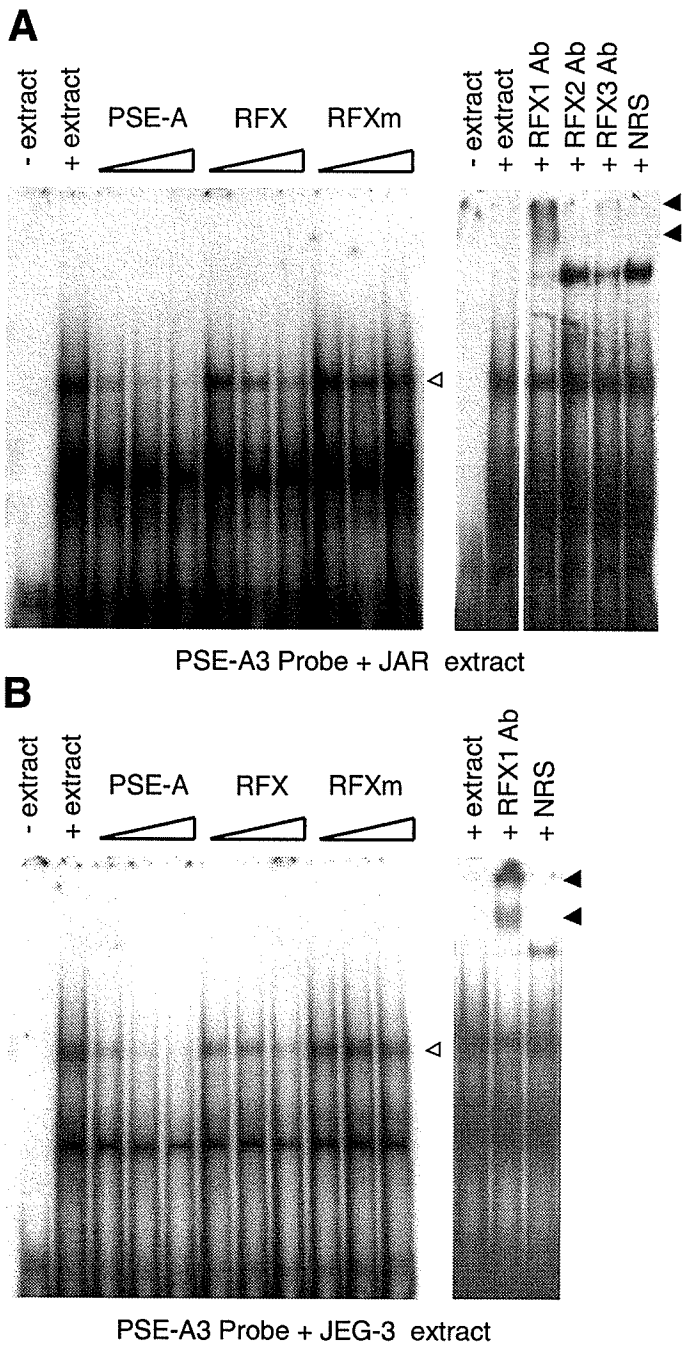


Figure 3.13: RFX1 associates with PSE-A3 in placental nuclear extracts.
Radioloabelled PSE-A3 was used as an EMSA probe with choriocarcinoma (**A**) JAR and (**B**) JEG-3 nuclear extracts. Competitor oligonucleotides were preincubated with extract at 25-fold, 50-fold, and 100-fold mass excess of probe. Specific complexes and supershifted bands of lower mobility are indicated by open and closed arrowheads respectively. (Ab) antibodies, (NRS) normal rabbit serum.

3.2.2 The PSE-B oligonucleotide is an NF-1 site.

3.2.2a NF-1 proteins can bind to PSE-B4 in pituitary GC nuclear extracts.

The PSE-B nuclease protected region was also analyzed with the MatInspector 2.2 program and TRANSFAC 4.0 database (165, 166). Matrix and core similarity parameters of 0.85 identified a single candidate PSE-B binding factor. A consensus NF-1 element (5´-TGGN$_6$GCCAAT-3´) without any variation in the two core binding regions (186, 187) was identified within the 27 bp PSE-B region (Fig. 3.14A). As was done with PSE-A, a truncated 18 bp oligonucleotide (PSE-B4) was utilized in place of the entire PSE-B region for further analysis of NF-1 as a candidate PSE-B binding protein.

To determine whether NF-1 binds directly to PSE-B4, EMSA competitions from a PSE-B4 probe were done using specific oligonucleotides and NF-1 antibodies. The PSE-B4 probe resulted in specific complexes with GC cell nuclear protein (Fig. 3.14B, open arrowheads) defined through their competition by unlabelled PSE-B4 (2.5-50-fold mass excess of probe). The NF-1 consensus, but not the mutant NF-1 oligonucleotide, was an efficient competitor of PSE-B4 specific complexes at as low as 2.5-fold mass excess of PSE-B4 probe. *The competition of PSE-B4-specific complexes by NF-1 antibodies, and the appearance of a supershifted band of lower mobility (closed arrowhead) confirmed direct binding of NF-1 to PSE-B4. In terms of relative affinity, it appears that PSE-B4 is a lower affinity NF-1 site than the NF-1 consensus, as the NF-1 site is a more efficient competitor of PSE-B4 specific complexes than PSE-B4 is itself.*

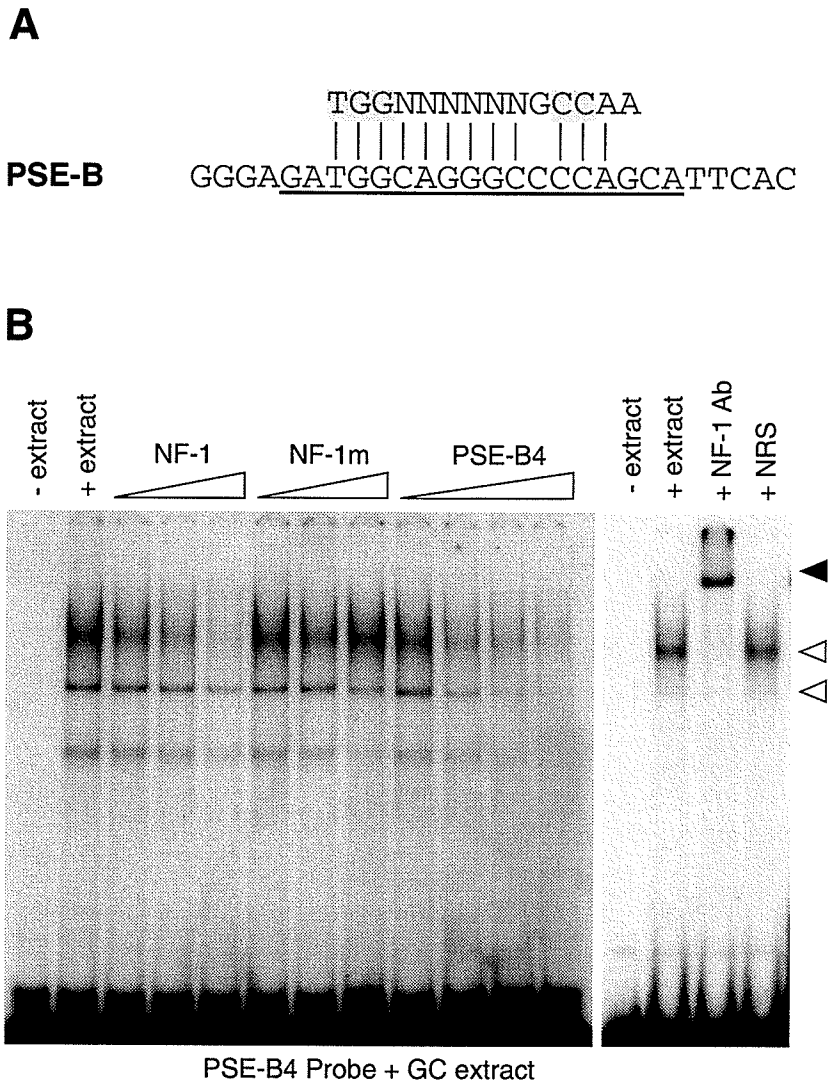


Figure 3.14: Direct binding of NF-1 from pituitary GC nuclear extract to PSE-B4. **(A)** The sequence of the PSE-B nuclease protection region is shown with the PSE-B4 oligonucleotide underlined. Aligned above the PSE-B sequence is the consensus binding site for NF-1. Core regions within the NF-1 consensus sequence, determined to be essential for the association of NF-1 (186, 187), are highlighted. **(B)** Radioloabelled PSE-B4 was used as an EMSA probe with pituitary GC nuclear extract. Competitor oligonucleotides were preincubated with extract at 2.5-fold, 5-fold, and 25-fold mass excess of probe (NF-1 and NF-1m) or 2.5-fold, 5-fold, 25-fold and 50-fold mass excess of probe (PSE-B4). The PSE-B4 specific complexes and supershifted bands are indicated by open and filled arrowheads, respectively. (Ab) antibodies, (NRS) normal rabbit serum. Data reported in (151).

3.2.2b NF-1 binding is responsible for PSE-B4 repressor activity.

The retention of repressor activity by PSE-B4 was assessed by inserting PSE-B4 upstream of the CS-A promoter in CSp.Luc (PSE-B4CSp.Luc) and determining the luciferase activity of this construct relative to CSp.Luc in transiently transfected GC cells. The presence of PSE-B4 resulted in a consistent and significant repression of CS-Ap activity. In seven separate experiments, the average activity of PSE-B4CSp.Luc was 28% of CSp.Luc luciferase values (range of 19-39%). A representative PSE-B4 experiment can be seen in Figure 3.15, where PSE-B4CSp.Luc has 28% of CSp.Luc activity (n=6, $p<0.05$). *These experiments demonstrated that, like the 263P fragment, PSE-B4 is a repressor of CS-A promoter activity in the absence of the CS-B downstream enhancer region in transiently transfected GC cells.*

Transient transfection experiments with luciferase reporter gene constructs were also used to assess the role of NF-1 in PSE-B4 repressor activity. An initial experiment examined the ability of a known NF-1 binding site to functionally substitute for PSE-B4. The NF-1 consensus element was inserted upstream of CS-Ap in CSp.Luc to create NF-1CSp.Luc. *Like PSE-B4CSp.Luc, NF-1CSp.Luc had significantly less luciferase activity than CSp.Luc in GC cells (Fig. 3.15, 62%, n=12, $p<0.05$), demonstrating the ability of a known NF-1 element to function as a repressor of CS-A promoter activity in GC cells.*

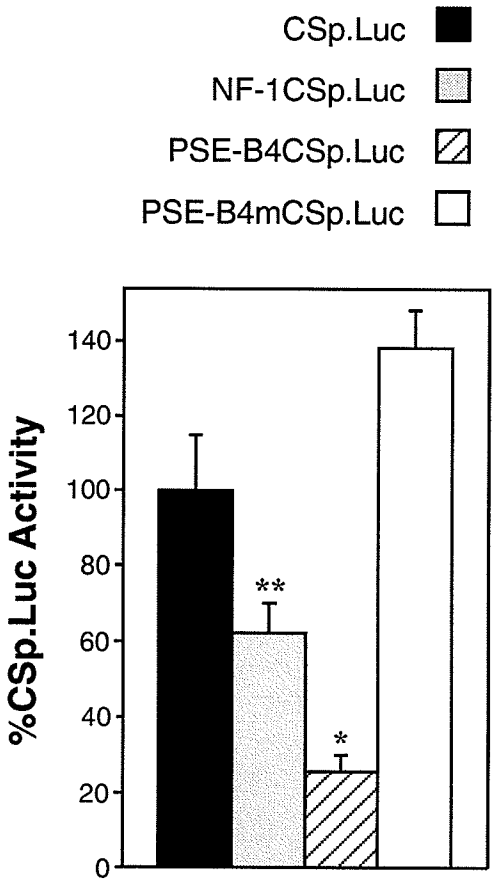


Figure 3.15: NF-1 is responsible for PSE-B repressor activity.
Hybrid luciferase (Luc) genes were used to assess the role of NF-1 in the repressor activity of PSE-B4. Constructs were transiently transfected into pituitary GC cells, and to control for DNA uptake cells were co-transfected with RSVp.CAT. Corrected values are expressed as a percentage of CSp.Luc activity, which was arbitrarily set to 100%. Bars represent SEM. *$p<0.05$. **$p<0.005$. Data reported in (151).

A second set of experiments was designed to test if the binding of NF-1 to PSE-B4 was responsible for CS-A promoter repression. Base pair substitutions were made within PSE-B4 in the core of the NF-1 recognition sequence, creating PSE-B4m (Fig. 3.16A). This mutation was then assessed for the loss of NF-1 binding by EMSA. When PSE-B4m was used as an EMSA probe with GC nuclear extract, the pattern was noticeably different from that of the PSE-B4 probe, specifically, the NF-1 complexes that formed on PSE-B4 were no longer observed (Fig. 3.16B). The PSE-B4m specific complexes, defined as those competed by excess unlabeled PSE-B4m, were not competed by PSE-B4, the

consensus NF-1 oligonucleotide, or the NF-1 antibodies. The appearance of the NF-1 supershifted band that was seen with the PSE-B4 probe, was also no longer observed with the PSE-B4m probe. *These results confirmed that, as predicted by the designed mutation, NF-1 does not bind to PSE-B4m.* In addition, PSE-B4m was not an efficient competitor of PSE-B4 complexes, even at 1000-fold mass excess of probe (Fig. 3.16C).

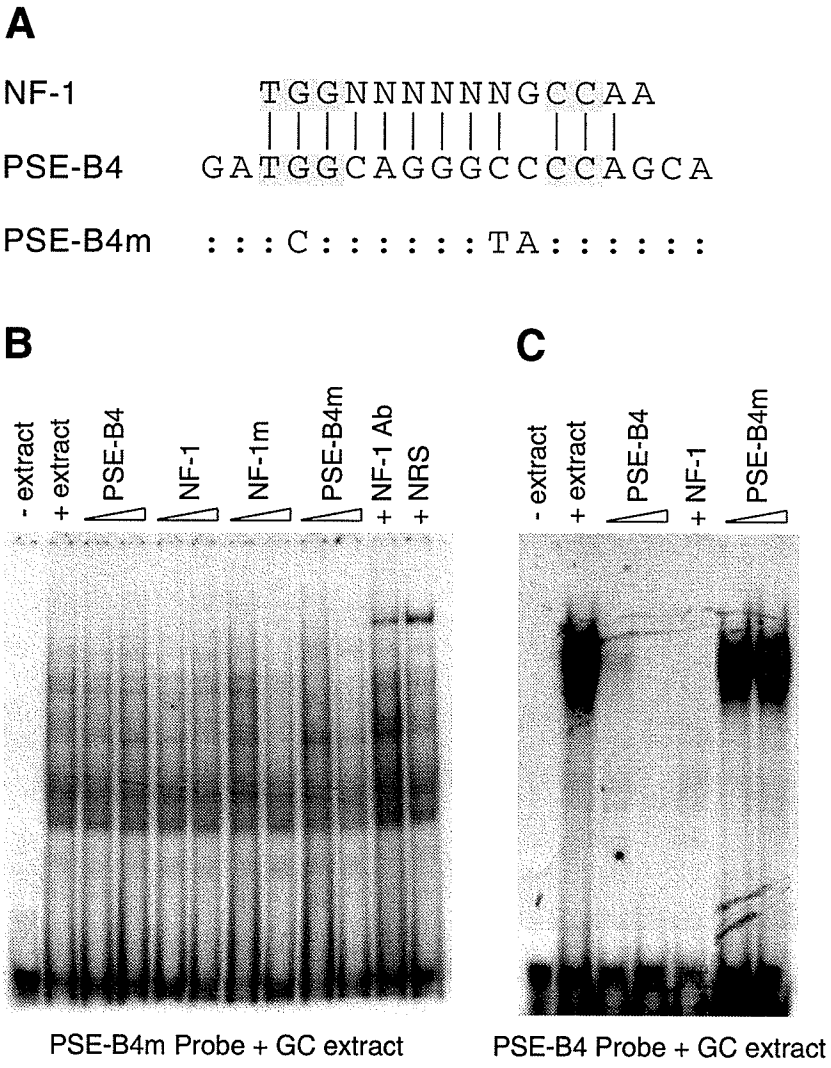


Figure 3.16: PSE-B4m does not bind NF-1.
(A) Sequence alignment of NF-1 consensus, PSE-B4 and PSE-B4m. The NF-1 core region is highlighted. **(B)** Radiolabelled PSE-B4m was used as an EMSA probe with GC nuclear extract. Competitor oligonucleotides were preincubated with extract at 5-fold and 25-fold mass excess of probe. (NRS) normal rabbit serum. **(C)** Radiolabelled PSE-B4 was used as an EMSA probe with GC nuclear extract. Competitor oligonucleotides were preincubated with extract at 10-fold (NF-1) or 100-fold and 1000-fold mass excess of probe (PSE-B4 and PSE-B4m).

When PSE-B4m was placed upstream of the CS-A promoter (PSE-B4mCSp.Luc) and compared to PSE-B4CSp.Luc, the ability to repress CS-A promoter activity was no

longer detected (Fig. 3.15, n=9). In fact, a slight 1.3-fold increase, although not significant (p=0.06), was observed. *Taken together, these data indicated that the repressor activity of PSE-B4 is due to the direct binding of NF-1 to this element.*

3.2.2c Minor variations outside of the NF-1 core binding region in PSE-B4 affect functional activity.

The repressor activity of PSE-B4 can be accounted for by the direct binding of NF-1 (Fig. 3.15 and 3.16). There is a significant difference, however, between the 72% repression of CS-A promoter activity with PSE-B4CSp.Luc and the 38% repression with NF-1CSp.Luc (Fig. 3.15, n=12, $p<0.005$). NF-1 is not a single transcription factor. There are four genes in vertebrates that comprise this family; NF-1A, NF-1B, NF-1C (also known as CTF-1), and NF-1X (192-194). The NF-1 family all recognize the same DNA consensus sequence, due to >90% homology in an NH_2-terminal DNA binding domain (153, 170, Gounari, 1990 #107, 181, 195). Outside of this conservation, the remainder of the NF-1 protein varies greatly between the different members (181, 195). The variability of the family beyond the DNA-binding domain likely explains why roles in both transcriptional repression (182, 196-199) and activation (153, 192, 200-202) have been reported for NF-1. PSE-B4 is clearly an NF-1 binding site, however, as with other known NF-1 regulatory elements, it is unclear which NF-1 family member(s) is responsible for repression of the CS-A promoter in GC cells. In the literature, there is some evidence that although all members of the NF-1 family recognize and associate with the same consensus sequence, their individual affinities for different sites can vary (186, 187). For example, all members of the NF-1 family bind to the adenovirus origin of replication site with higher affinity than to the GST-P gene NF-1 site (170). If, however, you examine only the adenovirus origin of replication site, the NF-1 family member with the greatest affinity is NF-1A followed by NF-1X, NF-1B, and NF-1C (170). In contrast, the NF-1 site in the GST-P gene has a different profile of relative affinities, where NF-1X has the highest affinity, followed by NF-1A, NF-1B, and then NF-1C (170). Since the COOH-terminal region varies so extensively between the different family members, one can imagine that although all NF-1s are capable of binding to a particular site, the family member with the greatest relative affinity for that site would be the predominant associating factor and thus

the major determinant of the activity (for example activation versus repression) exerted by that element. Since the relative expression of all family members should be the same in the GC cells transiently transfected with PSE-B4CSp.Luc and NF-1CSp.Luc, the difference in their capacity to repress the CS-A promoter may lie in different NF-1 family member affinity profiles.

Although PSE-B4 contains an intact NF-1 core region, there are minor sequence variations that exist when aligned with the NF-1 consensus sequence used in NF-1CSp.Luc (Fig. 3.17A). These variations occur in both the NF-1 recognition site as well as the six base pair spacer region. Minor sequence variations between NF-1 elements have been shown to affect the relative affinity profile for different members of the NF-1 family (170, 186, 187), however, a functional consequence of this had never been examined. The question arose as to whether the sequence variations between the NF-1 consensus and PSE-B4 elements could be responsible for their functional differences. To examine this possibility, mutations in PSE-B4 were created that should not affect the ability of NF-1 to bind, but may have an impact on the relative affinities of different family members for the sites (Fig. 3.17A). The NF-1 consensus includes a 5′-GCCAA-3′ half-site, of which the dinucleotide "CC" is essential for NF-1 binding. The half-site in PSE-B4 is 5′-CCCAA-3′. Thus, a single base pair conversion was introduced in PSE-B4 to change the cytosine (C) of this sequence to the guanine (G) found in the NF-1 consensus element, creating PSE-B4Gm. A change in the "spacer" region of the NF-1 site (or N) was also made within PSE-B4, converting adenosine (A) to cytosine (C), and this was named PSE-B4Cm. In addition, a double mutation of both changes (PSE-B4CGm) was made.

The PSE-B4 mutations were first assessed by EMSA to ensure that they had retained the ability to bind NF-1 proteins (Fig. 3.17B). The NF-1 complexes of PSE-B4 were retained by all three mutants when they were used as EMSA probes with pituitary GC nuclear extract. These complexes were competed by both NF-1 consensus oligonucleotide and NF-1 antibodies, but not the NF-1 mutant oligonucleotide (Fig. 3.17B, open arrowhead). *The retention of NF-1 binding to the PSE-B4 mutants was confirmed by the observation*

of a supershifted band when NF-1 antibodies were included in the reaction (Fig. 3.17B, closed arrowhead).

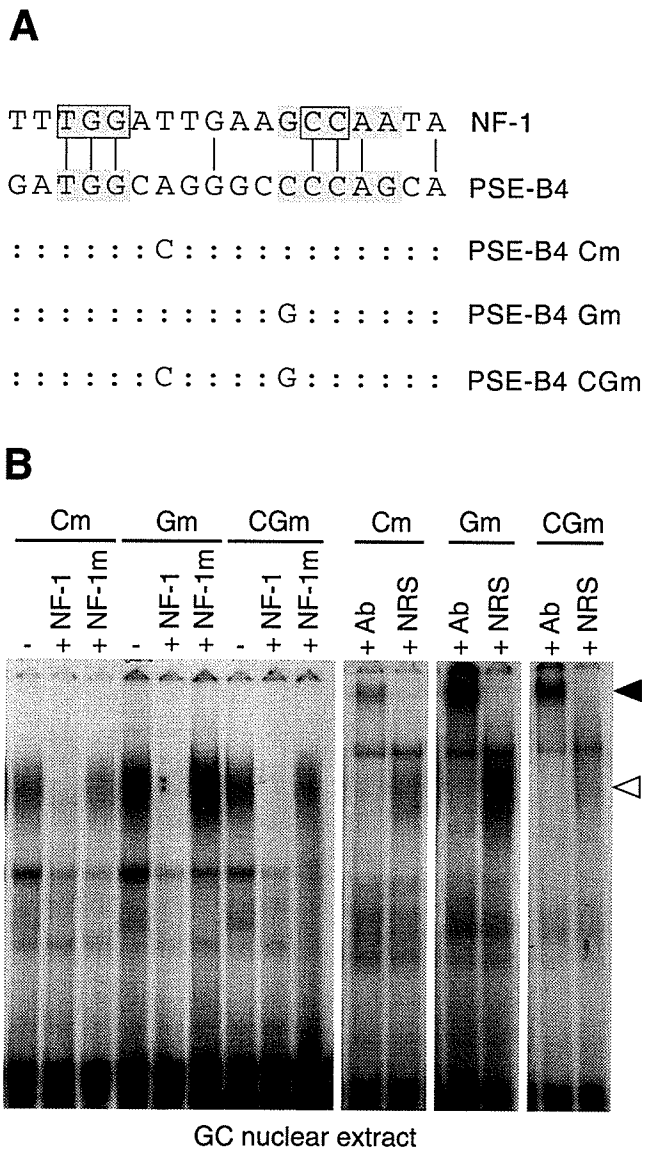


Figure 3.17: Minor variations outside of the NF-1 core binding region in PSE-B4 do not disrupt NF-1 binding.

(A) Sequence alignment of NF-1 consensus, PSE-B4 and the PSE-B4 mutants. The NF-1 recognition sequence is shaded and core region boxed. **(B)** The PSE-B4 mutants were radiolabelled and used as EMSA probes with pituitary GC nuclear extract. The probe used in each reaction is indicated at the top with competitors listed below. Competitor oligonucleotides were preincubated with extract at 25-fold mass excess of probe. (Ab) NF-1 antibodies, (NRS) normal rabbit serum. Specific complexes and supershifted bands are indicated by open and closed arrowheads respectively. Data reported in (151).

Once the ability of the PSE-B4 mutations to bind NF-1 was confirmed, they were placed upstream of the CS-A promoter and their functional effects were evaluated in transiently transfected GC cells (Fig. 3.18). All three mutations were significant repressors of CS-A promoter activity. The presence of PSE-B4Gm resulted in luciferase activity that was 77% of CSp.Luc (arbitrarily set to 100%, n=6, $p<0.05$) while both PSE-B4CmCSp.Luc and PSE-B4CGmCSp.Luc displayed 61% of CSp.Luc luciferase activity (n=6, $P<0.005$). The PSE-B4CSp.Luc and NF-1CSp.Luc contsructs were included in these experiments for comparison. The mutation in PSE-B4Gm resulted in a significantly different effect on CS-Ap activity compared to both PSE-B4 and NF-1, whereas PSE-B4CmCSp.Luc and PSE-B4CGmCSp.Luc were significantly different from PSE-B4, but not NF-1. *These results demonstrated for the first time that variations in an NF-1 recognition site, including the 'spacer' region, could affect the functional activity of individual NF-1 elements*. Underlying these results is that the degree of CS-A promoter repression by PSE-B4 relies on its specific NF-1 binding sequence, which may have a unique affinity profile for different members of the NF-1 family.

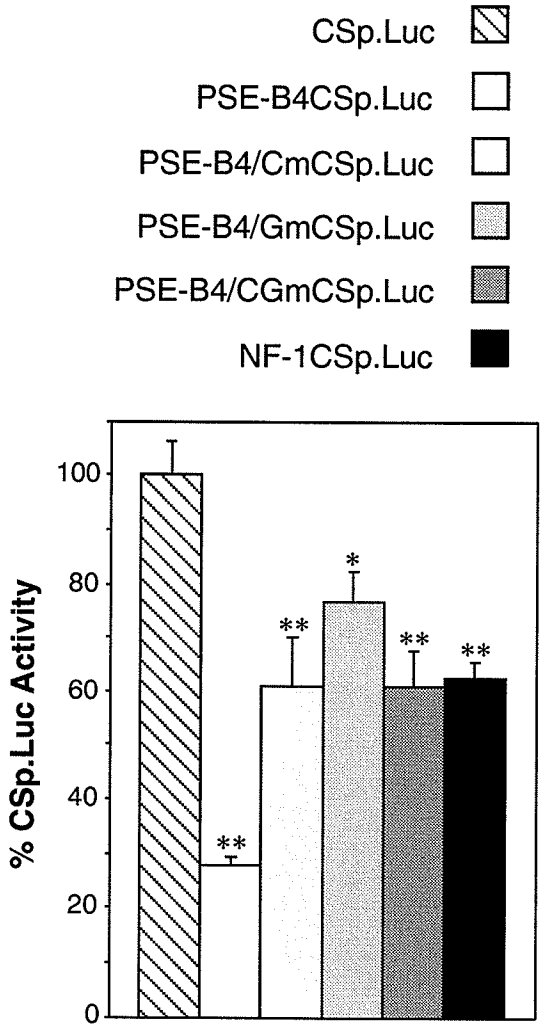


Figure 3.18: Minor variations outside of the NF-1 core binding region in PSE-B4 affect functional activity.
Hybrid luciferase (Luc) genes were used to assess the effect of point mutations on PSE-B4 function. Constructs were transiently transfected into GC cells, and to control for DNA uptake cells were co-transfected with RSVp.CAT. Corrected values are expressed as a percentage of CSp.Luc activity, which was arbitrarily set to 100%. Bars represent SEM. *p<0.05, **p<0.005. Data reported in (151).

3.2.2d Detection of NF-1A, NF-1C and NF-1X transcripts in rat pituitary GC cells.

Although NF-1 expression was previously demonstrated in pituitary cells(182), it was not determined which of the NF-1 family members were expressed. The expression of the different NF-1 family members was assessed in rat pituitary GC cells by northern

blotting. Due to the high homology in the NH_2-terminal DNA-binding domain, specific probes for each NF-1 were designed and tested for cross-reactivity. The rat NF-1A, NF-1B, NF-1C and NF-1X cDNAs were a generous gift from Dr. M. Imagawa (170). The cDNAs were restriction endonuclease digested and COOH-terminal regions were isolated. Slot blots containing 0.1 and 1.0 ng of the NF-1 cDNA plasmids were probed with each of the carboxyl-terminal probes (Fig. 3.19A). Specific binding to their respective cDNA, with very little evidence of cross-reactivity to other family members was observed. These same probes were then used to assess the presence of the different NF-1 family members in GC cells (Fig. 3.19B). For the northerns, 25 μg of polyA$^+$-enriched GC cell RNA was electrophoresed and transferred to nitrocellulose. *Using the specific probes, the expression of NF-1A, NF-1C and NF-1X was confirmed in GC cells (Fig. 3.19B).* Several transcripts were detected; NF-1A (5.1 and 4.4 kb), NF-1C (6.0 and 4.0 kb) and NF-1X (5.7 and 4.6 kb). NF-1B was undetectable in the GC RNA samples under these conditions. Previously, NF-1B was shown to be expressed in rat lung (170), and the ability of the NF-1B probe to detect NF-1B transcripts was confirmed by northern blotting with 30 μg of polyA$^+$-enriched mouse lung RNA (data not shown). The original film of the NF-1B lung RNA northern blot was forwarded to the editors of Biochemical Journal when Norquay *et al* (2001) was submitted for review. From these results, it appears that if NF-1B is expressed in GC cells, the levels are much lower than NF-1A, NF-1C, or NF-1X. Alternatively an NF-1B splice variant that is undetectable by this probe may be expressed in GC cells. An alternatively spliced NF-1B transcript has been described elsewhere that is truncated to the DNA-binding domain (203). This transcript does not contain sequences corresponding to the NF-1B specific probe designed here. *These experiments demonstrate that NF-A, NF-1C, and NF-1X are PSF-B candidates in pituitary GC cells. Although NF-1B transcripts were not detected, their presence cannot be ruled out, as there is the potential for expression of an alternative transcript that would not be detected by the NF-1B probe.*

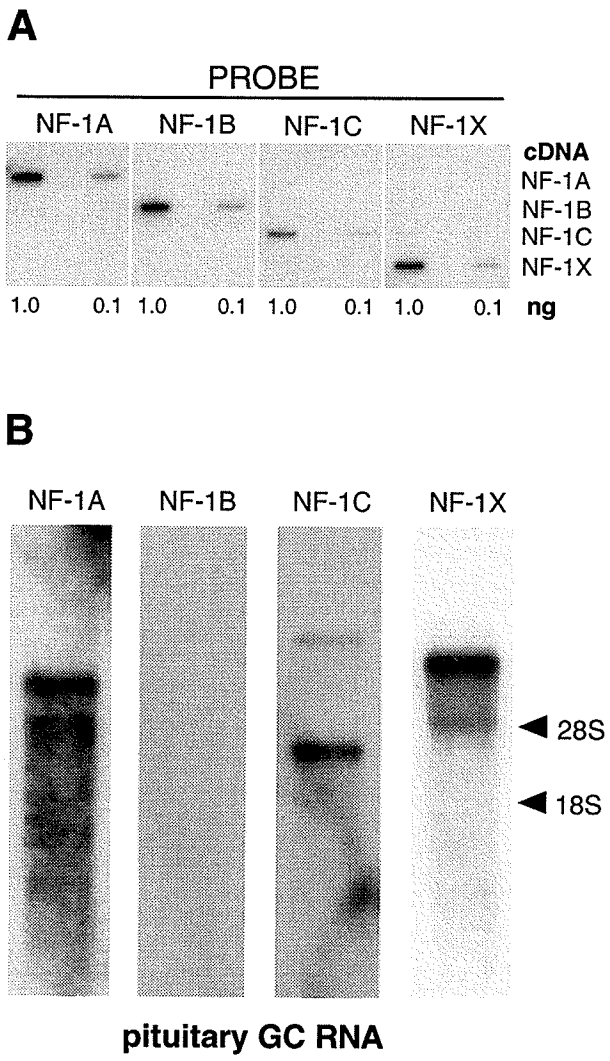


Figure 3.19: NF-1 transcripts were detected in pituitary GC cells by RNA blotting. **(A)** Specific NF-1 cDNA probes for NF-1A, NF-1B, NF-1C and NF-1X were created to the carboxyl-terminal regions, as they vary between family members. To test the specificity, samples (0.1 and 1.0 ng) of the NF-1 cDNA plasmids were immobilized on nitrocellulose with a slot-blot apparatus and probed with each of the radiolabelled carboxyl-terminal cDNA probes. **(B)** Northern blots of polyA$^+$-enriched RNA from pituitary GC cells (25 μg per lane). Each blot was probed with one of the specific NF-1 probes as indicated. The locations of 28S and 18S RNA correspond to 4.7 and 1.9 kb respectively, and were used to estimate the sizes of the NF-1 transcripts. Data reported in (151).

3.2.2e The NF-1 family has the capacity to interact with RFX1.

As previously discussed, the ability of both PSE-A3 and PSE-B4 to compete for the 263P nuclease protection pattern suggested that if the factors associating with these elements were different, they may form interactions with each other (Fig. 3.3). This potential was further assessed through EMSA. At high levels of competitor (1000-fold mass excess of probe), the PSE-B4 and NF-1 oligonucleotides competed for RFX1 complexes from a PSE-A3 probe (Fig. 3.20). *These low affinity competitions were a further indication that the PSFs had the capacity to interact.*

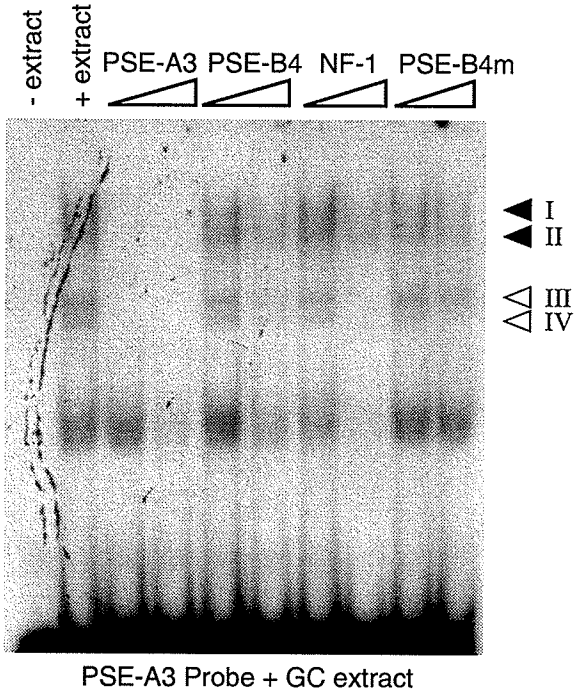


Figure 3.20: PSE-B4 can compete for RFX complexes from a PSE-A3 probe.
Radiolabelled PSE-A3 was used as an EMSA probe with pituitary GC nuclear extract. Competitor oligonucleotides were preincubated with nuclear extract at 100-fold and 1000-fold mass excess of probe. Specific PSE-A3 complexes are indicated by arrowheads. The RFX1 complexes are indicated by closed arrowheads.

The potential for RFX1 and the NF-1 family to interact was subsequently investigated through co-immunoprecipitation experiments. Specific antibodies for each of the NF-1

family members were not available. Therefore, to differentiate between the different NF-1 species, epitope tags were utilized. The NF-1 cDNAs for each family member were inserted downstream of a 39 bp *c-myc* epitope sequence in mammalian expression vectors (Clontech). These constructs were transiently co-transfected with a similarly constructed RFX1 expression vector (HA-epitope tag) in human choriocarcinoma (JAR) cells. A collaboration with Ms. Patricia Sheppard and Dr. Janice Dodd was arranged, due to their expertise in immunoprecipitation and protein blotting techniques. Protein blots of the transfected JAR lysates were probed with c-myc and RFX1 antibodies to ensure that the NF-1 family members and RFX1 were overexpressed in the transfected JAR cells (Fig. 3.21A and 3.21B). Immunoprecipitations were then carried out with the transfected JAR lysates and RFX1, or as a negative control, unrelated rodent Prostate dorsolateral protein (DLP)-specific antibodies. Probing the immunoprecipitates with the RFX1 antibody confirmed the specificity of the protocol (Fig. 3.21C). When the RFX1 antibody was used to immunoprecipitate the lysates, evidence of NF-1 co-immunoprecipitation was seen by bands that corresponded in size to the tagged NF-1 family members when the western blots were probed with c-myc antibodies (Fig. 3.21D). These same bands were not detected in the negative control DLP immunoprecipitation (Fig. 3.21E). *These results demonstrated the capacity for the NF-1 family to interact with RFX1. All members of the NF-1 family appear to be capable of these interactions. This suggests that the NF-1/RFX1 interaction may occur through a domain that is common to the NF-1 family, such as the NH_2-terminal DNA-binding domain.*

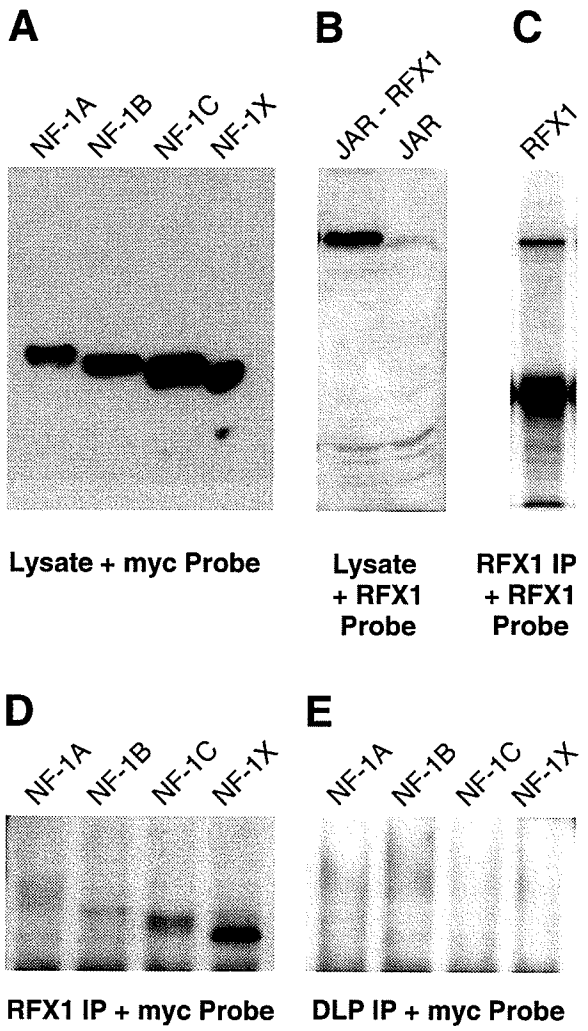


Figure 3.21: The NF-1 family has the capacity to interact with RFX1.
(A) Transiently transfected JAR lysates (50 μg) were separated by SDS-PAGE and transferred protein blots were probed with c-myc antibody to confirm overexpression of the epitope-tagged NF-1 family members. The approximate sizes of the NF-1 proteins were 60kDa (NF-1A), 57 kDa (NF-1B), 54 kDa (NF-1C), and 50 kDa (NF-1X). **(B)** Transiently transfected JAR lysate (50 μg) was separated by SDS-PAGE and the transferred protein blot was probed with RFX1 specific antibody to confirm overexpression of RFX1. **(C)** The specificity of the RFX1 immunoprecipitation was confirmed by probing RFX1 immunoprecipitate with the RFX1 antibody. **(D)** The presence of NF-1 family members in the RFX1 immunoprecipitations was revealed by probing with c-myc antibody. **(E)** Immunoprecipitation of the transiently transfected JAR lysates with unrelated DLP antibodies was done as a negative control. Probing of the DLP immunoprecipitates with c-myc antibody did not detect NF-1 family members. Immunoprecipitations and western data provided by Ms. Patricia Sheppard. Data reported in (130).

Co-immunoprecipitation experiments were also used to demonstrate the ability of endogenous NF-1 and RFX1 proteins to interact (Fig. 3.22). Immunoprecipitations with commercial NF-1 and RFX1 antibodies (Santa Cruz) were done from pituitary GC nuclear extract. The immunoprecipitates were run in SDS-PAGE gels, and transferred protein blots were probed with RFX1 (Fig. 3.22A) and NF-1 (Fig. 3.22B) antibodies. The specificity of the immunoprecipitations was confirmed, as NF-1 and RFX1 were both visible when their respective immunoprecipitates were probed with the same antibody. When RFX1 was used to immunoprecipitate the GC nuclear extract, a band of the appropriate size was visible when the blot was probed with NF-1 antibody (Fig. 3.22B). *This result demonstrates that the RFX1 and NF-1 proteins in GC nuclear extracts are capable of participating in common complexes.* The reverse, however, was not observed, as a band that corresponded to RFX1 was not visible when the NF-1 immunoprecipitation was probed with RFX1 antibody (Fig. 3.22A). The NF-1 antibody specifically recognizes the NF-1 DNA binding domain. If in fact the interpretation of the epitope-tagged immunoprecipitation data (Fig. 3.21) is true, and RFX1 interacts with a common NF-1 domain such as the DNA-binding domain, it is possible that this explains the result. In this interpretation, the NF-1 and RFX1 that are interacting in the GC nuclear extract are doing so through the NF-1 DNA binding domain. The possibly exists that this interaction could mask the NF-1 epitope that the antibody detects. The NF-1 that is immunoprecipitated would be the remainder of the NF-1 population, that is not associated with RFX1. *Thus, although these data do not prove that NF-1 and RFX1 interact through the DNA binding domain of the NF-1 family, taken together, the results from Figures 3.21 and 3.22 are consistent with this interpretation.*

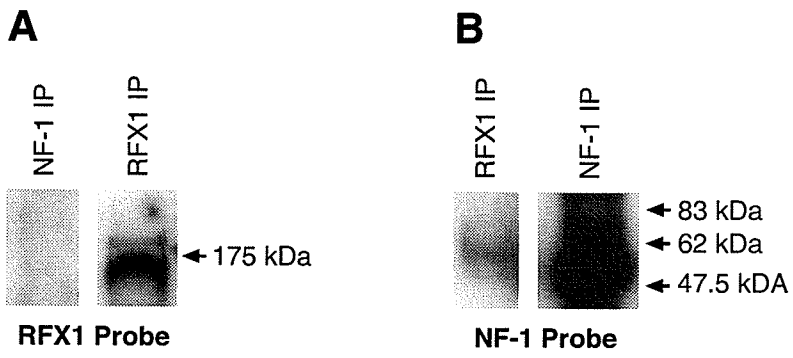
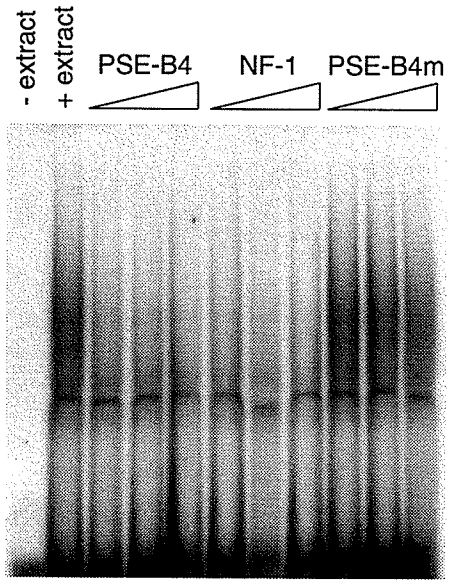


Figure 3.22: Interactions between endogenous NF-1 and RFX1 proteins are detected through co-immunoprecipitations of pituitary GC nuclear extract.
(A) RFX1 and NF-1 immunoprecipitations (IP) were separated by SDS-PAGE and transferred protein blots were probed for RFX1. The specificity of the RFX1 immunoprecipitation was confirmed by probing RFX1 immunoprecipitate with the RFX1 antibody. **(B)** RFX1 and NF-1 immunoprecipitations (IP) were separated by SDS-PAGE and transferred protein blots were probed for NF-1. The specificity of the NF-1 immunoprecipitation was confirmed by probing NF-1 immunoprecipitate with the NF-1 antibody. These blots were provided by Ms. Xiaoyang Yang.

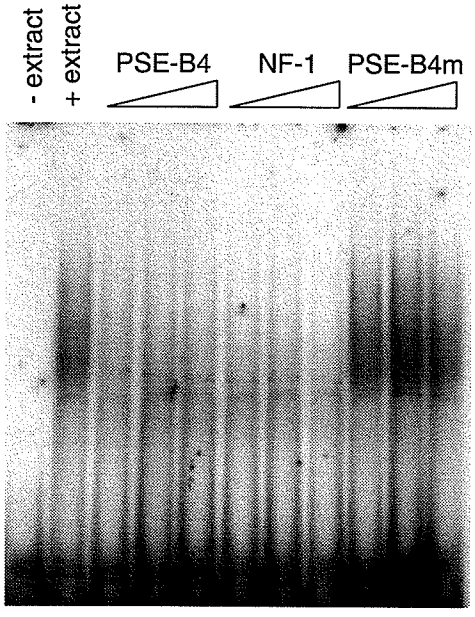
3.2.2f Direct binding of NF-1 to PSE-B4 in placental nuclear extracts.

As well as the ability to repress the CS-A promoter in pituitary cell lines (133), P sequences have the capacity to enhance reporter gene expression in transgenic mouse placenta (150). This suggests that if this single element has the capacity to activate gene expression in one cell type and repress gene expression in another, then a pituitary/repressor complex may exist that is different from a placental/enhancer complex. The observation of NF-1 association with PSE-B4 is intriguing in this respect, as this family has been implicated in both repression (182, 196-199) and activation (153, 192, 200-202). The question arose, therefore, as to whether the NF-1 family could associate with PSE-B4 in placental as well as pituitary cells. Radiolabelled PSE-B4 probe was incubated with nuclear extracts from human term placenta tissue and choriocarcinoma (JAR and JEG-3) cell lines in EMSAs (Fig. 3.23). Although the complexes were not as clear as those seen with pituitary GC nuclear extract, evidence of competition by the NF-1 consensus sequence was observed with all three nuclear extracts. As a negative control, the PSE-B4m oligonucleotide, which does not bind NF-1

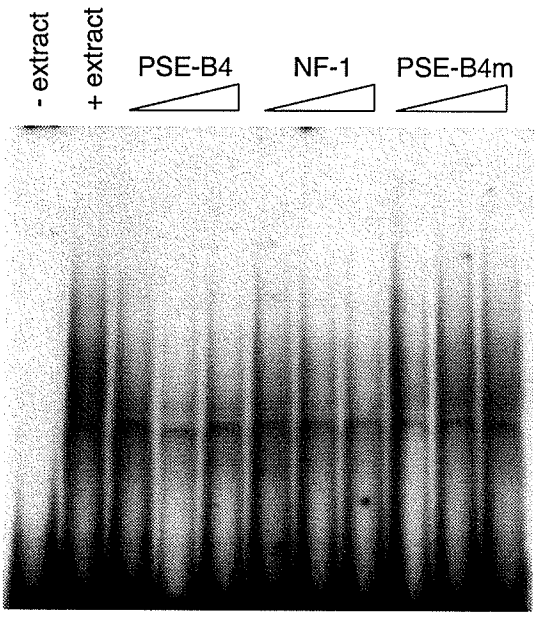
(Fig. 3.16), was used as a competitor and did not result in competition of the PSE-B4 placental complexes. *These results provide evidence that the NF-1 family members are PSF-B candidates in the placenta as well as the pituitary.*



PSE-B4 Probe + human placenta extract



PSE-B4 Probe + JAR extract



PSE-B4 Probe + JEG-3 extract

Figure 3.23: NF-1 from placental nuclear extracts associates with PSE-B4.
Radioloabelled PSE-B4 was used as an EMSA probe with human term placenta, JAR, and JEG-3 nuclear extracts as indicated. Competitor oligonucleotides were preincubated with extract at 25-fold, 50-fold, and 100-fold mass excess of probe.

3.3 The P sequence complex in pituitary GC cells.

The identification of candidate PSFs using the individual PSEs demonstrated the potential for both the NF-1 family and RFX1 to be participants in the P sequence complex *in vitro*, as well as the capacity for the NF-1 family and RFX1 to interact. The experiments presented to this point also identified PSE-A as a composite element, and the binding of RFX1 and NF-1 to PSE-A3 as mutually exclusive. As previously discussed, the demonstration of mutually exclusive NF-1 and RFX1 binding to PSE-A has important implications in the interpretation of PSE-A results. That is, the context of surrounding sequences, in combination with cellular environment, may influence which of these two factors will associate with the PSE-A region. The observation that PSE-B is an NF-1 binding site raised the possibility that the factor(s) associating with PSE-A in the context of 263P may be influenced by the binding of NF-1 to PSE-B. Thus, DNA fragments that contain both PSE-A and PSE-B were evaluated for the participation of the NF-1 family and RFX1 in the *in vitro* pituitary P sequence complex.

3.3.1 An *in vitro* P sequence complex that contains RFX1 does not repress CS-A promoter activity in pituitary GC cells.

3.3.1a Direct binding of RFX1 to the 103P element.

To generate an element that contained the entire PSE-A and PSE-B nuclease protection regions, a 103 bp P sequence fragment (103P) was synthesized by PCR (Fig. 3.24).

263P

5′ - TCCTACAGGC CTGCCTGGAG AACAGCTCAC AGCACAGTGC CCTCCCAGCA

GATGATGAGT CTGGGGTGCT AGTCCAGTAA TGCTTCAGGA ATGACGGCAG

PSE-B

AAAAAGAGCT CTGTTTTCTG CTCTGAAAGT GGGGAGATGG CAGGGCCCCA

PSE-A

GCATTCACAT CCTAGGCCAC AGGGGTGTGG GTGTTCAATG TTGGTTGCCA

ACACCACTGC CAACCACTTC TGGAAGCGTT TGCCTGTTTG TTTGCTTGTG

TTTCTACAGA GT - 3′

Figure 3.24: P sequences upstream of the CS-A gene.
The 5′- 3′ sequence of 263P is shown with the PSE-A and PSE-B nuclease protection regions highlighted. The 103P fragment is underlined.

The 103P element was then used as an EMSA probe with GC nuclear extract for analysis of NF-1 and RFX1 binding (Fig. 3.25). The 103P probe resulted in the formation of four complexes (I-IV) with GC nuclear extract that are indicated by arrowheads. PSE-A3, PSE-A4, and the RFX EF-C/MDBP element were all efficient competitors of complexes I and II (Fig. 3.25A, closed arrowheads). In contrast, neither PSE-B4 nor the NF-1 consensus element were efficient competitors of 103P complexes under these same conditions (Fig. 3.25B). *Thus, despite the presence of NF-1 sites at PSE-A and PSE-B, direct binding of RFX1, but not NF-1, to 103P was detected with pituitary GC nuclear extracts.*

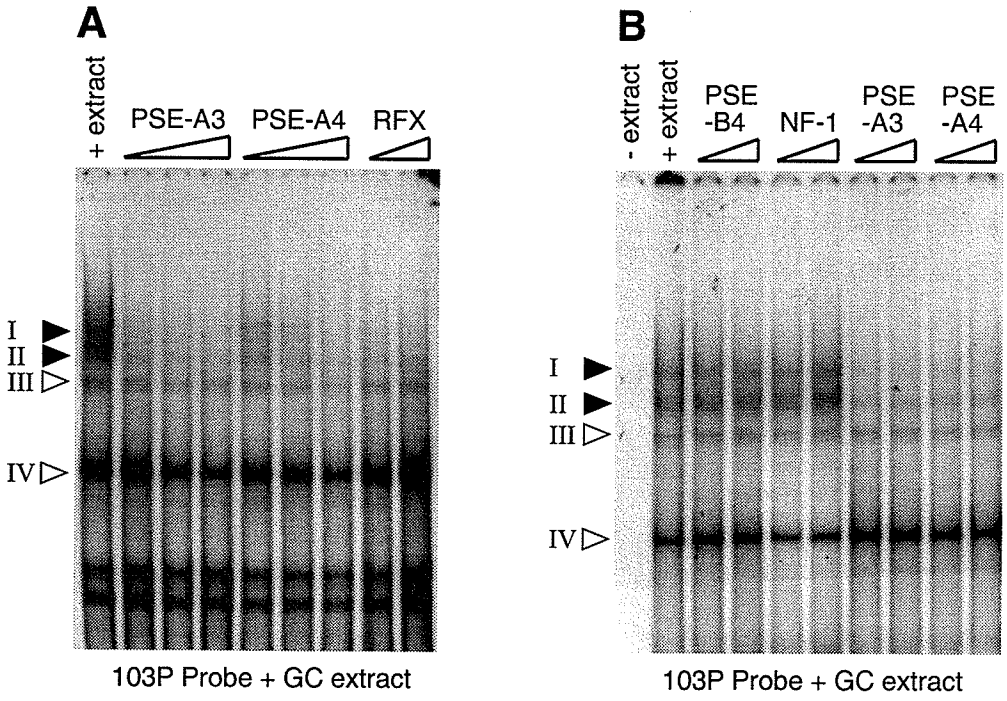


Figure 3.25: Direct binding of RFX1, but not the NF-1 family, to 103P.
Radiolabelled 103P was used as an EMSA probe with pituitary GC nuclear extract. Complexes are indicated by arrowheads and numbered. **(A)** Competitor oligonucleotides were preincubated with nuclear extract at 25-fold, 50-fold, and 100-fold (PSE-A3 and PSE-A4) or 25-fold and 50-fold (RFX) mass excess of probe. **(B)** Competitor oligonucleotides were preincubated with nuclear extract at 25-fold and 50-fold mass excess of probe.

3.3.1b The 103P element does not repress the activity of the CS-A promoter in pituitary GC cells.

Luciferase reporter genes were constructed and transiently transfected into GC cells to assess the effect of the 103P fragment on CS-A promoter function (Fig. 3.26). As a positive control, 263PCSp.Luc was included in the experiment and resulted in 36% of CSp.Luc activity (n=6, $p<0.0005$). In comparison, the presence of the 103P fragment (103PCSp.Luc) had no significant effect on CS-A promoter activity (n=6). *Thus, a functional effect of the 103P fragment on CS-A promoter activity was not detected in transiently transfected GC cells.* The lack of CS-A promoter repression with 103P indicated that the complex associating with this fragment did not represent the pituitary

repressor complex of 263P. The 263P fragment was therefore used as a probe in both nuclease protection and EMSA assays to assess the presence of NF-1 and/or RFX1 in the pituitary repressor P sequence complex.

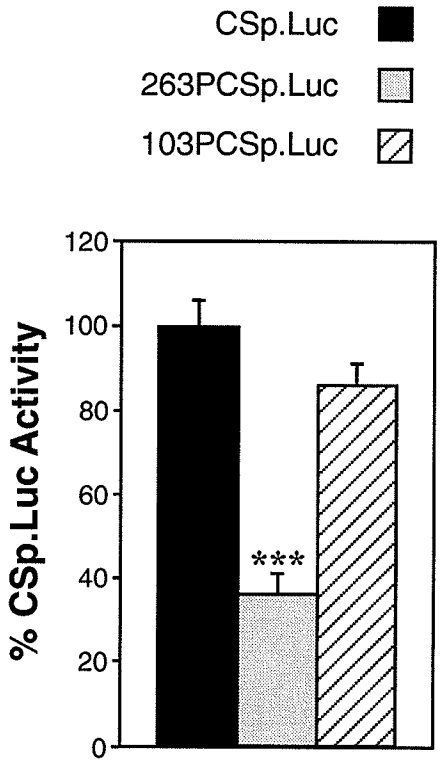


Figure 3.26: 103P does not affect CS-A promoter activity in transiently transfected pituitary GC cells.
Hybrid luciferase (Luc) genes were used to assess the effect of 103P on CS-A promoter (-496/+6, CSp.Luc) activity in GC cells. 263PCSp.Luc was included for comparison. To control for DNA uptake, cells were co-transfected with pRLTkp.Luc. Corrected values are expressed as a percentage of CSp.Luc activity, which was arbitrarily set to 100%. Bars represent SEM. ***$p<0.0005$.

3.3.2 NF-1, but not RFX1 binding is detected with 263P.

3.3.2a Competition with NF-1 elements disrupts protein binding to the 263P fragment.

The ability of the NF-1 family to interact with 263P was assessed through competition of 263P protein binding by NF-1 DNA elements, in EMSA and nuclease protection assays (Fig. 3.27). Both the consensus NF-1 element and a second NF-1 element (reported as a high affinity NF-1 binding site (162) were efficient competitors of the 263P nuclease protection pattern generated with GC nuclear extract at 1000-fold pmole excess of probe (Fig. 3.27A). As a negative control, the NF-1 mutant element was unable to compete for this pattern. It was noted that competition of the nuclease protection pattern with PSE-B4 required levels of competitor that were greater than 1000-fold pmole excess of probe. This agrees with the observation that PSE-B4 was a lower affinity NF-1 site than the consensus NF-1 site (Fig. 3.14). When the same radiolabelled 263P fragment was utilized as a probe in an EMSA with pituitary GC nuclear extract, competition of protein binding with the NF-1 consensus oligonucleotide, and not the non-NF-1 binding PSE-B4m element, was observed (Fig. 3.27B). *These results demonstrate the association of NF-1 from pituitary GC nuclear extract with 263P in vitro.*

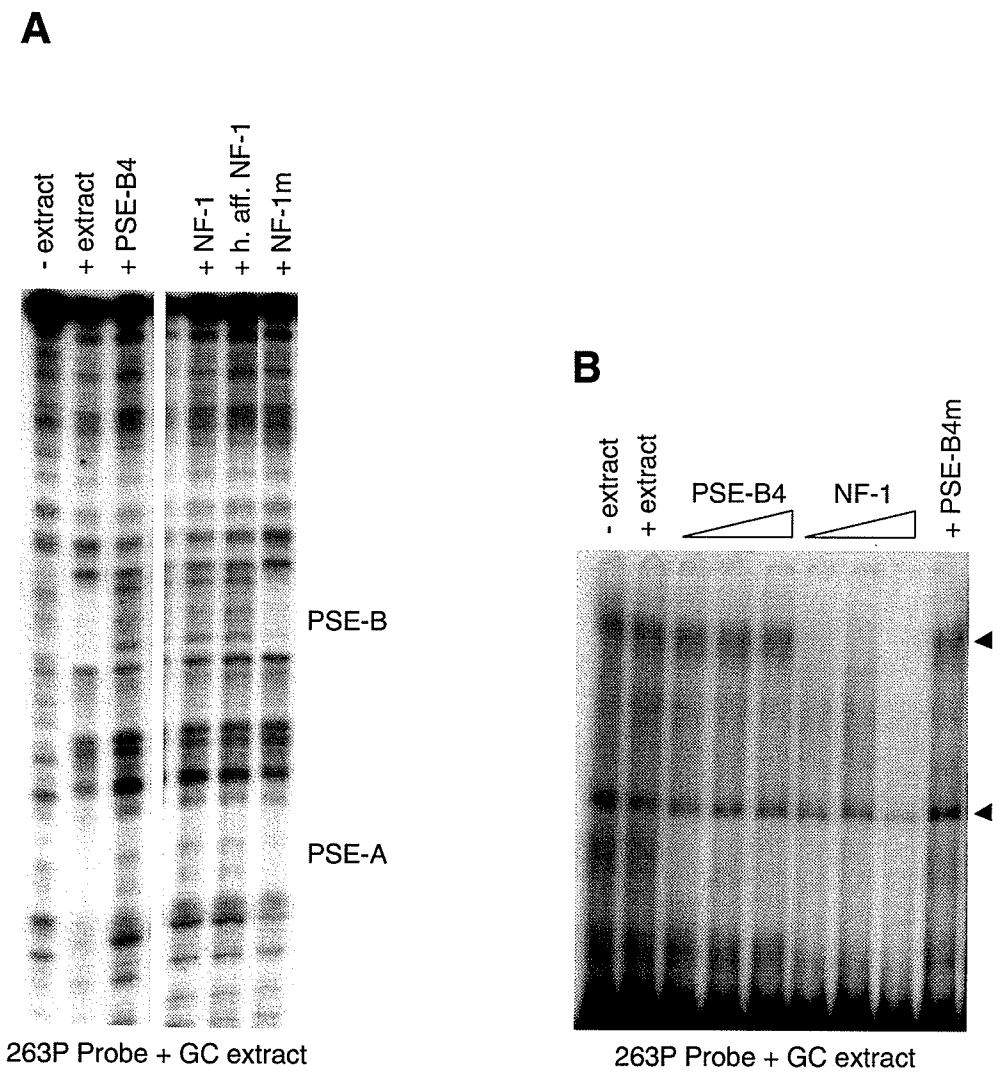


Figure 3.27: Competition with NF-1 elements disrupts protein binding to the 263P fragment.

(A) The 263P fragment was radiolabelled at the 3′end and incubated with or without pituitary GC nuclear extract before DNaseI digestion. Competitor oligonucleotides were included in the reactions at 1000-fold (NF-1s) or 5000-fold (PSE-B4) pmole excess of probe. The protected PSE-A and PSE-B regions are indicated. Data reported in (151) **(B)** The same radiolabelled 263P fragment used for nuclease protection assays was also utilized as an EMSA probe with GC nuclear extract. Competitor oligonucleotides were included in the reactions at 10000-fold, 20000-fold, and 50000-fold pmole excess of probe (PSE-B4 and NF-1 consensus) or 50000-fold pmole excess of probe (PSE-B4m). Competed complexes are indicated by closed arrowheads.

3.3.2b The EF-C/MDBP RFX element does not compete for protein binding to 263P.

The competition of 263P protein binding in EMSAs and nuclease protection assays was also used to assess the ability of the RFX family to interact with 263P (Fig. 3.28). In contrast to PSE-A3, the EF-C/MDBP RFX DNA element was unable to effectively compete for the 263P nuclease protection pattern at 5000-fold, or even as much as 20000-fold pmole excess of probe (Fig. 3.28A). The 263P radiolabelled probe was also used in EMSA with pituitary GC nuclear extract (Fig. 3.28B). The RFX oligonucleotide was unable to compete at levels where clear competition of protein binding to 263P was observed with the NF-1 consensus element. *Thus, despite evidence of RFX1 association with PSE-A3 and 103P, in vitro binding assays using the 263P fragment do not implicate RFX1 as a component of the 263P complex.*

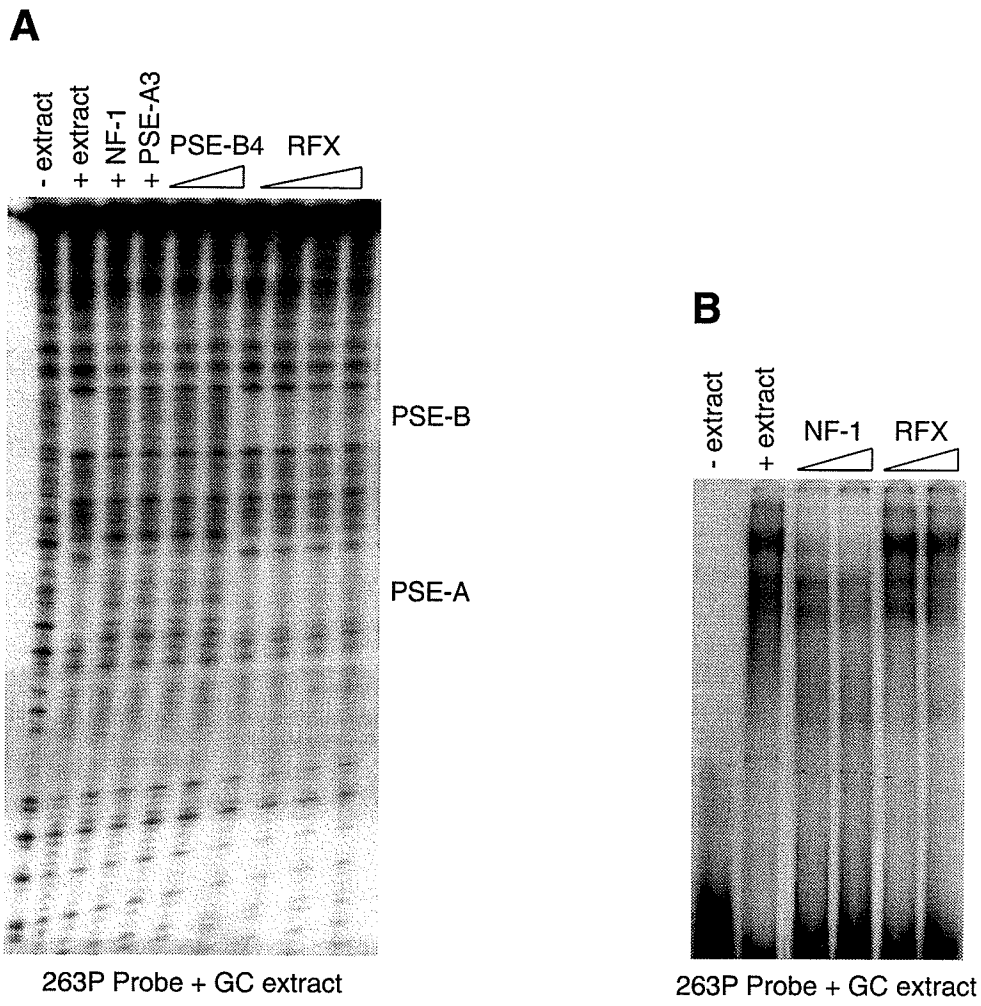


Figure 3.28: The EF-C/MDBP RFX element does not compete for protein binding to the 263P fragment.

(A) The 263P fragment was radiolabelled at the 3′end and incubated with or without pituitary GC nuclear extract before DNaseI digestion. Competitor oligonucleotides were included in the reactions at 5000-fold (NF-1 and PSE-A3), 5000-fold and 10000-fold (PSE-B4), or 5000-fold, 10000-fold, 15000-fold, and 20000-fold (RFX) pmole excess of probe. The protected PSE-A and PSE-B regions are indicated. **(B)** The same radiolabelled 263P fragment used for nuclease protection assays was also utilized as an EMSA probe with GC nuclear extract. Competitor oligonucleotides were included in the reactions at 20000-fold, and 50000-fold pmole excess of probe.

3.3.2c An NF-1 half-site in PSE-A3 contributes to its ability to compete 263P protein binding.

The PSE-A3 element is a composite element that primarily associates with RFX1, but in the absence of RFX1, is bound by the NF-1 family (Fig. 3.11). When used as a competitor of the 263P nuclease protection pattern, PSE-A3 behaves more like the NF-1 consensus element than the RFX element. The PSE-A3 scanning mutations and a 3′ truncated form of PSE-A (PSE-A4) were used to assess which region of PSE-A3 was important for its ability to compete 263P protein binding in nuclease protection assays (Fig. 3.29). Inefficient competition by any of these oligonucleotides would suggest some contribution of the modified sequences to the nuclease protection pattern. The 3′ mutation (PSE-A3m6) and a truncation (PSE-A4) of PSE-A3 are relatively inefficient competitors at 1000-fold pmole excess of the 263P probe. In contrast, at this same level of competitor, PSE-A3m2 and PSE-A3m3, which encompass the RFX1 binding site, as well as wild type PSE-A3, PSE-A3m1, PSE-A3m4 and PSE-A3m5 are all efficient competitors of the 263P nuclease protection pattern. *This result suggests that the 3′ end of PSE-A3, which localizes to one of the NF-1 half-sites, contributes more to its ability to compete for 263P protein binding than does the RFX1 binding region.* It should be noted that at 5000-fold pmole excess of probe, all of the PSE-A3 mutations compete for the nuclease protection pattern. This additional result can be explained by the fact that none of these mutations eliminate both putative NF-1 half-sites (Figs. 3.5A and 3.8A).

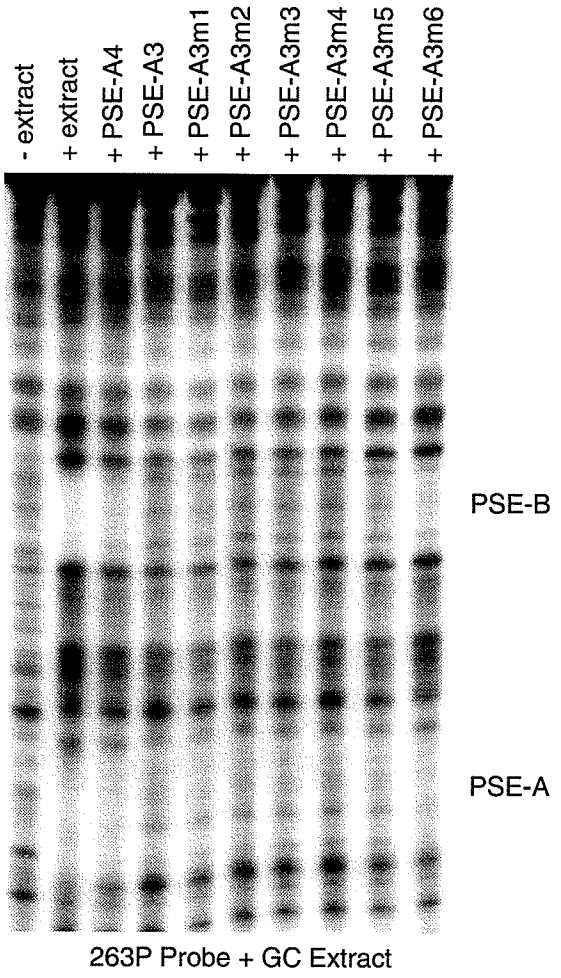


Figure 3.29: An NF-1 half-site in PSE-A3 contributes to its ability to compete for 263P protein binding.
The 263P fragment was radiolabelled at the 3′end and incubated with or without pituitary GC nuclear extract before DNase1 digestion. Competitor oligonucleotides were included in the reactions at 1000-fold pmole excess of 263P probe. The protected PSE-A and PSE-B regions are indicated. Data reported in (130).

3.3.3 Functional evidence in support of mutually exclusive NF-1 and RFX1 263P complexes.

Structural assessment of the 263P complex confirmed the presence of the NF-1 family at both PSE-A and PSE-B, but called into question association of RFX1. This result was one of the options predicted in Figure 3.3, where related proteins (members of the NF-1

family) recognize similar sequences (PSE-A and PSE-B). Nonetheless, binding of RFX1 in the absence of NF-1 was seen with both PSE-A3 and the 103P fragment. The involvement of RFX1 in the (non-repressing) 103P complex implies that two complexes are capable of forming on P sequence fragments; a non-functional RFX1-containing complex, and a repressor complex that contains NF-1 at both PSE-A and PSE-B. This interpretation is consistent with the observation of mutually exclusive binding of NF-1 and RFX1 at PSE-A3 (Fig. 3.11), and contributes to understanding the variable repressor activity with PSE-A3CSp.Luc in transiently transfected GC cells. When NF-1 is the predominant factor at PSE-A3 the resulting activity would be as a repressor, while RFX1 association at PSE-A3 would have no functional effect. This hypothesis, however, is based primarily on structural data. The involvement of NF-1 and RFX1 sites in 263P repressor activity was also characterized in functional assays.

3.3.3a The ability of 263P to repress CS-A promoter activity involves sequences at both the PSE-A/RFX and PSE-B/NF-1 sites.

The relative contributions of the PSE-B/NF-1 and PSE-A/RFX1 sites to 263P repressor activity was assessed by introducing specific mutations into the 263P fragment through site-directed mutagenesis (Fig. 3.30). The PSE-A3m2 mutation was substituted for PSE-A in 263P, and this fragment was placed upstream of the CS-A promoter to generate 263Am2CSp.Luc. The PSE-A3m2 mutation disrupts only the RFX site, leaving both of the NF-1 half-sites intact (Fig. 3.8) Likewise, the non-NF-1 binding PSE-B4 mutation was introduced into 263P and 263BmCSp.Luc was created. When these plasmids were transiently transfected into GC cells, a significant loss in repressor function was observed compared to 263PCSp.Luc. Whereas, 263PCSp.Luc displayed 43% ($n=9$, $p<0.0005$) of CSp.Luc activity (arbitrarily set to 100%), 263Am2CSp.Luc and 263BmCSp.Luc had 79% ($n=9$, $p<0.05$) and 75% ($n=9$, $p<0.005$) of CSp.Luc activity respectively.

Neither of the single mutations, however, resulted in a complete loss of repressor activity, as both of the constructs maintained significantly less luciferase activity than CSp.Luc ($p<0.005$ for both 263Am2CSp.Luc and 263BmCSp.Luc). A third construct was created, 263Am2/BmCSp.Luc, which contained both mutations upstream of CSp.Luc. When

263Am2/BmCSp.Luc was transiently transfected into GC cells, no significant difference was detected in the activity compared to CSp.Luc (Fig. 3.30, n=9). *Thus, loss of 263P repressor activity required mutation at both the PSE-A/RFX and PSE-B/NF-1 sites.*

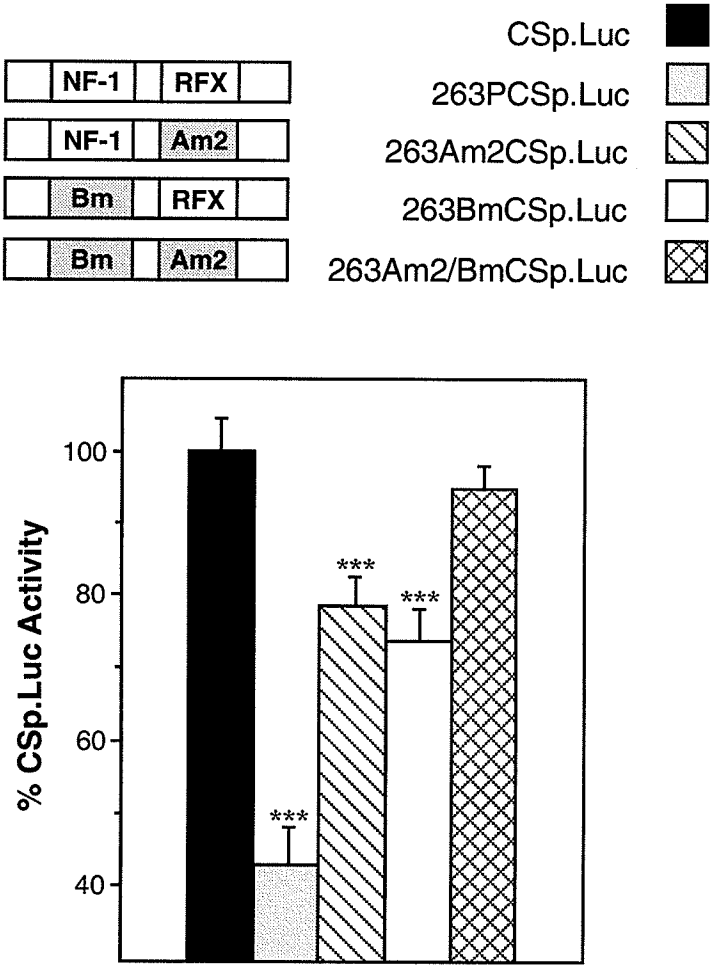


Figure 3.30: The ability of 263P to repress CS-A promoter activity involves sequences at both the PSE-A/RFX1 and PSE-B/NF-1 sites.
Hybrid luciferase (Luc) genes were used to assess the effect of 263P mutations on CS-A promoter (-496/+6, CSp.Luc) activity in pituitary GC cells. To control for DNA uptake, cells were co-transfected with pRLTkp.Luc. Corrected values are expressed as a percentage of CSp.Luc activity, which was arbitrarily set to 100%. Bars represent SEM. ***p<0.0005, compared to CSp.Luc. Data reported in (130).

Based on the hypothesis that RFX1 does not contribute to 263P repressor function, the disruption of repressor activity in 263Am2CSp.Luc was not expected. The PSE-B NF-1 site and the two PSE-A NF-1 half-sites, are maintained in this construct. It may be interpreted from this result that RFX1 binding is involved in the repressor function of 263P. It is important to remember, however, that RFX1 binding to 263P was not confirmed through either EMSA or nuclease protection competitions. Thus, it is an assumption that the result of the 263Am2 mutation is to affect RFX1 association. An alternate explanation for the 263Am2CSp.Luc result is that alteration of sequences immediately adjacent to one of the NF-1 sites may affect the function of this site. As demonstrated in Figure 3.18, this interpretation is a possibility, as sequences that do not affect NF-1 binding can affect the function of an NF-1 site. *The 263BmCSp.Luc result provides strong evidence that NF-1 binding at PSE-B contributes to 263P repression of CS-Ap activity. Additional experiments would be required to address whether RFX1 is involved in the 263P repressor complex.*

3.3.3b Overexpression of NF-1 in GC cells increases the degree of CS-A promoter repression by 263P.

To further investigate the involvement of NF-1 in the 263P pituitary repressor complex, expression vectors containing the cDNAs for different NF-1 family members were co-transfected into GC cells with 263PCSp.Luc (Fig. 3.31). The cDNA plasmids were the same c-myc epitope-tagged vectors previously described for the analysis of NF-1/RFX1 interactions (Fig. 3.21). A significant increase in the ability of 263P to repress the CS-A promoter was observed when 2 μg of NF-1B, NF-1C, or NF-1X expression plasmid was co-transfected with 263PCSp.Luc ($n=9$, $p<0.005$, $p<0.0005$, and $p<0.005$ respectively). These functional effects were specific to the 263PCSp.Luc plasmid, as the co-transfected pRLTkp.Luc was unaffected by NF-1 overexpression (data not shown). The lack of NF-1A effects on 263PCSp.Luc function may be due to the observation that, although identical with the exception of the cDNA insert to the other NF-1 expression vectors, the NF-1A expression vector has been consistently less efficient in protein production (Xiaoyang Yang, pers. comm.). *These experiments provide further support that NF-1 is the factor responsible for repression.*

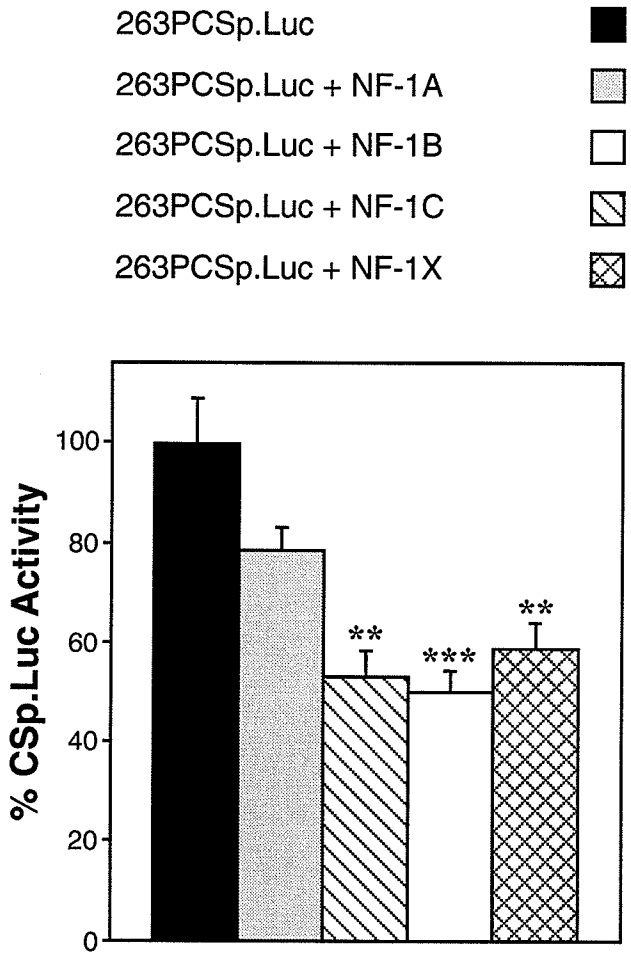


Figure 3.31: Overexpression of NF-1 in GC cells increases the degree of CS-A promoter repression by 263P.
Expression plasmids containing the cDNAs for NF-1A, NF-1B, NF-1C, and NF-1X (2 μg) were transiently co-transfected with 263PCSp.Luc into pituitary GC cells. To control for DNA uptake, cells were co-transfected with pRLTkp.Luc. There was no significant effect of NF-1 over-expression on pRLTkp.Luc activity (data not shown). Corrected values are expressed as a percentage of 263PCSp.Luc activity with no NF-1 over-expression, which was arbitrarily set to 100%. Bars represent SEM. **p<0.005, ***p<0.0005.

The increase in the degree of repression following NF-1 overexpression also supports the possibility that mutually exclusive NF-1 and RFX1 complexes form on P sequence fragments. In this model, both complexes exist on 263P in GC cells. The luciferase

activity of 263PCSp.Luc (which is present in a transfected cell in multiple copies) is the sum of the total activity from both NF-1 repressor complexes and non-functional RFX1 complexes. In this scenario, overexpression of NF-1 alters the balance of NF-1 and RFX1 proteins within the cell. As the NF-1 increases relative to RFX1, the likelihood of NF-1 repressor complex formation also increases. Therefore, the observed decrease in the luciferase activity of 263PCSp.Luc when NF-1 is overexpressed, would be explained by an increase in NF-1 repressor complexes and consequent reduction in non-functional RFX1 complexes.

3.3.3c Overexpression of RFX1 in GC cells enhances luciferase activity independent of 263P.

In light of the above interpretation, it would be predicted that overexpression of RFX1 in GC cells would have the opposite effect on the average luciferase activity of co-transfected 263PCSp.Luc. That is, RFX1 overexpression would alter the cellular balance of RFX1 to NF-1, such that the likelihood of non-functional RFX1 complex formation would increase. The result would be an increase in the average luciferase activity of 263PCSp.Luc, or derepression. To assess if RFX1 overexpression can affect 263P function, 2 μg of RFX1 expression plasmid was co-transfected with 263PCSp.Luc (Fig. 3.32). The overexpression of RFX1 in GC cells resulted in a 2.8-fold increase in 263PCSp.Luc activity ($n=18$, $p<0.0005$), consistent with the derepression hypothesis. The RFX1 enhancement, however, was not specific to 263PCSp.Luc. Although RFX1 overexpression had no significant effect on the activity of co-transfected pRLTkp.Luc, a significant 2.0-fold increase in 263Am2CSp.Luc ($n=9$, $p<0.005$) activity was observed. 263Am2CSp.Luc is the 263P construct with a mutation in the RFX1 binding site. *Therefore, although RFX1 overexpression enhanced the luciferase activity of 263PCSp.Luc, it could not be demonstrated that this increase was directly related to the proposed RFX site in 263P.*

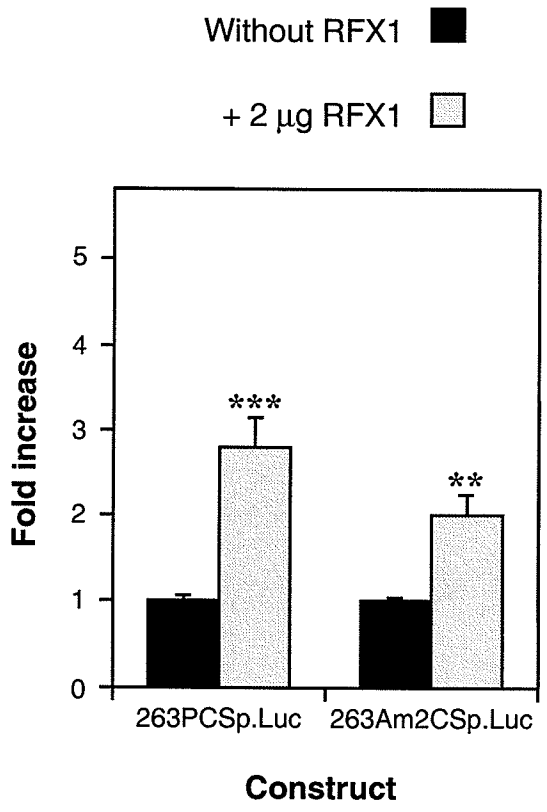


Figure 3.32: Overexpression of RFX1 in GC cells enhances luciferase activity independent of 263P.

The hybrid luciferase (Luc) genes 263PCSp.Luc, 263Am2CSp.Luc, and CSp.Luc were transiently transfected into pituitary GC cells with or without 2 µg of RFX1 expression plasmid as indicated. To control for DNA uptake, cells were co-transfected with pRLTkp.Luc. RFX1 over-expression had no effect on pRLTkp.Luc activity (data not shown). Corrected values are expressed as fold increase over transfections without RFX1 over-expression. Bars represent SEM. *p<0.05, **p<0.005, ***p<0.0005.

SUMMARY OF CHAPTER 3 RESULTS

- The data presented in this section indicate that 263P can act as a repressor in pituitary cells independent of the CS-B downstream enhancer region.

- To identify candidate PSFs, the PSE-A and PSE-B oligonucleotides were used. PSE-A3 was examined in isolation and found to be a composite element, capable of mutually exclusive NF-1 and RFX1 binding in pituitary GC cells. This statement is supported by the observations of: (i) specific RFX1 binding to PSE-A3 in GC EMSA assays (ii) the loss of RFX1 complexes and the appearance of a novel complex when the RFX site was mutated in PSE-A3m2 and (iii) evidence of NF-1 binding to PSE-A3m2 but not PSE-A3 or PSE-A3m6. These observations have important implications for the interpretation of PSE-A results. The presence of other factors and/or sequences when PSE-A3 is not in isolation, could potentially influence whether RFX1 or NF-1 is the predominant associating factor. That is, whether NF-1 or RFX1 associate with PSE-A may depend on the context of the element. The variability of the PSE-A3CSp.Luc construct in transient transfection experiments can be explained by the potential for at least two different PSE-A complexes to exist. The results presented in section 3.2.1 demonstrate that in the pituitary, both RFX1 and NF-1 are PSF-A candidates.

- The analysis of PSE-B4 in isolation implicated members of the NF-1 family as PSF-B in pituitary cells. This was supported by: (i) the competition of PSE-B4-specific EMSA complexes by NF-1 oligonucleotides and antibodies and (ii) the determination that PSE-B4 repressor activity required NF-1 binding. In addition, analysis of PSE-B4 provided the first evidence that variations in an NF-1 recognition site, including the spacer region, could affect the functional activity of individual NF-1 elements. The possibility was discussed that this functional effect was the result of alterations in the affinity profile for different members of the NF-1 family.

- The observation of 263P function in JAR and JEG-3 cell lines provides the first evidence for significant *in vitro* P sequence function in placental cells. When considered with previously reported observations (133), the data suggest that functional cross-talk may occur between 263P and the downstream CS enhancer region. This possibility is further discussed in Chapter 6.3.1. A nuclease protection pattern with JAR nuclear proteins was observed to be similar to the pattern that is evident with GC nuclear proteins. Taken together, these observations indicated that PSE-A and PSE-B may be useful for the identification of putative placental P sequence binding proteins.

- In choriocarcinoma cell lines, the analysis of PSE-A3 and PSE-B4 oligonucleotides through EMSA also supported RFX1 and NF-1 as PSF candidates. The NF-1 and RFX families have both been linked to repressor and enhancer function. The expression profiles of the NF-1 and RFX proteins in the pituitary and placenta may therefore be an important aspect in determining whether the P sequences have enhancer or repressor activity.

- The pituitary repressor complex was evaluated in the context of larger fragments that contained both PSE-A and PSE-B. Interpretation of the results in section 3.3 generates a model whereby at least two separate complexes have the potential to form on P sequence fragments in GC cells; an NF-1 repressor complex, and a non-functional RFX1 complex. This interpretation is supported by: (i) the observation of RFX1 and not NF-1 association with the non-functional 103P fragment (ii) the confirmation of NF-1 but not RFX1 binding to 263P and (iii) an increase in 263PCSp.Luc repressor activity following NF-1 overexpression. This model implicates the NF-1 family and not RFX1 in the repressor function of 263P *in vitro*. It agrees with the predicted complex in Figure 3.3, whereby two related proteins (the NF-1 family members) recognize highly similar sites in PSE-A and PSE-B. A more detailed discussion of the pituitary P sequence complex can be found in Chapter 6.1.1.

- The capacity for RFX1 and the NF-1 family to interact was supported by observations of: (i) low affinity EMSA competitions, (ii) co-immunoprecipitation of overexpressed epitope-tagged proteins and (iii) co-immunoprecipitation of endogenous NF-1 and RFX1 proteins from GC nuclear extracts. The possibility was discussed that these interactions may involve the NF-1 NH_2-terminal DNA-binding domain. Implications of RFX1 and NF-1 family interactions that pertain to P sequence complexes is discussed further in Chapter 6.1.1a.

CHAPTER 4

THE P SEQUENCE COMPLEX *IN SITU*

The *in vitro* examination of PSE-A and PSE-B implicated RFX1 and NF-1 as PSE-A binding factors (3.2.1) and NF-1 as PSF-B (3.2.2). However, analysis of 263P in pituitary GC cells, strongly supported involvement of the NF-1 family in the repressor complex, but called into question the participation of RFX1 (3.3). An important consideration in interpreting the data to this point has been the context of the fragment that is being investigated. For example, PSE-A3 was clearly a mutually exclusive RFX1 and NF-1 element that predominantly associated with RFX1 over NF-1 *in vitro*. In the context of 263P, however, only NF-1 binding to PSE-A was detected. When the P sequence elements are placed in the context of chromatin, this may also influence factors that associate. As discussed in Chapter 1, the structure of chromatin limits the accessibility of DNA sequences. This is an important consideration with the identification of NF-1 as a candidate PSF, as the ability of NF-1 to associate with its target site is reported to be reduced in the presence of nucleosomes (204). The association of RFX1 with chromatin compared to naked DNA has not been reported. Consequently, in determining which candidate PSFs associate with P sequences in the pituitary and placenta, it is important to include an assessment of their ability to associate in the context of chromatin. To examine if RFX1 and the NF-1 family were participating in pituitary and/or placental P sequence complexes *in situ*, a chromatin immunoprecipitation (ChIP) technique with human tissue samples was developed.

4.1 Development of the ChIP technique for analysis of human tissue samples.

The presence of the CS/GH-V duplication, and thus P sequence DNA, is exclusive to primate lineages (83). As a result, the pituitary GC cell line, which originates from a rat anterior pituitary adenoma, does not contain endogenous P sequences. This assumption is supported by NCBI BLAST analysis (http://www.ncbi.nlm.nih.gov/BLAST/). For

assessment of the pituitary P sequences complex *in situ*, post-mortem human pituitary tissue was obtained from the Human Pituitary Repository in the Protein and Polypeptide Laboratory at the University of Manitoba.

The decision was also made to use human term placenta tissue for analysis of the placental P sequence complex *in situ*. Expression of the endogenous GH-V and CS genes in human term placenta is much higher than in the human choriocarcinoma cell lines (160). P sequences were hypothesized to play a role in enhancing placental gene expression *in vivo* (150). Thus, the use of human term placenta tissue enables analysis of the P sequence complex in a context where the endogenous genes are highly active. Human term placentas were obtained from term deliveries at the Health Sciences Centre, Winnipeg, Manitoba.

Unlike cultured cell lines, where cross-linking can be done in culture dishes at the beginning of a ChIP experiment, human tissue samples require processing of the tissue prior to cross-linking proteins to the DNA. Placental syncytiotrophoblast is a syncytium, in which very few individual cells are present at term. For consistency between the two tissue types, nuclei were isolated from both tissues prior to cross-linking (Fig. 4.1).

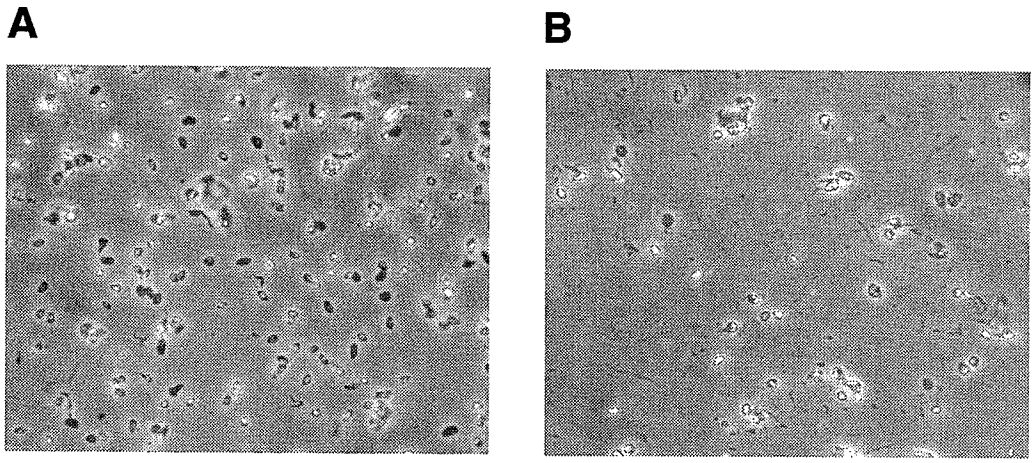


Figure 4.1: Isolated human placenta and pituitary nuclei.
For the ChIP protocol, nuclei were isolated from human term placenta and post-mortem pituitary tissue samples prior to formaldehyde cross-linking. Examples of **(A)** placenta and **(B)** pituitary nuclei preparations, under 100x magnification, are shown. The pituitary nuclei preparation that is shown here was provided by Mr. Scott Gregoire.

Following the isolation of intact nuclei from the tissue samples, the conditions established for ChIP using cultured cells were then used (Fig. 4.2). Within these established conditions, however, there are several steps that needed to be assessed, such as the cross-linking time, fragmentation of DNA by sonication, the amount of chromatin used for immunoprecipitation, specificity of the antibodies for immunoprecipitation, and the method used to analyze the final DNA product (205-207).

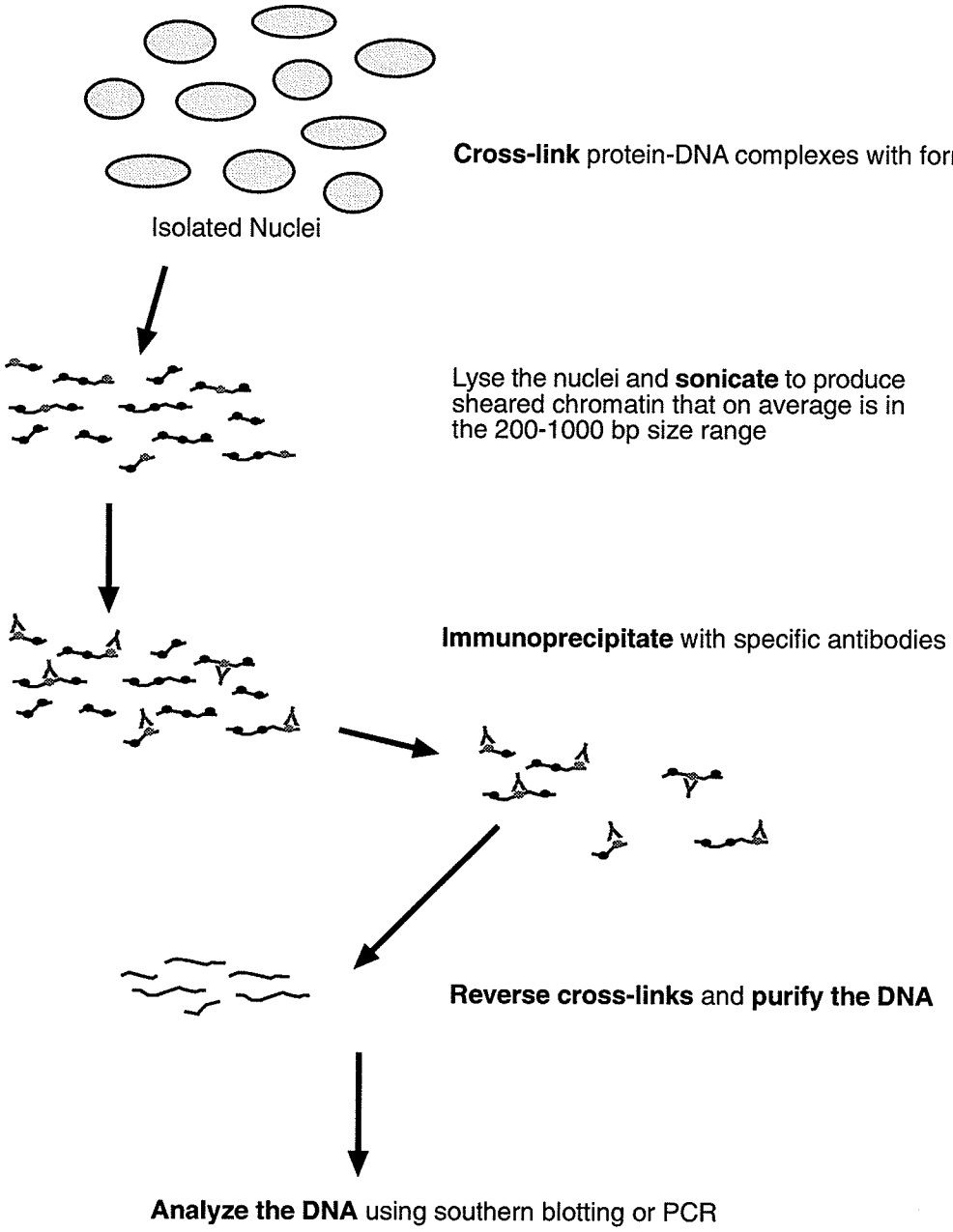


Figure 4.2: The ChIP protocol requires assessment of several key steps.
The general ChIP procedure is outlined. The amount of time the nuclei are cross-linked, the sonication and immunoprecipitation conditions, as well as the method to analyze the final DNA product, must all be established.

4.1.1 Determination of the cross-linking time.

The time used for cross-linking proteins to DNA in cultured cell lines are generally in the range of 5-15 minutes (206, 208, 209). As cells or nuclei are exposed to formaldehyde for increasing lengths of time, not only are the DNA-binding proteins cross-linked, but the proteins that are in contact with the DNA-binding proteins are cross-linked as well (205-207). This becomes a concern as cross-linking time increases, because cross-linking can extend from the DNA and DNA-binding proteins to the insoluble nuclear matrix (209). When the matrix is removed before immunoprecipitation, any of the DNA that has been cross-linked to the matrix will be removed as well. To evaluate the length of time that the nuclei would be cross-linked in this protocol, placental nuclei were incubated in a HEPES buffer (see Chapter 2.2.1 for more details) containing 1% formaldehyde for 5, 10 and 15 minutes, with all other components of the protocol kept constant. Placental nuclei were chosen over pituitary nuclei for this experiment due to the relative availability of the tissue samples. Following cross-linking, the nuclei were lysed and the chromatin was sonicated to fragment the DNA. The absorbance (A_{260}) values of the lysates were measured as an indication of total chromatin. The lysates were then microfuged to remove the insoluble nuclear matrix, and the A_{260} of the supernatant was measured. At 5 and 10 minutes of cross-linking time, 93% and 91% of the chromatin content was retained in the supernatant. At 15 minutes of cross-linking time, only 67% of the chromatin was retained. *Based on these observations, a cross-linking time of five minutes in 1% formaldehyde was chosen for the protocol.*

4.1.2 Optimization of the sonication conditions

The cross-linked nuclei are lysed and the chromatin is mechanically sheared by sonication. The amount of sonication must be optimized for each tissue type. The aim is to establish sonication conditions where the majority of the chromatin fragments fall into the size range of 200 – 1000 bp. This range is an indication of the resolution of the protocol. *When determining the cross-linking time, it was found that sonicating the chromatin from lysed placental nuclei at 40% output (Vibra Cell™, Sonics and Materials Inc., Danbury, CT, USA) for a total time of two minutes in 30 second bursts yielded average fragment sizes in the appropriate range.* The isolated DNA from three separate

experiments is shown as an example of the fragment size expected with this protocol (Fig. 4.3A). To determine the optimum sonication conditions for pituitary chromatin, pituitary nuclei that had been cross-linked for five minutes with 1% formaldehyde were lysed and sonicated in 20 second intervals. The DNA was isolated and run in a 2% agarose gel to assess the average length of the fragments. *For pituitary, sonicating for a total time of 1.5 minutes in 30 second intervals was chosen for the protocol, as this yielded an average fragment length within the appropriate range (Fig. 4.3B).*

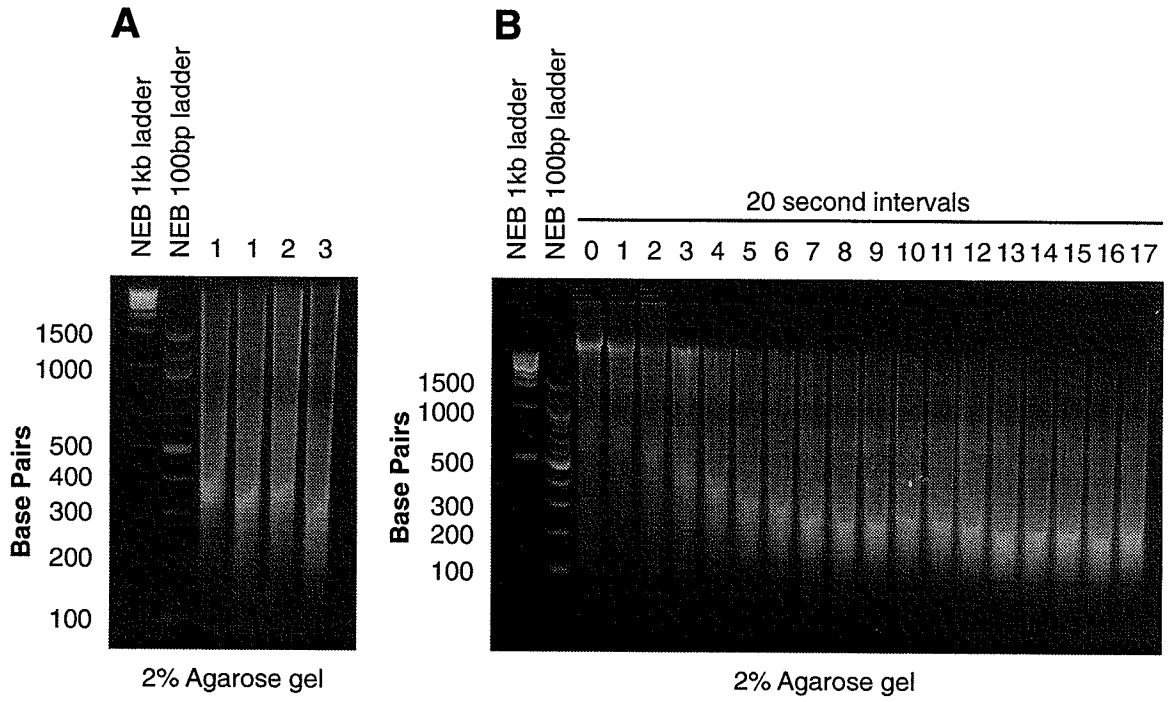


Figure 4.3: Optimization of the sonication conditions for placenta and pituitary chromatin.
The sonication conditions differ depending on the tissue type. **(A)** For the placenta nuclei, four intervals of 30 seconds each was used to shear the chromatin. The majority of the chromatin was in the size range of 300-400 base pairs (bp). The isolated DNA from three separate experiments (1,2,3) is shown. **(B)** Cross-linked pituitary nuclei were lysed and sonicated in 20 second intervals 0-17 times. The isolated DNA was run in a 2% agarose gel to determine the sonication conditions which would yield fragments in the appropriate size range. The sonication conditions for the pituitary ChIP experiments were determined to be three bursts of 30 seconds each. Photo of pituitary sonication data provided by Mr. Scott Gregoire.

4.1.3 Determining the amount of chromatin input for the immunoprecipitations

The ChIP procedure provides a method to enrich for sequences that are bound by a protein of interest; however a level of non-specific background DNA is expected (205-207). The background of non-specific sequences increases with the amount of chromatin used as the starting material for the immunoprecipitations (the input). To investigate the effect of increasing chromatin input, 2U (A_{260}), 5U, and 10U of placental chromatin were used in immunoprecipitations with an antibody to the hyperacetylated form of histone H4 (Upstate Biotechnology Inc.). A chromatin input of 2U was used as a starting amount, based on a protocol obtained from the laboratory of Dr. J. Davie (University of Manitoba). Immunoprecipitations with the same input levels and no antiserum were used to represent non-specific background. The DNA from these immunoprecipitations (the bound fraction) was isolated and used as the template in PCR reactions with 263P specific primers (Fig. 4.4). Previous studies indicated that 263P was hyperacetylated in human term placental chromatin (150). *The level of chromatin input chosen for the immunoprecipitations was 2U (approximately 100 µg), as this amount showed the greatest difference between the PCR product from the no antibody (control) and the hyperacetylated histone H4 (test) bound samples.*

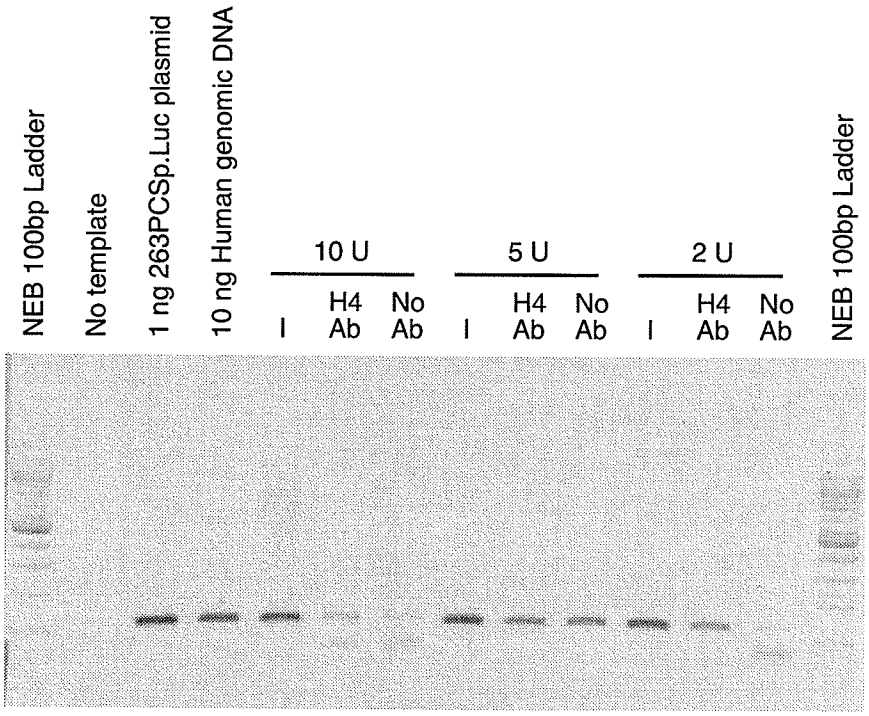


Figure 4.4: Chromatin input above 2U into the immunoprecipitations decreases the difference between the specific and background signals.
ChIP was done using 10 U, 5 U, or 2 U of chromatin input with hyperacetylated histone H4 antibody (H4Ab). Five μL of isolated DNA from each immunoprecipitation, as well as from the input chromatin, was used as a PCR template with internal 263P primers (103P forward and reverse). As positive controls, 1 ng of 263PCSp.Luc plasmid DNA and 10 ng of human genomic DNA (isolated from placenta tissue) were used as templates. As a negative control, sdH2O (no template) was used. For the PCR reactions, an annealing temperature of 60 °C was used for 30 cycles. Equal volumes of each reaction were run in a 2% agarose gel to compare the intensity of the PCR products.

4.1.4 Specificity of the NF-1 and RFX1 antibodies for use in ChIP

The hyperacetylated histone H4 antibody (Upstate Biotechnology Inc.) is recommended for use in ChIP assays; however neither NF-1 nor RFX1 had previously been assessed by ChIP. The specificity of the commercial NF-1 and RFX1 antibodies (Santa Cruz) for immunoprecipitation experiments was therefore assessed. JAR cells were transiently transfected with plasmids containing NF-1 and RFX1 cDNAs, cloned in frame as fusions to a 13 amino acid c-myc epitope tag. Lysates from the transfected cells were immunoprecipitated with the NF-1 and RFX1 antibodies. The immunoprecipitations were

separated by SDS-PAGE, transferred to PVDF membranes, and probed with a c-myc antibody (Fig. 4.5). When the NF-1 immunoprecipitations were probed with the c-myc antibody, bands that corresponded to the NF-1 bands in the transfected cell lysates were observed (Fig. 4.5A). Similarly, when the RFX1 immunoprecipitation was probed with the c-myc antibody, a band corresponding to RFX1 was observed (Fig. 4.5B). *These experiments established that the NF-1 and RFX1 commercial antibodies could be used for the ChIP assays.*

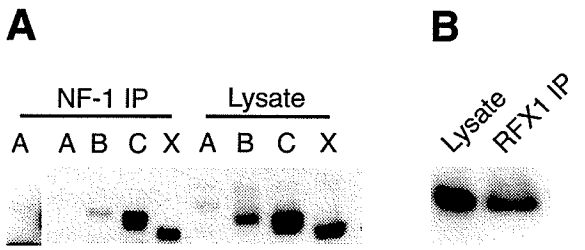


Figure 4.5: Specificity of the NF-1 and RFX1 antibodies for use in ChIP.
(A) Immunoprecipitations with NF-1 specific antibodies were run alongside lysates (100 μg) from JAR cells transfected with myc-NF-1A (A), myc-NF-1B (B), myc-NF-1C (C), or myc-NF-1X (X). The transferred blot was probed with c-myc antibody. A longer exposure of the NF-1 immunoprecipitation from myc-NF-1A JAR lysate is shown to the left (A′). **(B)** Lysate (100 μg) from JAR cells transfected with myc-RFX1 was run alongside a RFX1 immunoprecipitation. The transferred blot was probed with c-myc antibody. This data was provided by Mr. Scott Gregoire. Data reported in (130).

4.1.5 Analysis of the ChIP DNA.

The DNA that is isolated after immunoprecipitation can be analysed by Southern (DNA) blot or PCR (206, 208). Under specific ChIP conditions, the DNA yield from a single immunoprecipitation can be expected to be in the range of a few nanograms (206). Despite the decision to scale up the protocol (for more information see Chapter 2.2.1), sufficient DNA for Southern blot analysis was not expected. PCR was chosen as the method to analyze the DNA due to both the low yield, as well as the ability of PCR to differentiate between highly similar sequences.

In the literature, two methods can generally be found to correct ChIP PCR results. The first method involves comparing the intensity of the PCR product from the bound fraction of the specific antibody immunoprecipitation to the PCR product from the bound fraction of a control immunoprecipitation where either no antiserum or a different antiserum is used (210, 211). This method was used to differentiate the amount of background from different levels of chromatin input (Fig. 4.4). It has been reported, however, that antiserum alone can increase the amount of DNA that is recovered in the bound fraction even if the sequences are not specific (207). PCR results that are corrected through comparison of an antibody immunoprecipitation to a no-antibody immunoprecipitation may therefore, be inadequately controlled. An alternative method for correcting the PCR results was chosen, where a primer set to unrelated sequences is used to establish the background level of non-specific DNA in the bound fraction of a single immunoprecipitation (212-214). This method of correction has the additional advantage of providing a standard within each immunoprecipitation that can be used to compare and pool different experiments for statistical analysis. A primer set for the third exon of the human fibroblast growth factor (FGF)–16 gene was designed to represent the background of non-specific sequences in each immunoprecipitation. FGF-16 is a member of the FGF family that is specifically expressed in embryonic brown adipose and adult cardiac tissue (215). These sequences are not expected to be hyperacetylated in either placental or pituitary chromatin, and do not contain any putative binding sites for the RFX or NF-1 transcription factor families (166).

PCR products from the ChIP assays were electrophoresed and digital images were assessed by densitometry. The use of the FGF-16 primer set as a control for the specificity of the ChIP assay requires that the different primer sets in the experiment be comparable. Where possible, efforts were made to ensure that primer pairs were of a similar melting temperature (Tm) and the products were of a similar length. There are some variations, however, that are intrinsic; for example, the human genome contains four copies of the P sequence target for every one copy of FGF-16 exon 3 target. To correct for these inherent differences between the different primer sets, 10 ng of chromatin input was used as a template for PCR with each primer set. The values from

PCR reactions where the immunoprecipitated (bound) fraction served as a template were then expressed as a ratio of the input PCR value, the bound/input (B/I) ratio. An assumption in this method of correction is that there is a high degree of consistency in setting up the different PCR reactions. This assumption was tested by determining both the intra-experimental and inter-experimental variation (Fig. 4.6). The intra-experimental variation was assessed by setting up 15 PCR reactions with the FGF-16 exon 3 primers from the same stock of template DNA (10 ng) (Fig. 4.6A). Comparison of the PCR products by densitometry showed that with the FGF-16 exon 3 mean value set to 1, the ± standard deviation was 0.08, indicating very low levels of intra-experimental variation. To assess the inter-experimental variation, PCR reactions were set up in triplicate using FGF-16 exon 3 primers and 10 ng of input DNA from three separate experiments (Fig. 4.6B). There was no significant difference between the mean values of any of the three groups. *These experiments demonstrated low levels of both intra-experimental and inter-experimental variation. The high degree of consistency between the different PCR reactions allowed for confidence in the correction method that would be used for the PCR analysis of the ChIP assay DNA.*

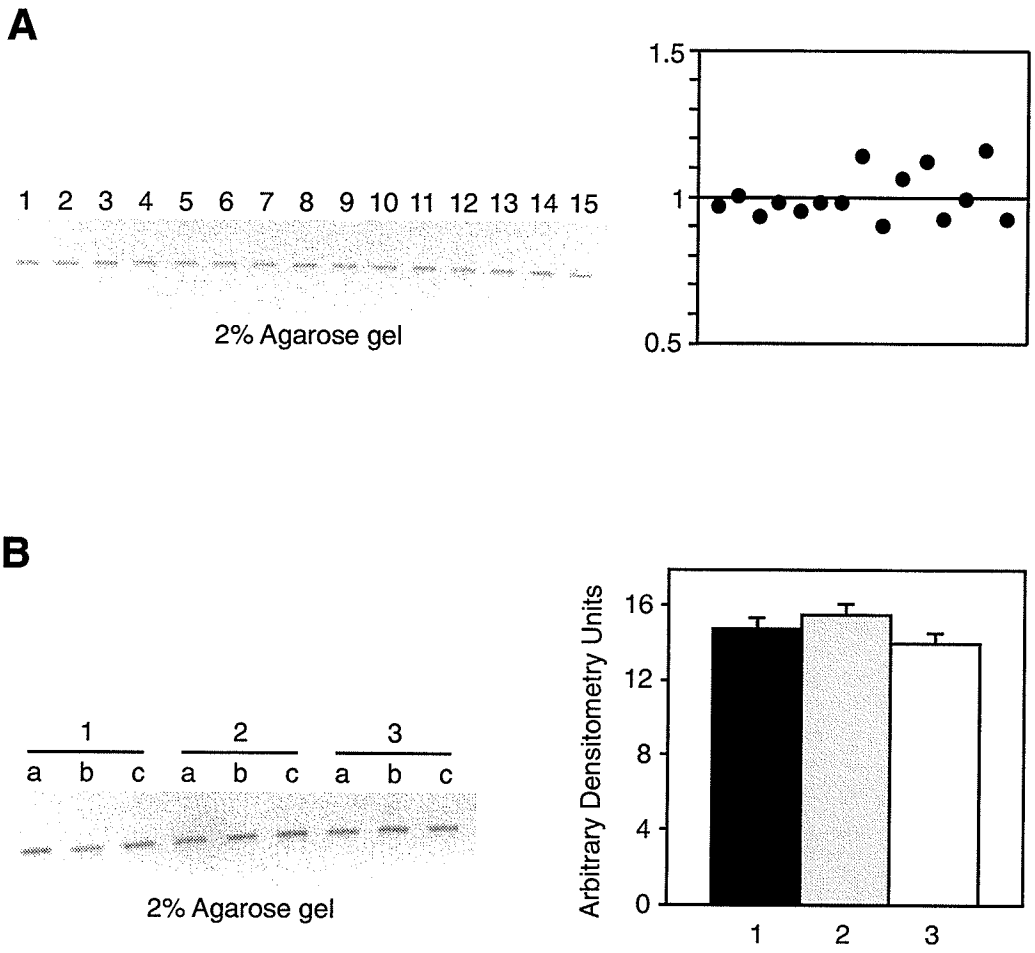


Figure 4.6: The intra-experimental and inter-experimental variations between PCR reactions are not statistically significant.

(A) Intra-experimental variation. Fifteen PCR reactions were set up using the same DNA template stock and FGF-16 exon 3 primers. The PCR products were electrophoresed and the average densitometry value was set to 1. Variation of each individual reaction from the mean is shown. The standard deviation from the mean was 0.08. **(B)** Inter-experimental variation. PCR reactions were set up in triplicate (a,b,c) from three different DNA samples (1,2,3) using FGF-16 exon 3 primers. The PCR products were electrophoresed and the average densitometry value for each of the DNA samples is shown. There is no significant variation between any of the samples.

4.2 Analysis of histone H4 hyperacetylation of the GH/CS locus in human pituitary and term placenta tissue.

4.2.1 Histone H4 hyperacetylation of the GH/CS locus in human term placenta.

Histone hyperacetylation of the human GH/CS locus in human term placental nuclei was reported previously using an alternate technique (150). In part to verify the technique that was developed in the previous section (4.1), the hyperacetylation of histone H4 in various regions of the GH/CS locus was analyzed (Fig. 4.7). Using a specific antibody for the hyperacetylated form of histone H4, evidence of hyperacetylation was observed in the upstream distal locus control region (LCR) at HS IV and HS III related sequences in term placenta chromatin (Fig. 4.7). The mean bound/input (B/I) ratio for HS IV was 1.36 and the mean B/I ratio for HS III was 1.90 (n=3). Both of these ratios were significantly higher than the 1.03 mean B/I ratio for FGF-16 exon 3 (n=3). This is in contrast to the previously published report that HS V and HS III, but not HS IV, were hyperacetylated in human term placenta nuclei (150). In the experiments reported in Figure 4.7, HS V hyperacetylation was not statistically significant , possibly due to variability. In two of the three experiments the B/I ratio for HS V averaged at 2.3, while in one experiment the B/I ratio was 1.3. *Thus, taken together, it appears that HS III is a strong candidate for a role in placental LCR function, and that the HS IV and HS V regions may be involved as well.* This is consistent with an apparent requirement for the LCR to achieve regulated expression of GH/CS genes in the placentas of transgenic mice (See Table 1.1).

Histone H4 hyperacetylation was also observed more proximally to the GH/CS genes. Both GH and CS promoters (GHp and CSp), as well as 263P and the downstream 241 bp CS enhancer regions (ENH), resulted in significantly higher B/I ratios than the FGF-16 exon 3 background control (n=3). ChIP with hyperacetylated H4 antibody yielded a mean B/I ratio for GHp of 1.77, for CSp of 2.57, for 263P of 1.8, and for ENH of 1.62. The previous report of placental hyperacetylation indicated that 263P was highly acetylated relative to other portions of the GH/CS locus (approximately 17-fold higher than background chromatin) (150). The results in Figure 4.7 instead suggest that a 'peak' of

hyperacetylation may not occur at 263P, and that multiple regions, including the CS enhancer region, which has been the focus of numerous *in vitro* studies (136-149, 172), may cooperate in placental regulation of the GH/CS locus.

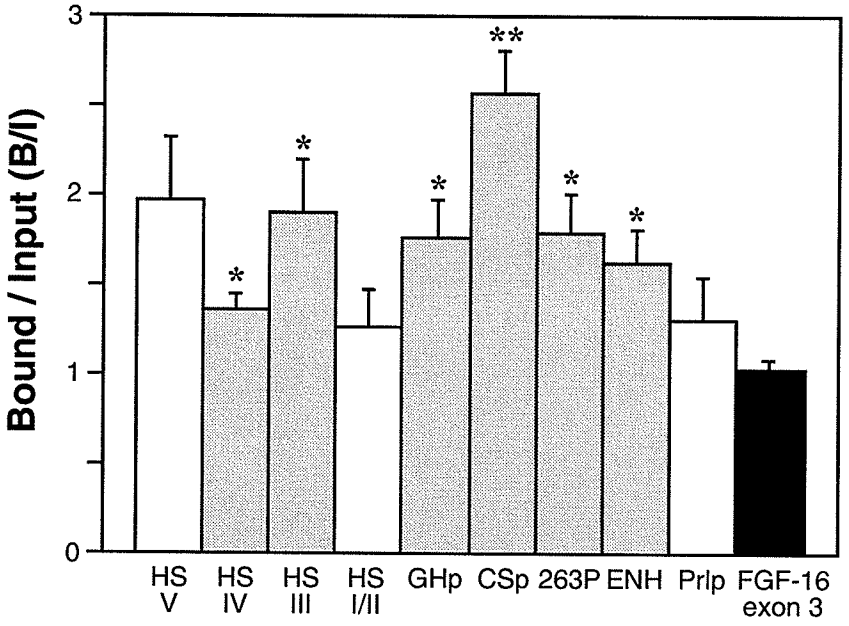


Figure 4.7: Histone H4 hyperacetylation of the GH/CS locus in human term placenta chromatin.
The mean B/I ratios for each primer set. Significant increases over FGF-16 exon 3 background indicates hyperacetylation of the region and is denoted by gray bars. Basal value for FGF-16 exon 3 was 1.03 ± 0.05. Bars represent SEM. *p<0.05, **p<0.005. HS, hypersensitive site; GHp, growth hormone promoter; CSp, chorionic somatomammotropin promoter; ENH, 241 bp CS downstream enhancer; Prlp, prolactin promoter.

4.2.2 Histone H4 hyperacetylation of the GH/CS locus in human pituitary tissue.

The analysis of hyperacetylation in the human GH/CS locus was extended to include the pituitary tissue samples. Previous demonstration of histone hyperacetylation in pituitary tissue came from a GH-secreting human pituitary adenoma sample and double transgenic mouse lines that were enriched in somatotrophs and contained an 87 kb fragment of human chromosome 17 (123, 129, 150). As mentioned in the previous section, the ChIP assay technique was developed to analyze the binding events at 263P in normal human

pituitaries taken post-mortem. Under my supervision, Ms. Xiaoyang Yang utilized the optimized ChIP protocol for analysis of histone H4 hyperacetylation in human pituitary samples. In analysis of this tissue, several regions of the locus were observed to be hyperacetylated, and the data were similar to previously reported observations from other pituitary tissue samples (123, 129, 150).

In the locus control region, HS I/II, HS III, and HS IV regions all resulted in significantly higher B/I ratios than the FGF-16 background control (Fig. 4.8). ChIP with hyperacetylated H4 antibody yielded a mean B/I ratio for HS I/II of 1.25, for HS III of 2.23, for HS IV of 1.27, and for FGF-16 exon 3 of 0.35 (n=3). As in the placenta ChIP assays, the HS V region was not significant due to variability. In two of three experiments the B/I ratio averaged at 1.87, while the third experiment was 0.56. In contrast to the previous studies, a 'peak' of hyperacetylation was not observed at HS I/II (123, 129, 150). A more detailed comparison of these results with those that have been previously reported can be found in Chapter 6.2.2a.

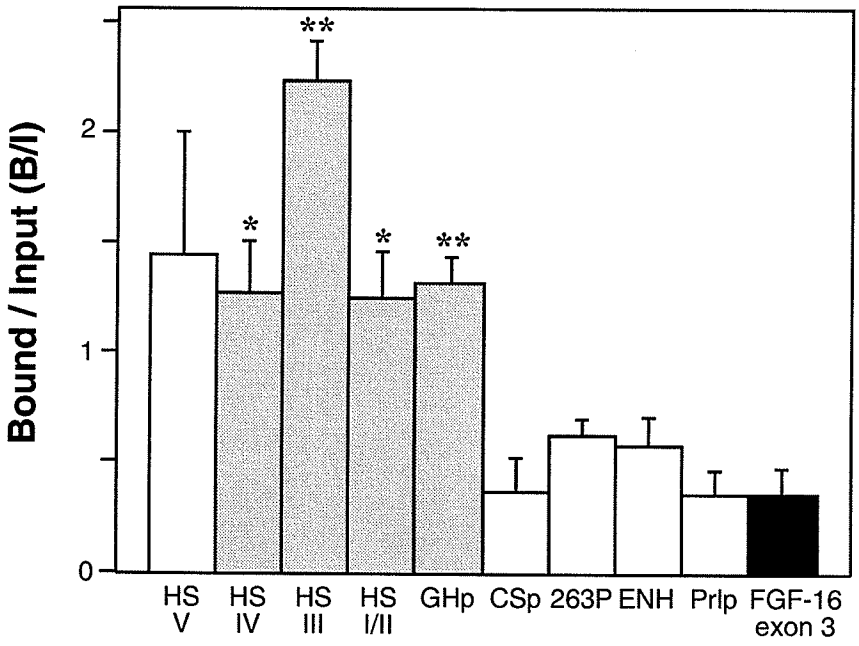


Figure 4.8: Histone H4 hyperacetylation of the GH/CS locus in human pituitary chromatin.
the mean B/I ratios for each primer set. significant increases over FGF-16 exon 3 background indicates hyperacetylation of the region and is denoted by gray bars. Basal value for FGF-16 exon 3 was 0.35 ± 0.12. Bars represent SEM. *p<0.05, **p<0.005. HS, hypersensitive site, GHp, growth hormone promoter, CSp, chorionic somatomammotropin promoter, ENH 241 bp CS downstream enhancer, Prlp, prolactin promoter.

An advantage to these experiments was the use of PCR for analysis of the data. Unlike previous reports that used a hybridization technique (slot blots) for the analysis of the ChIP DNA, through the design of specific primers, these experiments distinguished between events at the highly homologous GH and CS promoters. The GH promoter resulted in a mean B/I ratio of 1.31, which was significantly higher than background, while the CS promoter mean B/I ratio nearly equaled that of background at 0.37. This hyperacetylation pattern mirrors the expression of these genes in the pituitary.

4.3 NF-1 and RFX1 associate with 263P in human pituitary tissue.

The ChIP assays were then used to assess the association of NF-1 and RFX1 with 263P in human pituitary samples (Fig. 4.9). No indication of the length of time between death and the preservation of the tissue samples was available for the whole post-mortem human pituitaries. Therefore, as an assessment of the integrity of the pituitary chromatin in each assay, an immunoprecipitation with antibodies for the hyperacetylated form of histone H4 was run in parallel. Hyperacetylation of the HS I/II region was previously reported, and this region was shown to be required for high-level pituitary expression of the GH gene in transgenic mice (120, 122-124, 129, 150). Thus, the criteria established to determine whether data for a pituitary experiment would be used was that the B/I ratio for HS I/II PCR had to be significantly greater than the FGF-16 exon 3 B/I background levels. If hyperacetylation of the HS I/II region was not detected, the integrity of the chromatin was questionable and the assay was not used. When the NF-1 antibody was used for ChIP with human pituitary chromatin, the mean bound/input (B/I) ratio for 263P from four separate immunoprecipitations was 0.4 (Fig. 4.9B). This ratio was significantly higher than the mean B/I ratio of the FGF-16 exon 3 PCR from the same four immunoprecipitations (0.18, $p<0.05$). A representative PCR result is shown from one of the four NF-1 immunoprecipitations in Figure 4.36A. *These results indicate the association of NF-1 with P sequences in situ.*

When the RFX1 antibody was used for ChIP with human pituitary chromatin, the mean B/I ratio for FGF-16 exon 3 from four separate immunoprecipitations was 0.31. The B/I ratio for the 263P primers from these same four immunoprecipitations was 0.63, which was significantly higher than the FGF-16 background levels (Fig. 4.36B, $p<0.05$). A representative PCR result is shown from one of the four RFX1 immunoprecipitations in Figure 4.9A. *Therefore, the ChIP assays with human pituitary chromatin indicate that RFX1 associates with P sequences in situ.*

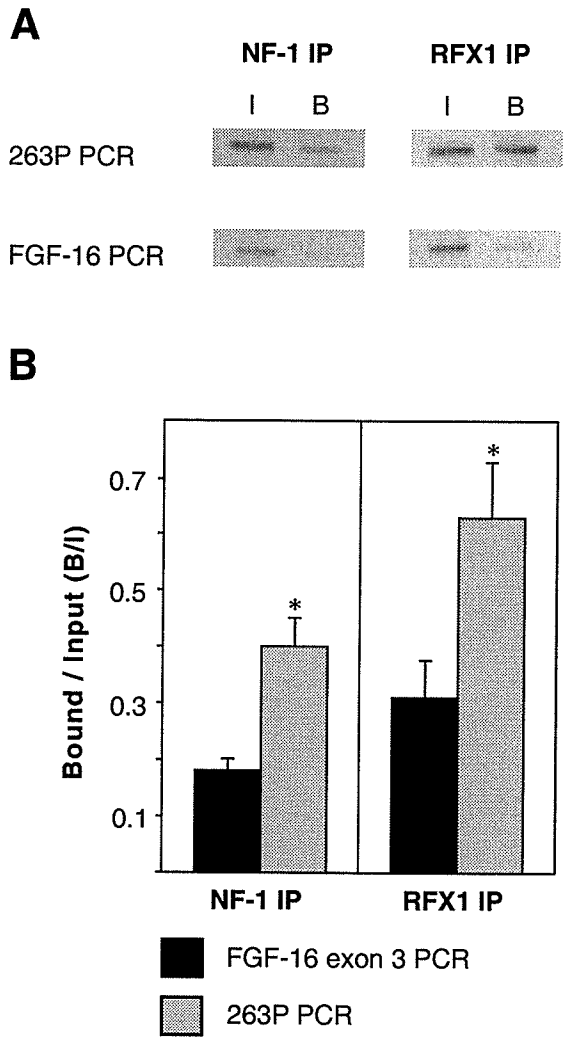


Figure 4.9: NF-1 and RFX1 associate with 263P in human pituitary tissue.
(A) A representative PCR result is shown for each of the primer sets for both the NF-1 and RFX1 immunoprecipitations (IP). PCR amplifications were done on 10 ng of input (I) DNA as well as 5 μL of immunoprecipitated/bound (B) DNA. **(B)** The PCR products were run in 2% agarose gels and digital images were assessed by densitometry. The mean B/I ratio from four separate immunoprecipitations with each antibody is shown. Bars represent SEM. *$p<0.05$. This data was provided by Ms. Xiaoyang Yang. Data reported in (130).

4.4 The association of NF-1 and RFX1 with P sequences was not be detected in human term placenta tissue.

The *in vitro* analysis of P sequence fragments also indicated that the NF-1 family and RFX1 were PSF candidates in the placenta. The optimized ChIP protocol was used to investigate whether NF-1 or RFX1 association with 263P could be detected in human term placenta samples. *In contrast to our observations in pituitary nuclei, ChIP assays using human placenta nuclei provided no evidence of either NF-1 or RFX1 association with 263P (Fig. 4.10).* Use of NF-1 antibody for ChIP yielded a mean B/I ratio for 263P of 0.99, which was not significantly different from the FGF-16 exon 3 B/I ratio of 0.67 (n=3). Likewise, when the RFX1 antibody was used for ChIP, the 263P B/I ratio of 0.72 was not significantly different from the FGF-16 exon 3 B/I ratio of 0.65.

The detection of NF-1 and RFX1 association with P sequences in human pituitary chromatin was the first report of these factors being used successfully in ChIP assays from human tissue samples. As a consequence, previously characterized examples of NF-1 or RFX1 association with defined sequences *in situ* were not available for use as positive controls. Thus, it is unclear from these experiments if in fact these factors do not associate with P sequences, or simply that the conditions used have not permitted their detection.

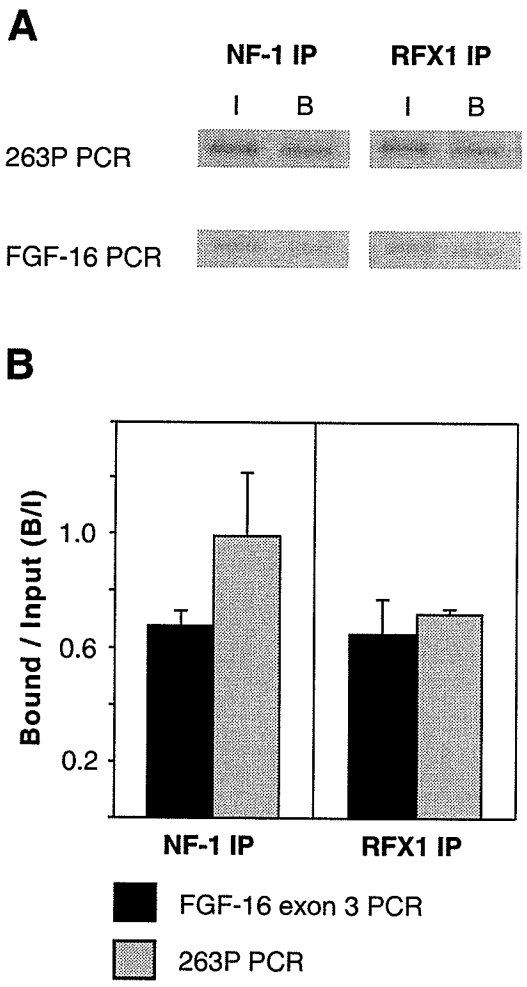


Figure 4.10: The association of NF-1 and RFX1 with P sequences was not detected in human term placenta tissue.
(A) A representative PCR result is shown for each of the primer sets for both the NF-1 and RFX1 immunoprecipitations (IP). PCR amplifications were done on 10 ng of input (I) DNA as well as 5 μL of immunoprecipitated/bound (B) DNA. **(B)** The PCR products were run in 2% agarose gels and digital images were assessed by densitometry. The mean B/I ratio from three separate immunoprecipitations is shown. Bars represent SEM.

SUMMARY OF CHAPTER 4 RESULTS

- To examine if RFX1 and/or the NF-1 family were participating in P sequence complexes *in situ*, a chromatin immunoprecipitation (ChIP) protocol was developed and optimized for use with human tissue samples (4.1).

- The conditions established for ChIP were used to assess regions of the GH/CS locus for hyperacetylation of histone H4 in human term placenta and pituitary tissue samples taken post-mortem (4.2). A comparison was made between the data obtained by this technique and a previously reported analysis of human term placenta chromatin (150). Previous reports of GH/CS hyperacetylation were extended, in that the results in Figures 4.7 and 4.8 represent histone H4 hyperacetylation, and not a mixture of histone H3 and histone H4 hyperacetylation (129, 150). Additionally, the analysis of normal human pituitaries taken post-mortem expanded the analysis of pituitary hyperacetylation patterns, which had previously only been reported for tissue samples in which GH is overexpressed (123, 129, 150). A comparison of the pituitary results from 'normal' pituitary samples and pituitary samples that overexpress GH can be found in Chapter 6.2.2a.

- In human term placenta chromatin, evidence of histone H4 hyperacetylation was observed in the distal locus control region (LCR) at HS IV and HS III, but not HS V related sequences. A previous report documented that HS V and HS III, but not HS IV, were hyperacetylated in human term placenta nuclei (150). When these two reports are considered together, HS III appears to be a strong candidate for contributing to placental LCR function. The HS IV and HS V regions may also be involved. This possibility is consistent with an apparent requirement for the LCR to achieve regulated expression of GH/CS genes in the placentas of transgenic mice (120, 122). A discussion of placental GH/CS locus regulation *in vivo* can be found in Chapter 6.3.2a.

- In the placental analysis, histone H4 hyperacetylation was observed throughout regions more proximal to the GH/CS genes, as well as in the LCR. A previous report demonstrated a 'peak' of hyperacetylation at 263P (150). In contrast, the GH and CS promoters as well as the downstream CS enhancer region, all resulted in hyperacetylation B/I ratios that approximated 263P levels. These observations suggest that multiple proximal regions may cooperate with the distal LCR in placental regulation of the GH/CS locus. This possibility is also discussed further in Chapter 6.3.2.

- ChIP of human pituitary chromatin samples demonstrated histone H4 hyperacetylation for the GH, but not the CS promoters, which correlates with the expression pattern of these genes in the pituitary. The ability to observe this difference was due to the use of specific primers and PCR for the analysis of the ChIP chromatin, as opposed to the use of a hybridization technique that cannot distinguish between events at the highly homologous sequences. Histone H4 hyperacetylation of other regions proximal to the genes, such as 263P and the downstream CS enhancers, was not observed in pituitary chromatin. The contribution of these observations to interpreting the mechanisms involved in the pituitary GH/CS expression pattern is discussed further in Chapter 6.2.2.

- The association of NF-1 and RFX1 with P sequences in pituitary chromatin was demonstrated by the ChIP assay (Fig. 4.9). This technique does not differentiate between factors that are associated in common complexes from those forming mutually exclusive complexes in different cells from the same tissue sample. Currently, there is not a protocol for human tissue samples that would differentiate between these two pituitary complex models. An interpretation of the P sequence complex *in situ*, and a possible mechanism for repression *in vivo*, are discussed in Chapter 61.1b and 6.2.2b, respectively.

- The NF-1 and RFX1 data were the first report in the literature (130) that documented the association of these transcription factors with specific sequences using ChIP and human tissue samples.

- The *in vitro* analysis of P sequence fragments also indicated that the NF-1 family and RFX1 were PSF candidates in the placenta. The optimized ChIP protocol was used to investigate whether NF-1 or RFX1 association with 263P could be detected in human term placenta samples. In contrast to the observations in pituitary nuclei, ChIP assays using human placenta nuclei provided no evidence of either NF-1 or RFX1 association with 263P (Fig. 4.10). The use of positive controls for NF-1 and RFX1 in the placental ChIP assays was not possible. It is unclear from these experiments if in fact these factors do not associate with P sequences, or simply that the conditions used have not permitted their detection. The placental P sequence complex is discussed further in Chapter 6.1.1c.

CHAPTER 5

CHARACTERIZATION OF A POTENTIAL PITUITARY REPRESSOR MECHANISM INVOLVING THE TRANSCRIPTION FACTOR PIT-1.

263P represses activity of the CS-A promoter in pituitary GC cells (Fig. 3.1). Activity of the CS-A promoter in GC cells is driven by the pituitary-specific transcription factor Pit-1 (117, 164, 216) (reviewed in Chapter 1.2.1). The original characterization of 263P, PSE-A, and PSE-B, demonstrated that the P sequence binding events were competed by high levels of Pit-1 oligonucleotide (133). This suggested that the then unidentified PSFs may interact with Pit-1. Taken together, these data indicate that the mechanism of 263P repressor activity in pituitary cells may involve interference with Pit-1 transactivation of the CS-A promoter.

5.1 263P repressor function is linked to Pit-1 activation of the CS-A promoter in pituitary GC cells.

5.1.1 263P enhances the activity of a minimal thymidine kinase promoter (Tkp).

To examine a possible link between 263P repressor activity and the CS-A promoter, 263P was placed upstream of an unrelated minimal promoter. The -81/+52 fragment of the viral thymidine kinase promoter (Tkp) does not contain Pit-1 binding sites. The 263P fragment was placed upstream of Tkp in the plasmid pT81luc (168) (Tkp.Luc) to create 263PTkp.Luc. The relative luciferase activities of 263PTkp.Luc and Tkp.Luc were assessed in transiently transfected pituitary GC, placental JEG-3, cervical carcinoma HeLa, and embryonic kidney 293 cells (Fig. 5.1). *Presence of 263P significantly enhanced the activity of Tkp in all cell lines tested.* In GC cells, 263PTkp.Luc had 2.4-fold higher luciferase activity than Tkp.Luc (n=6, $p<0.005$). In JEG-3 cells the increase

was 2.9-fold (n=12, $p<0.0005$), and in 293 cells the increase was 2.4-fold (n=12, $p<0.0005$). HeLa cells also demonstrated increased 263PTkp.Luc luciferase values relative to Tkp.Luc. The increase that was observed in HeLa cells could not be converted to a fold value, as Tkp.Luc activity was not above background levels (the background level is the luciferase activity measured in lysis buffer). The corrected activity of 263PTkp.Luc, however, was 1.2-7.2 fold greater than the corrected background luciferase values (n=12), indicating that 263P enhanced the minimal (undetected) activity of Tkp in HeLa cells. *These results indicate that 263P activity can be converted from a repressor to an enhancer by substituting Tkp for CS-Ap*. These observations raise the possibility that cross-talk between the 263P and promoter complexes determines the direction of functional activity.

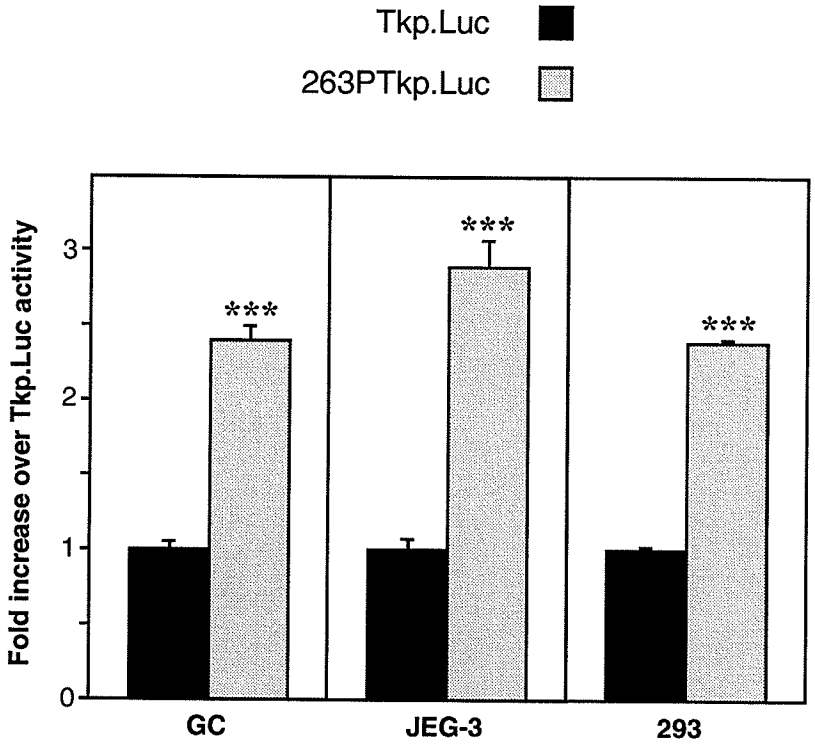


Figure 5.1: 263P enhances Tk promoter activity.
Hybrid luciferase (Luc) genes were used to assess the effect of 263P on activity of the Tk promoter in human JEG-3 and 293 cells. To control for DNA uptake cells were co-transfected with pRLTkp.Luc (JEG-3 and 293) or RSVp.CAT (GC). Corrected values are expressed as fold increases over Tkp.Luc activity, which was arbitrarily set to 1. The corrected basal activity for TKp.Luc was 0.15 ± 0.01 for GC, 2.49 ± 0.14 for JEG-3, and 81.1 ± 1.1 for 293. Bars represent SEM. ***$p<0.0005$.

5.1.2 263P represses Pit-1-mediated activity of the CS-A promoter in pituitary GC cells.

The above data suggest that the repressor activity of 263P in pituitary GC cells involves the CS-A promoter complex. In GC cells there are several possible CS-A promoter complexes, and the 'major' activating complex involves the transcription factor Pit-1 (reviewed in Chapter 1.2.1). As was observed in Chapter 3, 263P does not completely eliminate CS-A promoter activity in GC cells, as 263PCSp.Luc averaged at 44% of CSp.Luc activity. This result may be explained in at least two ways. The first possibility

assumes that the NF-1 repressor complex does not occupy all of the transfected 263PCSp.Luc vectors within the cells. This was discussed in reference to mutually exclusive NF-1 and RFX1 complexes in Chapter 3, and supported by the ability of NF-1 overexpression to increase the degree of 263P repression. A second possibility is that the NF-1 complex does not have the capacity to repress all permutations of CS-A promoter complexes that can occur in the GC cell. These two options are not mutually exclusive, and together may contribute to the inability of 263P to completely abolish CSp.Luc activity.

The presence of 263P in non-Pit-1 cell types, such as JEG-3 and JAR, still results in repression of CS-A promoter activity (Fig. 3.2). This indicates that 263P also has the capacity to function as a repressor in the absence of the Pit-1 driven CS-A promoter complex. In pituitary GC cells, however, the predominant activator of CS-A promoter activity is Pit-1. So much so, that disruption of the Pit-1 binding site by mutation eliminates the detection of reporter gene activity in a lower activity (chloramphenicol acetyltransferase) system (117). This was the same reporter gene system that was used in the original characterization of 263P repressor activity, where greater than 90% repression was observed (133). The inability to detect 'basal' or non-Pit-1 mediated promoter complex activity with this system suggests that the repressor mechanism of 263P involved Pit-1 activation. Although suggestive, evidence as to which CS-A promoter complex was repressed by 263P in pituitary GC cells was not determined. To address this issue, a mutation within the CS-A promoter Pit-1 binding site was introduced into CSp.Luc (CSpPit-1mLuc). CSpPit-1mLuc was created by *Nsi*I digestion of CSp.Luc, removal of the 3′ overhang nucleotides with Klenow fragment, and religation of the construct. This is the same mutation that was previously shown to eliminate binding of Pit-1 to the proximal site, and reduce promoter activity in pituitary GC cells (104, 133).

The 263P fragment was placed upstream of the modified CS-A promoter (263PCSpPit-1m.Luc) and the relative luciferase activities of these two constructs were assessed in transiently transfected GC cells (Fig. 5.2). For comparison, 263PCSp.Luc and CSp.Luc were included in this experiment. The 263PCSp.Luc construct resulted in 49% of the

luciferase activity of CSp.Luc, which was arbitrarily set to 100% (n=12, $p<0.0005$). The results in Figure 5.2 demonstrate that the CS-A promoter has the capacity to be activated independent of Pit-1. Consistent with this, CSpPit-1m.Luc was a viable promoter in GC cells, resulting in 25% of CSp.Luc activity (n=12, $p<0.0005$) (for comparison TKp.Luc activity was 3.5% of CSp.Luc activity in separate experiments, n=10, $p<0.0005$). The activity of CSpPit-1m.Luc represents CS-A promoter activity that is not mediated by Pit-1. Mutation of the Pit-1 binding site in the CS-A promoter eliminated the ability of 263P to function as a transcriptional repressor, as the activities of 263PCSpPit-1m.Luc and CSpPit-1m.Luc were not significantly different (n=12). *These results indicate that 263P is not a repressor of basal CS-A promoter activity in pituitary GC cells, and that 263P repressor function is linked to Pit-1 activation of the CS-A promoter in GC cells.*

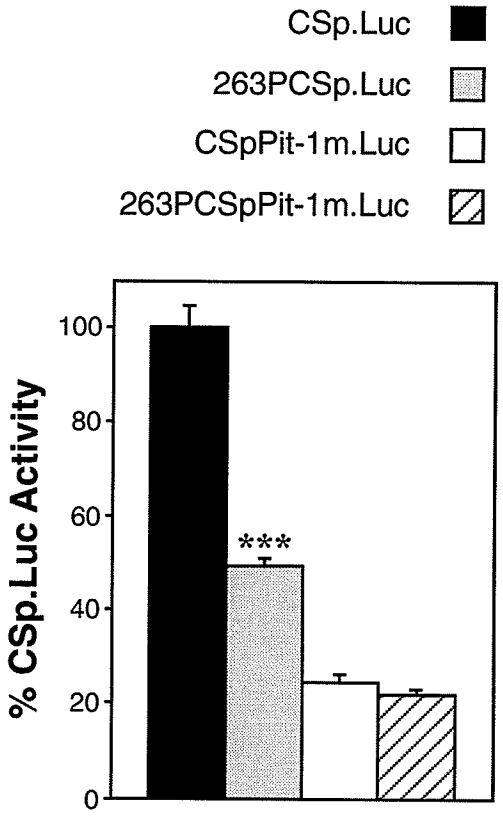


Figure 5.2: 263P does not repress non-Pit-1 mediated activity of the CS-A promoter in pituitary GC cells.
Hybrid luciferase (Luc) genes were used to assess the effect a mutation in the Pit-1 binding site of the CS-A promoter (-496/+6) on 263P repressor activity in pituitary GC cells. To control for DNA uptake cells were co-transfected with pRLTkp.Luc. Corrected values are expressed as a percentage of CSp.Luc activity, which was arbitrarily set to 100%. The corrected basal activity for CSp.Luc was 3.22 ± 0.15. Bars represent SEM. ***$p<0.0005$.

5.2 Potential interactions between Pit-1 and the PSFs.

5.2.1 EMSA and nuclease protection competitions suggest that Pit-1 has the capacity for interactions with the PSFs.

5.2.1a Pit-1 oligonucleotides compete for protein binding to P sequence elements.

In the original characterization of 263P, a potential for PSF and Pit-1 interactions was suggested by the ability of a Pit-1 DNA element to compete for protein binding from P sequence elements (133). These results were reassessed using two representative Pit-1 sites, the proximal Pit-1 binding site of the GH-N promoter spanning nucleotides –60/-100, called Pit-1(GH), and the proximal Pit-1 site from the rat prolactin promoter, called Pit-1(Prl). These oligonucleotides were used as competitors of the 263P GC nuclease protection pattern (Fig. 5.3). At 10,000-30,000 fold pmole excess of probe, both of the Pit-1 elements demonstrated some competition of protein binding to 263P. At these levels of competitor, an unrelated control oligonucleotide (RF) did not compete. The RF element is a component of the CS-B enhancer region (138). The 263P GC nuclease pattern appears to represent the NF-1-containing 263P complex (Chapter 3). *Competition of the 263P complex with Pit-1 elements suggested that Pit-1 may interact with NF-1.*

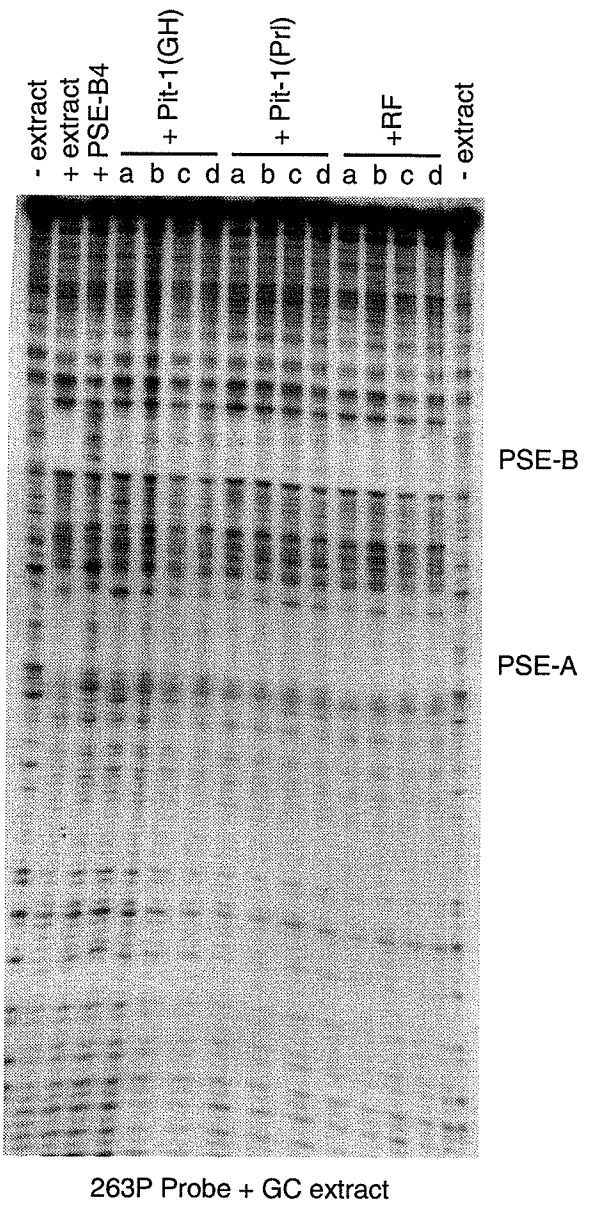


Figure 5.3: Pit-1 oligonucleotides compete for 263P protein binding.
The 263P fragment was radiolabelled at the 3′ end and incubated with or without pituitary GC nuclear extract before DNase1 digestion. Competitor oligonucleotides were included in the reactions at (a) 5000-fold, (b) 10000-fold, (c) 15000-fold, and (d) 30000-fold (GHpPit-1, PRLpPit-1, and RF) or 5000-fold (PSE-B4) pmole excess of probe. The protected PSE-A and PSE-B regions are indicated.

The Pit-1(GH) element was also used as a competitor of PSE-A3 and PSE-B4 EMSA complexes (Fig 5.4). When PSE-A3 was used as an EMSA probe with pituitary GC nuclear extract, competition of complexes III and IV by Pit-1(GH) was detected at 100-fold pmole excess of probe, and competition of complexes I and II was observed at 1000-fold pmole excess of probe (Fig. 5.4A). These competitions were specific, as they were not seen when an unrelated control oligonucleotide (Egr-1) was used. The Pit-1(GH) element was also able to compete for PSE-B4 complexes (Fig. 5.4B). When PSE-B4 was used as an EMSA probe with pituitary GC nuclear extract, Pit-1(GH) specifically competed complexes at 1000-fold pmole excess of probe (closed arrowhead). Competition of higher mobility complexes (open arrowheads) was observed at 100-fold pmole excess of probe but was determined to be non-specific, as competition of these complexes was also observed with the unrelated control Egr-1 oligonucleotide. *The competition of PSE-A3 and PSE-B4 EMSA complexes by Pit-1 elements suggested that Pit-1 may interact with RFX1, and was a further indication that Pit-1/NF-1 interactions may occur.*

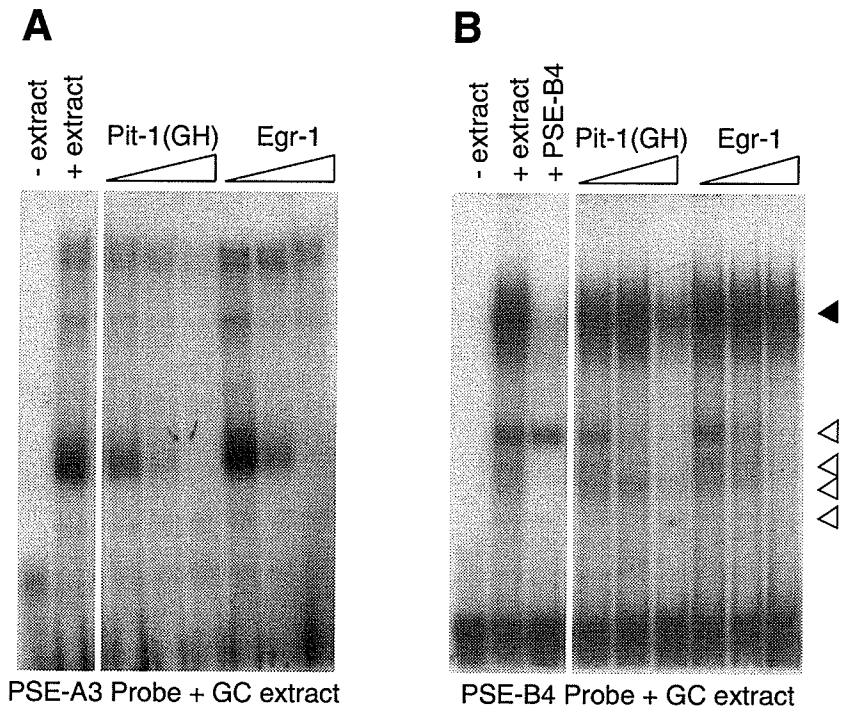


Figure 5.4: A Pit-1 oligonucleotide competes for PSE-A3 and PSE-B4 EMSA complexes.

(A) Radiolabelled PSE-A3 was used as an EMSA probe with GC nuclear extract. Competitor oligonucleotides were included in the reactions at 10-fold, 100-fold, and 1000-fold mass excess of probe. **(B)** Radiolabelled PSE-B4 was used as an EMSA probe with GC nuclear extract. Competitor oligonucleotides were included in the reactions at 10-fold, 100-fold, and 1000-fold mass excess of probe (GHpPit-1 and Egr-1) or 1000-fold mass excess of probe (PSE-B4).

5.2.1b Competition of Pit-1 binding to the CS-A promoter with an RFX element.

EMSA and nuclease protection competitions with P sequence probes, suggested that Pit-1 may associate with P sequence complexes. To address whether PSFs were connected with binding events at the CS-A promoter, competition of the CS-Ap GC nuclease protection pattern was attempted. A -281/+6 (*BamHI*/*Pst*I) fragment of the CS-A promoter was used as a nuclease protection probe with GC nuclear extract (Fig. 5.5). The proximal Pit-1 site is indicated and competed by the Pit-1(GH) oligonucleotide at 10,000-50,000 fold pmole excess of probe. Use of the EF-C/MDBP RFX element (RFX) as a competitor at 50,000-70,000 fold pmole excess of probe also resulted in competition of the Pit-1 protected region. The RFX mutant oligonucleotide (RFXm) was not a competitor at these same levels. Competition of Pit-1 binding to the CS-A promoter was not observed with PSE-Am2 (RFX element, Fig 3.11C), PSE-Am6 (NF-1 element, Fig. 3.11B), or the NF-1 consensus oligonucleotide. *These results imply that RFX proteins may associate with Pit-1 in GC cells.*

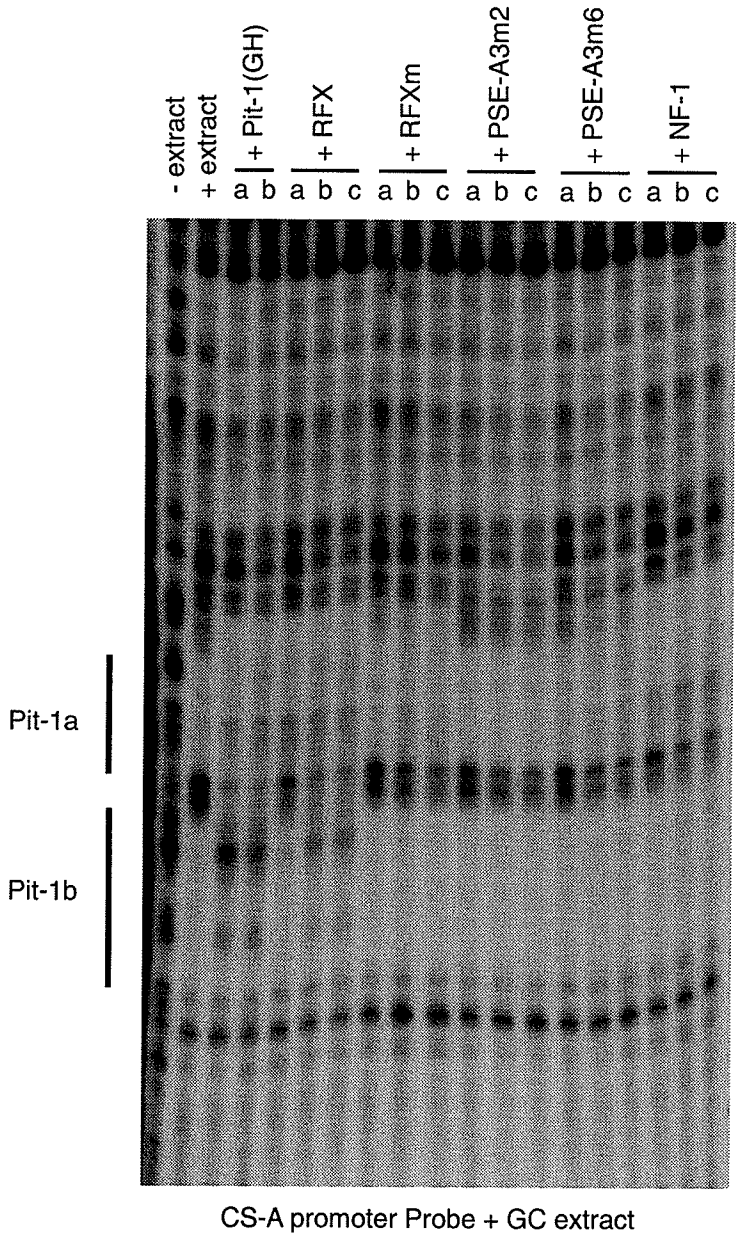


Figure 5.5: An RFX element competes for Pit-1 binding to the CS-A promoter.
A *BamH*1/*Pst*1 fragment of the CS-A promoter was radiolabelled at the 3′ end and incubated with or without pituitary GC nuclear extract before DNaseI digestion. Competitor oligonucleotides were included in the reactions at (a) 10000-fold, (b) 50000-fold, and (c) 70000-fold pmole excess of probe. The Pit-1 protected regions are indicated.

5.2.2 The capacity for PSF and Pit-1 interactions was not detected in the absence of DNA.

5.2.2a Pit-1 was not detected in immunoprecipitations of pituitary GC nuclear extracts with NF-1 or RFX1 antibodies.

The EMSA and nuclease protection competitions suggested the capacity for interactions between Pit-1 and the PSFs. The identification of NF-1 and RFX1 as PSFs (Chapters 3 and 4) meant that these interactions could be further investigated. Pituitary GC nuclear extracts were used to investigate if Pit-1/PSF interactions could be detected by co-immunoprecipitation (Fig. 5.6). Immunoprecipitations of GC nuclear extract were carried out with both NF-1 and RFX1 antibodies. The immunoprecipitated material was separated by SDS-PAGE and a western blot was probed with the Pit-1 antibody (Fig. 5.6A). Pit-1 bands of the appropriate size were detected in GC nuclear extract and 293 cell lysate where Pit-1 had been overexpressed. Pit-1 bands of the appropriate size were not detected in the NF-1 or RFX1 immunoprecipitation lanes. Western blots of these same immunoprecipitations were also probed with RFX1 (Fig. 5.6B) and NF-1 (Fig. 5.6C) antibodies and confirmed specificity of the protocol. Immunoprecipitation with the Pit-1 antibody and an attempt to detect NF-1 and/or RFX1 was also carried out (data not shown). The Pit-1 antibody did not, however, immunoprecipitate Pit-1 under these conditions. Dr. Harry Elsholz (University of Toronto) indicated in a personal communication that they have also attempted to immunoprecipitate with this and two other commercially available Pit-1 antibodies, using several conditions, and that the antibodies are not effective for immunoprecipitation. *Thus, interactions between Pit-1 and NF-1 or RFX1 were not detected in co-immunoprecipitation experiments from pituitary GC nuclear extract.*

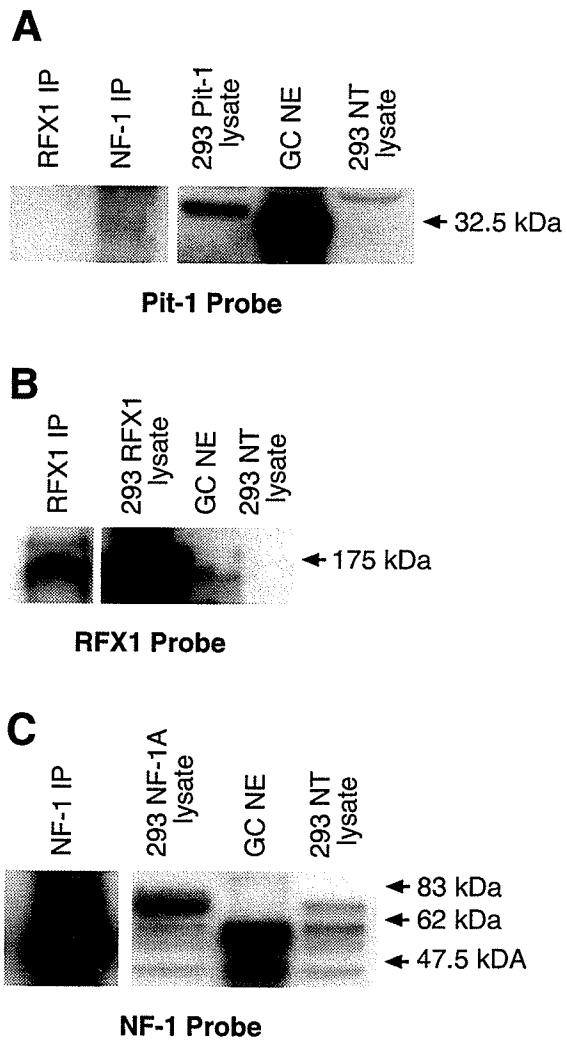


Figure 5.6: Pit-1 was not detected in the immunoprecipitations of pituitary GC nuclear extracts with NF-1 or RFX1 antibodies.
(A) RFX1 and NF-1 immunoprecipitations (IP) were separated by SDS-PAGE and a western blot was probed with Pit-1 antibodies. Included in the western blot was 30 μg of pituitary GC nuclear extract (GC NE), 293 cell lysate with overexpressed Pit-1 protein (30 μg), and non-transfected (NT) 293 lysate (30 μg). **(B)** The specificity of the RFX1 immunoprecipitation was confirmed by probing RFX1 immunoprecipitate with the RFX1 antibody. Included in the western blot was 30 μg of pituitary GC nuclear extract (GC NE), 293 cell lysate with overexpressed RFX1 protein (30 μg), and non-transfected (NT) 293 lysate (30 μg) **(C)** The specificity of the NF-1 immunoprecipitation was confirmed by probing NF-1 immunoprecipitate with the NF-1 antibody. Included in the western blot was 30 μg of pituitary GC nuclear extract (GC NE), 293 cell lysate with overexpressed NF-1A protein (30 μg), and non-transfected (NT) 293 lysate (30 μg). These blots were provided by Ms. Xiaoyang Yang.

5.2.2b Co-immunoprecipitation of Pit-1 with NF-1 or RFX1 was also not detected using overexpressed epitope-tagged proteins.

The use of epitope-tagged expression vectors for co-immunoprecipitation experiments was previously described for RFX1 and NF-1 (Chapter 3.2.2e). This system was also utilized to examine potential interactions between Pit-1 with RFX1 and the NF-1 family (Fig. 5.7). The Pit-1 cDNA was inserted into both the pCMV-HA and pCMV-Myc epitope-tagged mammalian expression vectors (Clontech). The HA-Pit-1 plasmid was co-transfected with the cmyc-NF-1 plasmids into 293 cells, and the cmyc-Pit-1 plasmid was co-transfected into 293 cells with the HA-RFX1 plasmid. Cell lysates from these transfections were run in SDS-PAGE gels and protein blots confirmed overexpression of epitope-tagged Pit-1 with RFX1 and the NF-1 family members (Fig. 5.7A-C). Immunoprecipitations were then carried out with RFX1 and c-myc antibodies. Western blots of the immunoprecipitated material were probed with Pit-1 antibody. Pit-1 was not detected in either the RFX1 or c-myc immunoprecipitations (Fig. 5.7D and E). *Thus co-immunoprecipitation experiments with overexpressed epitope tagged proteins did not detect Pit-1 interactions with NF-1 or RFX1.*

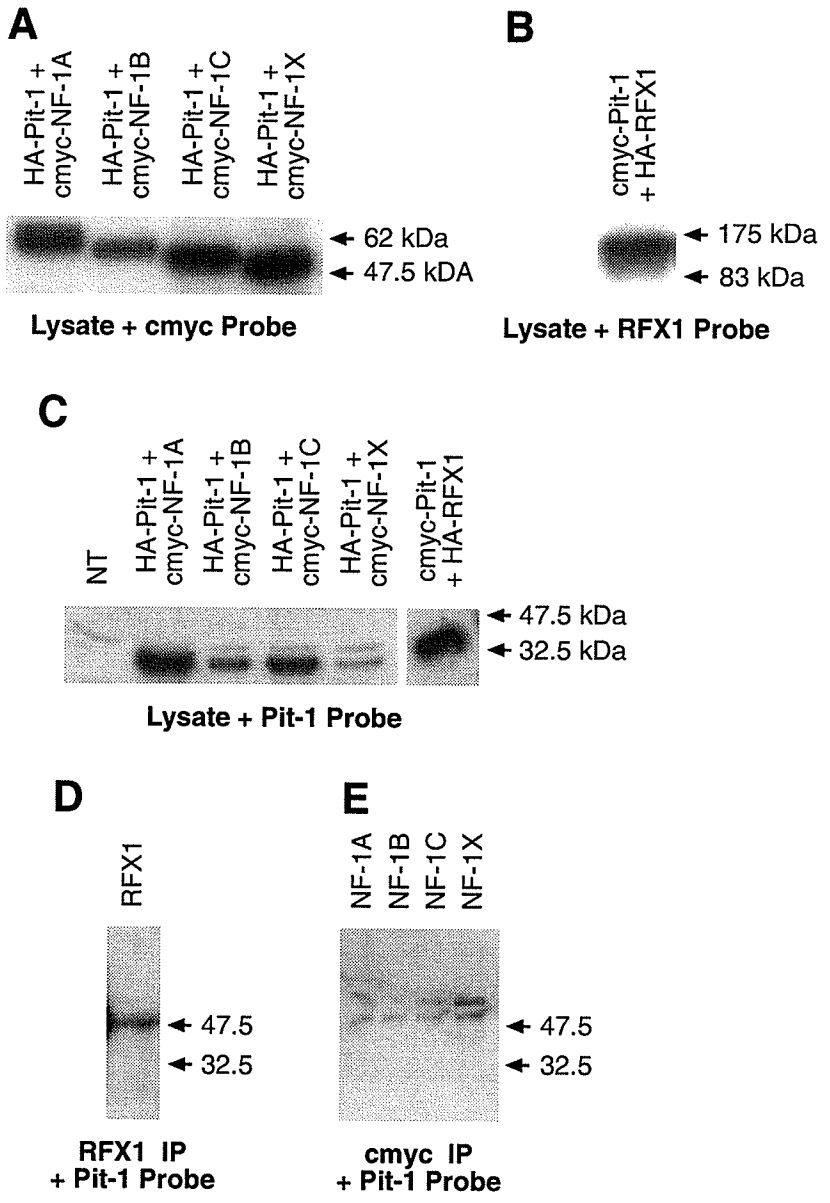


Figure 5.7: Co-immunoprecipitation of Pit-1 with NF-1 or RFX1 was not detected using overexpressed epitope-tagged proteins.

(A) Transiently transfected JAR lysates (50 μg) were separated by SDS-PAGE and transferred blots were probed with c-myc antibody to confirm overexpression of the epitope-tagged NF-1 family members. The approximate sizes of the NF-1 proteins were 60kDa (NF-1A), 57 kDa (NF-1B), 54 kDa (NF-1C), and 50 kDa (NF-1X). **(B)** Transiently transfected JAR lysate (50 μg) was separated by SDS-PAGE and the transferred blot was probed with RFX1 specific antibody to confirm overexpression of RFX1. **(C)** Transiently transfected JAR lysates (50 μg) were separated by SDS-PAGE and transferred blots were probed with Pit-1 antibody to confirm overexpression of the epitope-tagged Pit-1. Approximate size of Pit-1 is 33 kDa. (NT) non-transfected. **(D)** Probing the RFX1 immunoprecipitation with Pit-1 antibody did not reveal the presence of

Pit-1. **(E)** Probing the c-myc/NF-1 immunoprecipitations with Pit-1 antibody did not reveal the presence of Pit-1. Immunoprecipitations and western data provided by Ms. Xiaoyang Yang.

SUMMARY OF CHAPTER 5 RESULTS

- Replacement of the CS-A promoter with the minimal Tk promoter resulted in 263P function switching from repressor to enhancer activity. The 263P fragment was an enhancer of TkpLuc activity in pituitary GC, placental JEG-3, cervical carcinoma HeLa, and embryonic kidney 293 cells. These data imply that the repressor activity of 263P involves the CS-A promoter complex. The 'switch' in the direction of 263P when the promoter is replaced is an important observation, as it implies that it may be possible to change the function of P sequences from a repressor in the pituitary to an enhancer in the placenta, through altering the complexes that assemble at the promoters in the respective tissues.

- In pituitary cells, multiple complexes are capable of associating with the CS-A promoter (reviewed in Chapter 1.2.1). The main transactivator of transfected CS-A promoter activity in pituitary GC cells is the Pit-1 complex. A link between 263P repressor activity in GC cells and Pit-1 mediated activation of the CS-A promoter was established. A mutation was created in the Pit-1 binding site of the CS-A promoter (CSpPit-1m.Luc). The activity of this construct in GC cells is considered to represent the basal (or non-Pit-1-mediated) CS-A promoter activity. 263P was unable to repress the activity of CSpPit-1m.Luc, thus displaying a requirement for the Pit-1 promoter complex for its ability to repress CS-A promoter activity in GC cells.

- The potential for interactions between Pit-1 and the PSFs that were suggested in the original characterization of 263P (133) were reassessed using EMSA and nuclease protection experiments. These results suggest that interactions between Pit-1 with RFX1 and the NF-1 family are possible. An extension of the original analysis is the ability to compete Pit-1 binding from a CS-A promoter fragment with the EF-C/MDBP RFX element. The inability to compete Pit-1 binding with the PSE-Am6 RFX element may reflect a difference in the affinity of RFX proteins between these two elements.

- Co-immunoprecipitation experiments were carried out to further investigate if interactions could be detected between Pit-1, RFX1, and the NF-1 family. In contrast to the EMSA and nuclease protection experiments no indication of Pit-1/RFX1 or Pit-1/NF-1 interactions were observed. It is possible that the interactions occur but are not sufficiently stable to be detected by this technique. It is also possible that some degree of stability may be conferred by the inclusion of DNA in the EMSA and nuclease protection competition experiments. A discussion of the results presented in this chapter can be found in Chapter 6.2.1.

CHAPTER 6

DISCUSSION

6.1 P sequence complexes

6.1.1 Interpretation of the *in vitro* and *in vivo* complexes.

The P sequence elements are located approximately 2 kb upstream of each of the placentally expressed genes of the human GH/CS locus. The conservation and location of these elements led to the hypothesis that they were involved in the differential expression pattern of this gene family (83). Subsequently, 263P was demonstrated to repress the activity of the CS-A promoter in pituitary cells (133), and confer placental enhancement in transgenic mice (150). The data presented in this thesis identify NF-1 and RFX1 as factors that associate with P sequences *in vitro* and *in situ*.

6.1.1a The pituitary P sequence complex *in vitro*.

One of the research objectives of this thesis was to identify candidate P sequence factors (PSFs) in pituitary GC cells. 263P has a characteristic nuclease protection pattern that is observed with pituitary GC as well as other nuclear extracts, such as HeLa (human cervical carcinoma), C6 (rat glioma), and U-87 (human brain glioblastoma/astrocytoma) (133) and data not shown). The regions that are protected in 263P are PSE-A and PSE-B, with either element functioning as an efficient competitor of both protected regions (Fig. 3.3). Competition experiments such as these, led to the prediction that the nuclease protection pattern of 263P resulted from one of three potential complexes (Fig. 6.1): (a) a single protein that contacted two regions of the DNA, (b) related proteins that recognize similar sequences, or (c) two different proteins that interact with each other to form a more stable complex, such that disruption of one binding event also disrupts the second binding event.

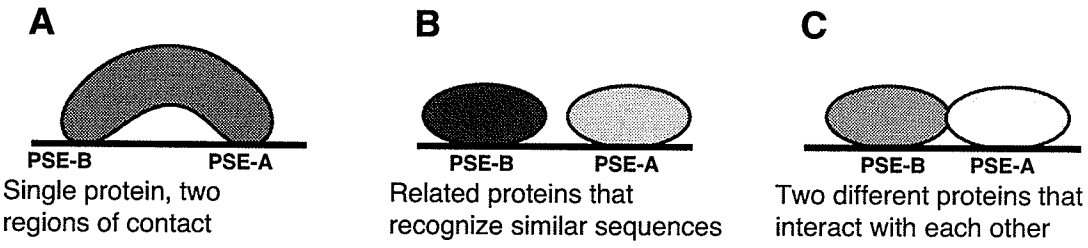


Figure 6.1: Potential PSF complexes based on nuclease protection data.
Based on competition of the 263P nuclease protection pattern with either PSE-A3 or PSE-B4, the pattern is interpreted to represent one of three potential PSF complexes.

Assessment of PSE-A and PSE-B in isolation, demonstrated that PSE-A is a composite element capable of mutually exclusive NF-1 and RFX1 binding (Chapter 3.2.1), and that PSE-B is an NF-1 element (Chapter 3.2.2). From these data, the nuclease protection pattern of 263P may be interpreted as representing the third of the three potential complexes, where different proteins that interact with each other associate with PSE-A and PSE-B (Fig. 6.2).

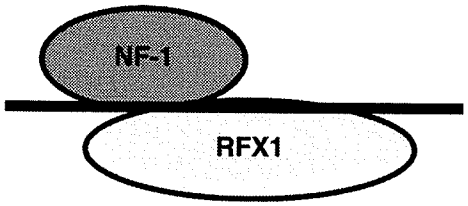


Figure 6.2: Model of a potential *in vitro* P sequence complex.
The demonstration of RFX1 association with PSE-A and NF-1 association with PSE-B allows for this model of the *in vitro* P sequence complex to be proposed.

The interpretation in Figure 6.2 is further supported by the demonstration of interactions between RFX1 and the NF-1 family. These interactions appear to occur through the DNA binding domain of NF-1 proteins (Chapter 3.2.2e), and have the capacity to form even in the presence of DNA elements (Fig. 3.20). Despite this, the participation of RFX1 as a component of the 263P complex *in vitro* was not confirmed by assessment of DNA-

protein interactions, as RFX elements did not compete for 263P protein binding in either EMSA or nuclease protection assays (Fig. 3.28). These additional results call into question the interpretation of the *in vitro* P sequence complex that is shown in Figure 6.2.

In contrast to the RFX element, NF-1 binding sites were efficient competitors of the entire 263P nuclease protection pattern and EMSA complexes (Fig. 3.27). Furthermore, competition of the 263P nuclease protection pattern by PSE-A was shown not to require the RFX binding site, instead, appearing to involve NF-1 half sites within the PSE-A3 sequence (Fig. 3.29). Thus, the 263P element structural data, as assessed by DNA-protein interactions, strongly support that the nuclease protection pattern represents the second of the three predicted complexes, and that NF-1 proteins are the 263P PSFs *in vitro* (Fig. 6.3).

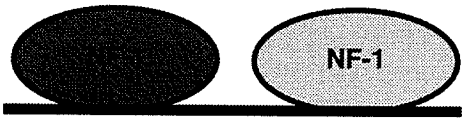


Figure 6.3: Model of a second potential *in vitro* P sequence complex.
The structural data in context of the 263P element strongly support the interpretation of the *in vitro* P sequence complex as an NF-1 complex, without evidence of RFX1 participation.

While Figure 6.3 agrees with the structural data, the functional data presented in Chapter 3 expand this model to an interpretation of at least two separate complexes that have the capacity to form on 263P *in vitro*. These include an NF-1 repressor complex such as the one shown in Figure 6.3, and a non-functional complex containing RFX1. There are several observations that support this interpretation: (i) only NF-1 association with 263P could be demonstrated (Fig. 3.27), (ii) RFX1 associated with 103P (Fig. 3.25), which did not repress CS-A promoter activity (Fig. 3.26), and (iii) overexpression of NF-1 in GC cells resulted in a decrease in co-transfected 263PCSp.Luc activity. Thus, there are two components to the interpretation of the pituitary *in vitro* complex. Firstly, taken together,

the results indicate that the 263P repressor complex is an NF-1 complex, likely represented by the schematic shown in Figure 6.3. Secondly, these data also suggest that the NF-1 complex is not the only P sequence complex that exists *in vitro*. The decrease in 263PCSp.Luc activity upon NF-1 overexpression demonstrates that repression of CS-A promoter activity can be increased. This was interpreted to occur through increasing the amount of NF-1 in the cell, with the assumption that 263PCSp.Luc was not fully occupied by NF-1 repressor complex. Mutually exclusive binding of NF-1 and RFX1 to PSE-A3 (Fig. 3.11) indicates that in the absence of NF-1 occupation, RFX1 may associate with 263PCSp.Luc at PSE-A. This RFX1 complex would not function as efficiently in the repression of CS-A promoter activity, supported by the lack of function observed with the 103P element (Fig. 3.26).

In terms of structure, the non-functional RFX1 complex may resemble the schematic shown in Figure 6.2. A second possibility for the RFX1 complex is that the RFX1 does not associate with the DNA directly, but through protein-protein interactions with NF-1. This is supported by the ability to detect interactions between RFX1 and the NF-1 family (Fig. 3.21). In this possible model, the association of RFX1 would inhibit the ability of the NF-1 to function as a repressor. Alternatively, the RFX1 complex could contain RFX1 alone, based on the inability to detect NF-1 association with 103P (Fig. 3.25).

Therefore, taking the data presented in Chapter 3 as a whole, the model for the 263P *in vitro* complex is a composite of at least two separate complexes. As previously mentioned, however, the association of RFX1 with 263P was not detected in structural assays. How then is this apparent discrepancy between the structural data and the dual complex model explained? The capacity for two complexes to exist on 263P *in vitro* suggests that the nuclease protection pattern is a composite of these two complexes. The presence of composite nuclease protection patterns has been previously demonstrated for the GH-N promoter, where Pit-1 and Sp1 binding sites overlap (119, 217). In the scenario of a composite 263P nuclease protection pattern, competing for RFX1 with the RFX element would not show a disruption of the complex, because in the place of RFX1, NF-1 could associate with PSE-A. This then explains the inability to compete with RFX

elements (Fig. 3.28). This scenario would also imply that the reverse competition would produce the same result, and that competition of the protected regions with NF-1 elements would leave a region of PSE-A, representing RFX1 association, protected. The results do not support this, as competition with the NF-1 elements does not result in a 'RFX1 footprint' (Fig. 3.27A). It is possible, however, that the association of RFX1 with 263P is either not stable or predominant enough to be detected under the *in vitro* assay condition used, and the negative result must therefore not be over-interpreted.

6.1.1b The pituitary P sequence complex *in situ*.

A second research objective of this thesis was to assess the potential for candidate PSFs identified *in vitro* to associate with P sequences *in situ*. To investigate this possibility, a chromatin immunoprecipitation (ChIP) technique with human pituitary tissue samples was developed. Previously, the histone hyperacetylation pattern of the GH/CS locus was documented for a human GH-secreting pituitary adenoma tissue sample and the pituitaries of double transgenic mice for a large fragment of the GH/CS locus and GH releasing hormone overexpression that resulted in pituitaries enriched in somatotrophs (129). In development of the ChIP protocol, analysis of chromatin modifications of the GH/CS locus was extended to normal post-mortem human pituitary samples. This analysis confirmed hyperacetylation of the HS I/II region, which was then used as a marker of chromatin integrity for subsequent experiments.

An added advantage to the data presented in Chapter 4 over previous reports of GH/CS chromatin status (129, 150) was the use of PCR for analysis of the ChIP chromatin. Through the use of specific primers, highly homologous sequences such as the GH and CS promoters were distinguished. This allowed for the observation that the GH, but not the CS promoters were hyperacetylated in human pituitary chromatin (Fig. 4.8). These results are consistent with the role of hyperacetylated sequences in gene regulation, as they mirror the expression pattern of these genes *in vivo*.

As well as a lack of CS promoter hyperacetylation, hyperacetylation of histone H4 was not evident for the P sequences in the human pituitary samples (Fig. 4.8). This lack of

hyperacetylation correlated with the observation of NF-1 and RFX1 association (Fig. 4.9). Therefore, not only do NF-1 and RFX1 play a role in the formation and function of P sequence complexes *in vitro*, but they are candidates for regulation of the GH/CS genes by P sequences in the pituitary as well. The inverse relationship between NF-1 and RFX1 association and 263P and CS promoter hyperacetylation suggests that the pituitary repressor mechanism may involve blockage of the hyperacetylation that is observed in placental chromatin. This potential mechanism will be further discussed in section 6.2. The implications of NF-1 and RFX1 involvement in regulation are further discussed in sections 6.1.2 (NF-1) and 6.1.3 (RFX1).

The ChIP assay results do not resolve whether the two complexes that were interpreted to occur *in vitro* also occur *in vivo*. These assays do not distinguish between factors that participate in a common complex from those forming mutually exclusive complexes in different cells from the same tissue sample. Given the data presented, an *in vivo* model of the P sequence complex that resembles Figure 6.2 is just as likely as an *in vivo* model that is a composite of the complexes in Figures 6.2 and 6.3. To this point, a technique is not currently available that will differentiate between these possibilities for the analysis of P sequence complex(s) *in situ*.

6.1.1c The placental P sequence complex.

The data presented in Chapter 3 provide the first evidence of significant P sequence function in placental cell lines (Fig. 3.2). When 263P was analyzed by nuclease protection with choriocarcinoma JAR nuclear proteins, a pattern that resembled that of pituitary GC nuclear extracts was observed (Fig. 3.4). This result indicated that PSE-A and PSE-B may be useful elements for the identification of candidate placental P sequence factors *in vitro*. The use of PSE-A3 and PSE-B4 supported RFX1 and NF-1 as candidate PSFs (Figs. 3.13 and 3.23). In the pituitary cells, the *in vitro* repressor complex was interpreted to be an NF-1 complex (Fig. 6.3). A second and non-functional RFX1-containing complex was also supported by the *in vitro* data. Based on these interpretations, and the known participation of both RFX1 and the NF-1 family in both repressor and activator functions, it can be hypothesised that the 'switch' from pituitary

repressor activity to placental enhancer activity could be achieved through altering the specific members of these families that are involved in each complex. The enhancer activity of 263P in transgenic mice was specific to placental tissues (150), thus the availability and expression of the NF-1 and RFX families in this tissue type may be an important aspect in the activity of this element *in vivo*.

Despite the enhancer activity of 263P in transgenic mice, the direction of the activity in the human placental cell lines was as a repressor. An explanation for this discrepancy could be that the cell lines were human in origin and the transgenic placenta was murine. It is noted, however, that the tissue restricted expression of the GH/CS locus is mirrored in transgenic mice (122). In addition, placental expression of other human genes, such as CYP19 can be accurately reproduced in transgenic mice, despite the lack of an endogenous murine gene (218, 219). This suggests that the mouse placenta is an accurate model to reflect human placenta gene expression. The explanation for the discrepancy between the *in vitro* and *in vivo* functions likely resides in the context of the two assays. The function of the 263P in transgenic mice as an enhancer correlates with the observation of histone hyperacetylation in human term placenta chromatin, which often reflects regions of enhancer activity. The *in vitro* gene transfer assay is not designed to detect such processes. Thus, either the protein binding events at 263P in placental cells *in vitro* do not reflect the binding events that occur *in vivo*, or the same binding events occur, but the ability to detect the appropriate functional effect *in vitro* is not permitted by the experimental design. With the second of these two options as a possibility, the candidate placental PSFs from the *in vitro* structural data were directly evaluated *in situ* using ChIP assays.

In the pituitary, the *in vitro* and *in situ* data correlated. In contrast, when human term placenta chromatin was used for ChIP assays, neither RFX1 nor NF-1 association was confirmed (Fig. 4.10). There are three possible explanations for this result. The first is that RFX1 and/or NF-1 associate with P sequences in human placenta, but that some aspect of the tissue and/or technique compromised the ability to detect these interactions. The second possibility is that different members of the RFX family, such as RFX2 or

RFX3 are the PSFs in human placental chromatin. A final possibility, is that the factors which associate with P sequences in the placenta are not related to the RFX or NF-1 families, and the PSFs in human placental chromatin are separate factors altogether. In summary, the composition of the placental P sequence enhancer complex was not determined from the assays that were used.

As the placental *in vitro* complex does not represent the *in vivo* complex in either function or composition, the development of an experimental system that demonstrates placental enhancer activity is required for further investigation of potential placental PSFs. An approach for the development of such a system as well as a more detailed discussion of possible experiments to identify the placental PSFs can be found in section 6.4.2b.

Although the ChIP assays with human term placental chromatin did not identify P sequence complex participants, development of this technique contributed information regarding chromatin modifications of the GH/CS locus in this tissue. A previous report of placental histone hyperacetylation patterns in the locus used a mixture of antibodies to hyperacetylated histone H3 and hyperacetylated histone H4 (150). According to the histone code hypothesis, nucleosomal modifications confer a specific set of instructions for gene expression based on different combinations of modifications (50), much as the different combinations of nucleotides in DNA sequences code for different information. In this study, antibodies to hyperacetylated histone H4 were used in isolation. These data demonstrated a pattern of hyperacetylation that differed from the previously reported combined antibody approach, and likely contributes to explaining the observed 1.8-fold hyperacetylation of 263P (Fig. 4.7) compared to the 17-fold hyperacetylation originally reported (150). This observation, combined with similar hyperacetylation ratios for the CS promoters and downstream enhancer regions, suggests that in the placenta, multiple regions proximal to the genes of the locus may cooperate in the regulation of their expression. In addition to these differences, HS IV hyperacetylation was also observed, an observation previously not reported (150). A potential mechanism for placental

regulation of the GH/CS genes based on these and other data is discussed further in section 6.3.

6.1.2 The implications of NF-1 involvement.

NF-1 was initially identified as a factor involved in the replication of adenoviral DNA (220, 221). Subsequently, it was determined that the transcription factor CTF-1 and NF-1 were one and the same (192). This was one of the first observations that a factor could be involved in both DNA replication and transcription. The dual function of NF-1 provides some clues to its role in these processes, as in both cases the recruitment of a polymerase and the remodeling of DNA structure are key events. In the activation of adenoviral transcription, NF-1 is involved in both recruitment and remodeling. NF-1 interactions with adenoviral DNA polymerase are essential for formation of the pre-initiation complex (PIC) (222-224). NF-1 also induces changes to the structure of the adenoviral DNA template (225, 226), resulting in an increase in DNA polymerase association and the efficiency of replication. This structural role in PIC assembly can be viewed as somewhat analogous to the bending of the DNA template by TATA-binding protein (TBP) that enables the association of RNA polymerase II and initiation of transcription (227). From this information, it is possible to speculate that the role of NF-1 in transcription may involve mechanisms similar to its participation in adenoviral replication. The polymerase for each process is a different molecule, however, NF-1 may still be involved in recruitment of the pre-initiation complex. Likewise, the template for these two processes differs, as the adenoviral DNA is free of nucleosome structures. Nonetheless, modifications of these chromatin units by other factors may permit the function of NF-1 in transcription to resemble its structural role in adenoviral replication.

The identification that NF-1 interacts with P sequences, and may play a role in regulating expression of the GH/CS family, is one of many instances where NF-1 has been implicated in transcriptional regulation. Since the identification that NF-1 was a transcription factor as well as a replication factor, NF-1 has been implicated in the

expression of more than 100 cellular and viral genes (181). NF-1 is one of several transcription factors, such as the zinc-finger Sp1, that are frequently found to be involved in the regulation of gene expression. Despite the frequency with which NF-1 is found in regulatory units, the result of NF-1 involvement is not always the same, as NF-1 is both an activator and repressor of transcription (153, 182, 192, 196-202, 228). Additionally, in spite of its common usage and ubiquitous expression pattern, NF-1 regulates the expression of several genes that are restricted in their expression patterns. This regulation can be spatial (expression in one cell type and not another) (182, 192, 200, 201, 228) or temporal (expression at a certain stage of cell development, or in response to environmental signals) (196, 197, 229-231). The data presented in this thesis extend the list of genes that involve NF-1 in regulation of their differential gene expression patterns. The underlying question with NF-1 involvement in transcriptional regulation, is how a ubiquitous factor utilized in many systems, results in such a wide array of specific responses.

6.1.2a Diversity of the NF-1 family.

As discussed in Chapter 3, NF-1 is not a single factor, but a family of transcription factors. This diversity is the reason that NF-1 generates specific functions such as activation and repression in regulation of differential gene expression patterns. The P sequences are hypothesized to activate expression of the GH/CS genes in the placenta and to repress the expression of these genes in the pituitary, thereby participating in the differential gene expression pattern of this family. The diversity of the NF-1 family and how this contributes to the different activities that can be achieved by NF-1, is therefore particularly relevant to the consideration of NF-1 as a P sequence complex participant.

There is a single NF-1 gene in the genomes of both *Caenorhabditis elegans* and *Drosophila melanogaster* (181, 232). In higher organisms, however, gene duplication events have led to four NF-1 genes, NF-1A, NF-1B, NF-1C, and NF-1X (193, 194, 232). The origin of the NF-1 family in higher organisms by gene duplication events is supported by their linkage with the Jun family in both humans and mice (195), and conservation of their DNA binding domains (181). This highly conserved DNA binding

domain is what categorizes NF-1 proteins into a family of transcription factors, as it is not related to other known DNA binding domains, such as zinc fingers, leucine zippers, or homeodomains (153).

The NF-1 DNA binding domain is contained within the NH_2-terminal 200-220 amino acids of the protein. This region is not only responsible for DNA binding, but is the domain that is necessary for adenoviral replication, nuclear localization, and dimerization of NF-1 molecules (153, 233, 234). Homology of the NH_2-terminus is approximately 90% between the four NF-1 genes from humans, mice, and chickens (181). As was previously discussed in Chapter 3, however, the NF-1 family members can vary in their affinities for a particular NF-1 site (170, 186, 187), despite homology of the DNA binding domain. The different affinities of the NF-1 proteins for NF-1 DNA elements result from COOH-terminal region influences on the NH_2-terminal DNA binding domain (202, 203, 235).

The COOH-terminal region is where the diversity between NF-1 family members occurs (181). This diversity is the result of differences between the sequences of the four genes, as well as alternative splicing of NF-1 transcripts (170, 202, 203, 235-238). Although common structural features exist, such as proline-rich activation domains (153, 202, 239), there are also components that are particular for individual family members. For example NF-1A (198, 239) and NF-1X (202, 240) contain distinctly different repressor domains in their COOH-terminal regions, and NF-1C variants can contain a single copy of the RNA pol II COOH-terminal domain (CTD) repeat (241-244) that is not contained in other members of the NF-1 family. The diversity of the NF-1 family is further compounded by the formation of hetero- and homo-dimers (153, 242, 245). NF-1 complexes require dimerization to associate with DNA (222), and both DNA binding domains are required for that association (233). The capacity for different NF-1 variants, combined with multiple possiblilities of NF-1 dimerization partners, results in the diversity of function that is possible within this family of transcription factors.

Although NF-1 is ubiquitously expressed, this statement applies to the family, in that there is NF-1 found in all cell types. The relative expression of the different genes and their variants varies between cell types (170, 199, 202, 203, 246, 247), in response to environmental stimuli (196, 197, 248-250), and can be affected by growth and differentiation (202, 247, 250, 251). The generation of NF-1 complexes that are involved in cell-specific functions can result from variations in the relative expression of different NF-1 family members. Thus, it is possible that in the case of the GH/CS family, there are particular members or variants of the NF-1 family that are responsible for pituitary repressor activity, while other members of the NF-1 family may contribute to placental activation.

6.1.2b The mechanisms of NF-1 activity

The gene duplications that led to the generation of the NF-1 family resulted in conserved structures such as the highly homologous DNA binding domains and proline-rich activation domains (153, 181, 202, 239). The activator and repressor functions of NF-1, however, are not as redundant as the conserved structures may originally imply. Perhaps the best demonstrations of the specificity of NF-1 function are the phenotypes of the NF-1 knockout mice. In a knockout model of the NF-1A gene, the expression of the three remaining NF-1 genes does not compensate for the loss of NF-1A, and the phenotype consists primarily of severe abnormalities in neural development (252). In contrast, the NF-1B knockout model phenotype displays abnormalities in lung development (253).

All four NF-1 genes have the capacity to activate transcription through the COOH-terminal region (254). The diversity of the NF-1 family, and the non-redundancy of NF-1 function, suggests that NF-1 family members activate transcription through different mechanisms. An example has been documented with NF-1C variants. Certain NF-1C variants utilize a domain containing a single copy of the RNA polymerase II CTD repeat in transcriptional activation (241-244). This domain, however, is absent from other NF-1C variants, which retain the capacity to activate transcription (235). The repressor activity of NF-1 also appears to involve different mechanisms, depending on the NF-1 family member involved. Repressor domains in NF-1A and NF-1X proteins are not

homologous (198, 202, 239, 240) and, therefore, are more likely to involve different than similar mechanisms. As it is unclear from this thesis which NF-1 family member(s) may be involved in P sequence function, a number of documented NF-1 mechanisms are potential explanations for activation and repression through P sequences.

There are two general mechanisms through which NF-1 activator and repressor activities are achieved. Both of these general mechanisms have potential relevance for P sequence function *in vitro* and *in vivo*. The first mechanism involves recruitment of other factors/complexes, such as basal transcription components (198, 244, 255, 256), specific co-activators (257), and specific co-repressors (199). A second mechanism by which NF-1 activities are achieved uses NF-1 binding to block the association of other factors (200, 258-261).

6.1.2c NF1 is a context dependant regulator.

The NF-1 family member that predominantly associates with an NF-1 site determines the function of that element. It is the COOH-terminal region of NF-1 that influences the affinity of the DNA binding domain for a particular NF-1 site. In addition to this, interactions with surrounding factors may stabilize NF-1 association, which suggests that the context of an NF-1 element has the capacity to influence the function of an NF-1 element.

A functional effect of altering NF-1 binding site sequences was demonstrated by creating mutations in PSE-B4 that did not affect the ability of the NF-1 family to associate with PSE-B, but did result in changes in the functional activity of the elements (Figs. 3.17 and 3.18). These observations suggested that the binding site itself was involved in dictating the function of an NF-1 element, presumably by affecting the predominant NF-1 family member that associated with a particular NF-1 site. This explanation was supported by the fact that alterations in an NF-1 site had previously been demonstrated to affect the predominant NF-1 family member that associated *in vitro* (170). There is also evidence that not only does the sequence of the NF-1 site influence function, but that the number of NF-1 sites will impact the mechanism of activity. For example, the RNA polymerase

CTD motif of NF-1C variants is not required for the transactivation from a reporter containing a single NF-1 site, but is involved in transactivation of a reporter containing six NF1 binding sites (235). This suggests that the number of NF-1 sites in a regulatory element can dictate the mechanism by which that element exerts its effects. With regard to P sequence function, this may be a relevant consideration. There are two NF-1 sites in 263P (at PSE-A and PSE-B), whereas PSE-B4 contains a single NF-1 site. While 263P is an activator of Tkp.Luc activity (Fig. 5.1), PSE-B4 and the consensus NF-1 site have no functional effect (151). Therefore, the documented ability of context to influence NF-1 function correlates with observations that P sequence function was influenced by the context of the element.

The influence of the context of NF-1 elements also involves the impact of surrounding factors. With regard to P sequences, when the surrounding factors or complexes are modified, as was the case with the Pit-1 mutation of the CS-A promoter, the NF-1-mediated repressor effect of 263P was eliminated (Fig. 5.2). Alterations in promoter complexes that influence the direction of NF-1-mediated function have been documented elsewhere as well. The distal regulatory region of the human Pit-1 gene contains two NF-1 sites (182). This region enhances the activity of the heterologous SV40 promoter, while reducing the activity of the homologous Pit-1 promoter in pituitary cells (182). This example is particularly relevant for two reasons. First, it supports the interpretation of Figure 5.2, that a specific promoter complex can be involved in determining the direction of NF-1 activity. Additionally, this observation is a second example of an NF-1 repressor mechanism that appears to involve a Pit-1 promoter complex, as the promoter of the Pit-1 gene contains Pit-1 binding sites (110, 262). The ability of the related POU domain protein Oct-1 to stabilize NF-1 binding has been demonstrated for a separate system (263), and may indicate a potential mechanism for Pit-1/NF-1 mediated repressor effects.

The cellular environment will also influence NF-1 function. Cysteine residues in the NF-1 DNA binding domain (264) and the transactivation domain (265, 266) can be oxidized. Under certain conditions, oxidation will inhibit NF-1 association with NF-1 sites (264, 267), while separate conditions do not alter DNA binding, but repress transactivation

(265, 266). Oxidation of NF-1 has thus been proposed to provide a rapid and reversible method for regulating NF-1 activity from a variety of cellular conditions and stimuli (268, 269). In addition, signalling will also affect NF-1 activity. For example, transforming growth factor-β (TGF-β) signalling and an associated increase in cytosolic calcium was shown to increase transactivation by NF-1 through phosphorylation of a TGF-β responsive domain (249). Phosphorylation of NF-1 has also been demonstrated after insulin stimulation (197).

6.1.3 The implications of RFX involvement in P sequence complexes.

6.1.3a The RFX family of transcription factors.

Like the NF-1 proteins, RFX1 belongs to a family of transcription factors characterized by a common DNA binding domain. The RFX family includes five members in vertebrates, RFX1-RFX5, which are each the product of a separate gene (155, 270-275). A single member of the RFX family in each of the following non-vertebrate species has also been identified; *Saccharomyces cerevisiae* (*Crt1*) (276), *Schizosaccharomyces pombe* (*Sak1*) (277), the fungus *Acremonium chrysogenum* (*CPCR1*) (278), *Caenorhabditis elegans* (*DAF-19*) (279), and *Drosophila melanogaster* (*dRFX*) (280).

6.1.3b DNA binding in the RFX family.

The amino acid sequence of the RFX family DNA binding domain is conserved 76-96% (188). This domain was originally believed to be unique (274), however, subsequent characterization demonstrated that it consisted of a winged-helix structure, and this catagorized the RFX family as members of the winged-helix HNF-3/forkhead subfamily of the helix-turn-helix classification (189). The uniqueness of the RFX family, is that unlike other members of this subclass, they contain only one 'wing' in their DNA binding domain (189). The consensus DNA binding site of the RFX family is referred to as an X-box, but shows considerable variability in the nucleotides required for binding (161, 190). Although the DNA binding domains are highly conserved, different members of the family will have variable degrees of affinity for the range of recognized RFX sites (155,

281). These differences in affinity are likely due to diversity outside of the DNA binding domains (188), which is supported by the observations that RFX5, the most divergent vertebrate member of the RFX family (275), requires interactions with other proteins to associate with DNA (282-287). Although RFX1-3 do not require other factors for DNA association, it is possible, based on the observations of RFX5, that interactions with other factors could influence the stability of their association. In the pituitary chromatin environment, where both RFX1 and NF-1 association with P sequences was detected (Fig. 4.9), the capacity for NF-1 and RFX1 to interact (Figs. 3.21 and 3.22) could contribute to the formation and stability of a common complex.

In vitro, RFX1 associates with DNA in a 1:1 and 2:1 stoichiometric ratio (161), which was interpreted to represent monomer and dimer formation (161, 274). Many classes of transcription factors require dimerization for association with DNA sites, however, in the case of RFX1, the DNA binding and dimerization domains are clearly separable (274). Nonetheless, RFX1 dimers have three to four orders of magnitude greater stability than RFX1 monomers (161). When the structure of RFX1 with its cognate binding site was assessed, there was no evidence for contact between the two RFX1 molecules (189). It was suggested that the increased affinity of 'dimerized' RFX1 complexes did not result from protein-protein interactions, but through DNA-protein effects, as the geometry of the DNA was altered by association of RFX1 (189). From this perspective, dimerization of RFX1 does not occur, however, the binding of two RFX molecules will induce more favourable structural changes than the association of a single RFX molecule. In terms of relevance to the P sequence complex, the RFX binding site in PSE-A is a half-site, which would be expected to bind a single molecule of RFX1. Based on observations of other 'monomers' compared to 'dimers' (161, 288, 289), this structure would be less stable than a complete X box site. An important consideration is that these interpretations are based on 'naked DNA' fragments used for the *in vitro* assays. The presence of nucleosomes induces significant structural changes to the geometry of the DNA (8). These induced structural changes may have the capacity to increase the affinity of RFX1 association, as is seen *in vitro* when two molecules are binding cooperatively, and thus may help to explain why RFX1 association was observed *in situ*, but not *in vitro*.

6.1.3c RFX activity as an activator and repressor of transcription.

A variety of target genes for the RFX family have been identified that implicate them in a number of cellular processes. These include the DNA damage and replication checkpoint in *S. cerevisiae* (276), the mitotic cell cycle in *S. pombe* (277), and control of the immune response in vertebrates (275). RFX sites have also been observed in the regulation of a number of viral promoter and enhancer sequences (155, 171, 288, 290). In these processes, the role of RFX1 has been as both activator and repressor. This capacity is consistent with the hypothesis that P sequences are involved in for both pituitary repression and placental activation.

6.1.4 The possibility of additional PSFs.

Due to the observation that RFX1 and NF-1 associated with PSE-A3 in a mutually exclusive manner, it was questioned whether the presence of NF-1 binding at PSE-B may affect the factor associating with PSE-A. The PSE-A region was therefore assessed in P sequence fragments that contained both PSE-A and PSE-B. When the 103P element was analyzed by EMSA, no evidence of NF-1 association was detected, despite inclusion of the PSE-B site. As was seen with PSE-A3, the binding of RFX1 to 103P was observed. In contrast, the *in vitro* binding assays demonstrated NF-1 and not RFX1 association with the 263P fragment. From these and other results, a model was presented in the Chapter 3 summary and discussed in section 6.1.1a of two separate complexes that had the capacity to form on P sequences in pituitary GC cells, an NF-1 repressor complex, and a non-functional RFX1 complex. A question arises from this model, how is the RFX1 complex of 103P converted to the NF-1 complex observed with 263P? One possibility is the hypothesis that 263P sequences, which are not contained in 103P, contribute to the association of additional factor(s) that favour formation of the NF-1 complex over the RFX1 complex. To assess the possibility of candidate PSFs that would contribute to this hypothesis, an analysis of the sequences in 263P, outside of 103P, was undertaken.

6.1.4a Candidate PSFs.

The 103P fragment spans nucleotides 127 to 229 of 263P (Fig. 3.24). Putative binding factors outside of the 103P region were determined by two parameters; (i) the presence of transcription factor binding sites in the DNA sequence and (ii) an expression pattern of the candidate factor that did not exclude pituitary cells.

The region upstream of 103P, from nucleotide 1 to nucleotide 132 of 263P (Fig. 6.4) was analyzed with the MatInspector 2.2 program and TRANSFAC 4.0 database (matrix similarity of 0.85 and a core similarity of 0.85) (165, 166). The sequence analysis was ended at nucleotide 132, which is the start of the PSE-B nuclease protection region, instead of nucleotide 127, the start of 103P, to include factors whose binding sites may span the junction of 103P and the upstream region.

5′- ^{1}TCCTACAGG ^{10}CCTGCCTGGA ^{20}GAACAGCTCA ^{30}CAGCACAGTG
^{40}CCCTCCCAG ^{50}CAGATGATGA ^{60}GTCTGGGGTG ^{70}CTAGTCCAGA
^{80}ATGCTTCAG ^{90}GAATGACGGC ^{100}AGAAAAAGGA ^{110}GCTCTGTTTC
^{120}TGCTCTGAA ^{130}AGTG -3′

Figure 6.4: The sequence of 263P upstream of the PSE-B nuclease protection region. The 5′-3′ sequence of the CS-A 263P is shown from nucleotides 1-132. The PSE-B nuclease protection region starts at nucleotide 133. The 103P fragment starts at nucleotide 127, and the nucleotides that are part of the 103P sequence are highlighted.

This analysis revealed several putative factors, some of which were excluded based on known expression patterns. The remaining factors are shown in Table 6.1. Beyond these two parameters, however, these putative factors were not considered to be excellent candidate factors that contribute to the formation of the NF-1 complex over the RFX1 complex at 263P. Presumably, if the association of these factors favoured or stabilized the NF-1 complex, this would be reflected in the 263P nuclease protection pattern, however, the region upstream of PSE-B does not appear to be protected by the presence

of additional binding factors, as the presence of additional nuclease protection regions upstream of PSE-B was not observed, in either the original characterization of 263P, which included assessment of the opposite strand (133), or the current nuclease protection experiments.

Table 6.1
Results of the database search for additional putative PSFs upstream of PSE-B.

Factor	Core Similarity	Matrix similarity	Location
AP-4	1.0	0.909	(+ 21) aaCAGCtcac
AP-4	1.0	0.868	(+ 28) caCAGCtcac
Ik2	1.0	0.928	(- 39) tgctGGGAgggc
Lyf1	1.0	0.863	(- 41) gctGGGAgg
AP-4	1.0	0.907	(+ 44) ccCAGCagat
E47	1.0	0.869	(+ 45) ccaGCAGatgatgag
C/EBPβ	0.930	0.854	(+ 70) tagtccaGTAAtgc
C/EBPβ	0.873	0.892	(+ 81) tgcttcaGGAAtga
TCF11	1.0	0.976	(- 83) GTCAttcctgaag
CREB	1.0	0.883	(+ 88) ggaaTGACggca
ATF		0.867	(+ 89) gaaTGACggcagaa
AP-1	1.0	0.895	(+ 90) aaTGACggcag

The location of the match is designated (+) or (-) to identify the strand where the match occurs.

In contrast to this, an 'extended footprint' has been observed for the sequences downstream of the PSE-A nuclease protection region. In a previous attempt to purify the PSFs, a DNA affinity column was made with PSE-B sequences. Enrichment for the then unknown PSF-B resulted in an 'extension' of the PSE-A nuclease protected region (173). This result implied that the factor(s) responsible for the extended footprint and PSF-B were interacting. This would be consistent with the hypothesis that additional 263P PSFs function to favour or stabilize the NF-1 complex. The sequences downstream of 103P were also analyzed with MatInspector 2.2 and the TRANSFAC 4.0 database (165, 166), using the same parameters that were used for the upstream sequence. The sequences used for this analysis spanned from nucleotides 223, the end of the PSE-A nuclease protection

region, to 263 (Fig. 6.5). As with the upstream region, the end of the PSE-A nuclease protection region was used instead of the end of 103P at nucleotide 229, to ensure that binding sites that may span the junction of 103P and the downstream region were not excluded. Putative factors from this analysis, which were not excluded based on expression patterns, are shown in Table 6.2.

5′- ^{223}GGAAGCGTTT ^{233}GCCTGTTTGT ^{243}TTGCTTGTGT ^{253}TTCTACAGAG ^{263}T -3′

Figure 6.5: The sequence of 263P downstream of the PSE-A nuclease protection region.
The 5′-3′ sequence of the CS-A 263P is shown from nucleotides 223-263. The PSE-A nuclease protection region ends at nucleotide 222. The 103P fragment ends at nucleotide 229, and the nucleotides that are part of the 103P sequence are highlighted.

Table 6.2
Results of the database search for additional putative PSFs downstream of PSE-A.

Factor	Core Similarity	Matrix similarity	Location
HNF-3β	1.0	0.925	(+ 231) ttgccTGTTtgtttg
HFH-2	1.0	0.987	(+ 233) gccTGTTtgttt
HFH-3	1.0	0.971	(+ 233) gccTGTTtgtttg
HFH-8	1.0	0.910	(+ 233) gccTGTTtgtttg
HNF-3β	1.0	0.914	(+ 235) ctgttTGTTtgcttg
HFH-2	1.0	0.924	(+ 237) gttTGTTtgctt
HFH-3	1.0	0.915	(+ 237) gttTGTTtgcttg
HFH-3	1.0	0.857	(+ 247) ttgTGTTtctaca
HFH-8	1.0	0.876	(+ 247) ttgTGTTtctaca
C/EBPβ	0.986	0.870	(- 249) ctctgtaGAAAcac

The location of the match is designated (+) or (-) to identify the strand where the match occurs.

HNF-3β and the HFH factors belong to the HNF-3/forkhead family of transcription factors (174). Like the RFX proteins, they associate with DNA sites through a winged

helix DNA binding motif, which is a subclass of the helix-turn-helix binding domain (291). The HNF-3/forkhead family appears to recognize sites with varying degrees of affinity. This results in sites that are recognized by more than one family member, as well as sites that are specific for a particular family member (292, 293). HNF-3 proteins have been previously demonstrated to be involved in regulatory complexes that include NF-1. HNF-3 and NF-1 regulate basal promoter activity of the microtubule-associated protein 1A gene (294) as well as function in regulating expression of several genes in liver cells including, collagen XVIII (295), methionine adenosyltransferase (296), protein C (297), serum albumin C4BPB (298), and transferrin (299). These reports strengthen the consideration of the HNF-3/forkhead family as candidate 263P PSFs that favour the formation of the NF-1 repressor complex.

The CCAAT enhancer binding protein (C/EBP) family recognize the same DNA binding site through a basic leucine zipper DNA binding domain (300). These proteins have also been demonstrated to regulate genes in cooperation with NF-1, examples of which are the human insulin receptor promoter (301), the mouse insulin-like growth factor-binding-protein-5 promoter (302), the mouse vas deferens protein gene (303), and the rat alpha-fetoprotein promoter (229).

The HNF-3/forkhead and C/EBP factors are not necessarily mutually exclusive candidates for additional PSFs. Evidence for these factors functioning together in regulatory complexes has been demonstrated for the rat alpha-fetoprotein far upstream enhancer (304), apolipoprotein B (305, 306), cytochrome P450 3A4 (307), and transthyretin (308) genes. In addition, NF-1, C/EBP, and HNF-3 factors are all components of the liver 6-phosphofructo-2-kinase transcriptional regulatory complex (309).

The 263P NF-1 repressor complex appears to repress CS-A promoter activity in pituitary GC cells through inhibition of the Pit-1 promoter complex (to be further discussed in section 6.2). Further support for C/EBP and HNF-3/forkhead factors as participants in this mechanism comes from documented interactions between these factors with Pit-1 or

other POU domain proteins. Interactions between HNF-3/forkhead family members and POU domain proteins have been described (310-312). C/EBPβ has been observed to regulate the gonadotropin-releasing hormone gene expression in cooperation with the POU domain protein Oct-1 (313), and C/EBPα cooperates with Pit-1 in the expression of the rat growth hormone (314) and prolactin promoters (315, 316) in pituitary cells.

6.1.4b Future analysis of the candidate PSFs.

The HNF-3/forkhead and C/EBP families of transcription factors are candidate PSFs that may favour or stabilize the formation of the 263P NF-1 repressor complex. This is based on (i) the observation of an 'extended PSE-A footprint' during an attempted purification of PSF-B, (ii) the presence of consensus binding sites in the PSE-A downstream region and (iii) literature support for the participation of these factors in regulatory complexes with NF-1 and POU domain proteins. The ability of these factors to participate in P sequence complexes should be investigated further experimentally.

An approach for the future analysis of the candidate PSFs is suggested, that begins with the use of *in vitro* structural assays. The consensus binding sites for these families can be tested for their ability to compete the NF-1 repressor complex from 263P in EMSA and nuclease protection assays. Confirmation of their association with the PSE-A downstream region can also be assessed by EMSA with the 41 bp 223-263P fragment as a probe. Functional *in vitro* assays should also be conducted. Mutation of HNF-3/forkhead and C/EBP sites in the context of 263PCSpLuc can be evaluated upstream of CSp.Luc in pituitary GC cells, to determine if these mutations have the capacity to effect 263P repressor function. These mutations could also be utilized as nuclease protection probes to assess their effect on the 263P nuclease protection pattern. To confirm the hypothesis that the region downstream of PSE-A, and not the region upstream of PSE-B, contributes to the formation of the 263P NF-1 repressor complex over the 103P RFX1 complex, P sequence fragments from 1-229P (103P + the PSE-B upstream region) and 127-263P (103P + the PSE-A downstream region) should be analysed structurally and functionally. Finally, the presence of these factors as members of P sequence complexes should be assessed *in situ* using the ChIP assays.

6.2 Potential mechanisms of pituitary repression through P sequences.

6.2.1 Repression of CS-A promoter activity in pituitary GC cells.

6.2.1a The 263P repressor complex in GC cells.

The analysis of 263P repressor activity in pituitary GC cells supported involvement of the NF-1 complex in the repressor mechanism. Mutation of the Pit-1 binding site in the CS-A promoter demonstrated the capacity for basal (or non-Pit-1 mediated) CS-A promoter activity in GC cells (Fig. 5.2). Repression of this 'basal' CS-A promoter by 263P was not observed. This result implies that the 263P repressor mechanism in pituitary GC cells involves interference with Pit-1 activation of the CS-A promoter (Fig. 6.6). In this interpretation, additional complexes would also have the capacity to form at both the CS-A promoter and 263P. These additional complexes likely account for the retention of reporter gene (luciferase) activity that is detected with the 263PCSp.Luc plasmid.

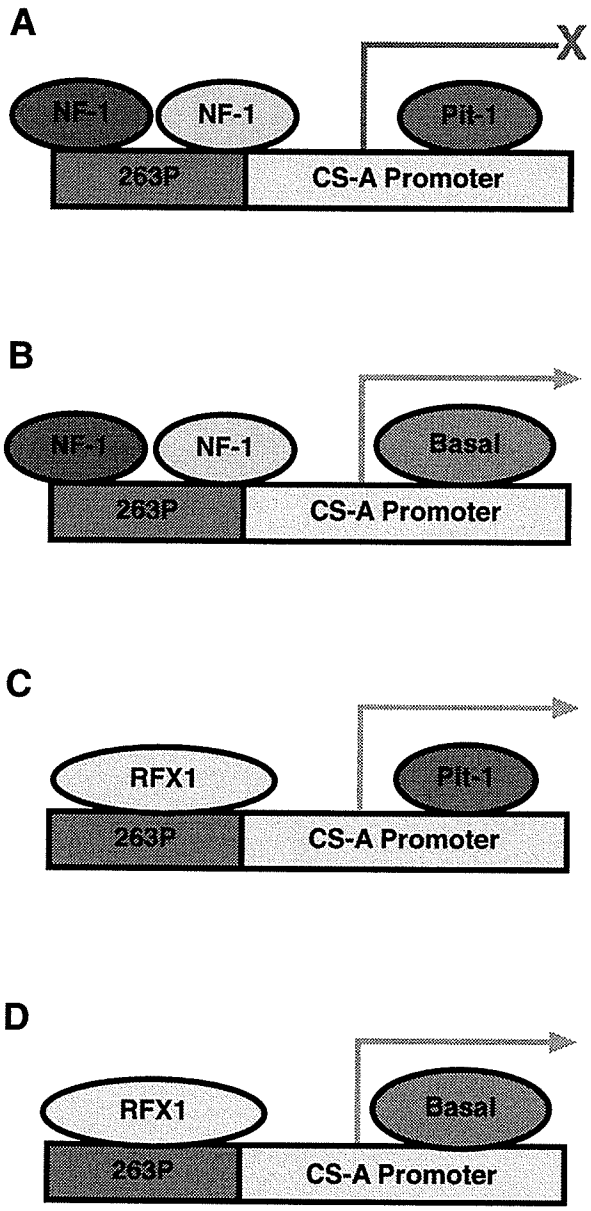


Figure 6.6: The pituitary repressor mechanism appears to result from inhibition of the activity of the Pit-1 promoter complex by the NF-1 263P repressor complex. Based on the data presented in Chapter 3 and Chapter 5, the above model is proposed for repression of the CS-A promoter in pituitary GC cells. The 263P NF-1 complex represses Pit-1 mediated (**A**), but not basal (**B**) activity of the CS-A promoter. (**C**) and (**D**) The 263P RFX1-containing complex does not repress activity of either CS-A promoter complex.

6.2.1b Potential mechanisms for *in vitro* repression.

Diversity in the COOH-terminal domain of NF-1 family members (discussed in section 6.1.2) implies that repression of CS-A promoter activity is mediated by a particular member(s) of the NF-1 family. The variability in the COOH-terminal region affects both DNA binding affinity as well as interactions with other proteins. This results in different mechanisms for achieving function, depending on the NF-1 family member that is involved. In the case of pituitary GC cells, transcripts for several members of the NF-1 family were observed (Fig. 3.19). It was not determined, however, which of these NF-1 family member(s) were involved in the repressor function.

Some insight into a potential mechanism comes from overexpression of NF-1 cDNAs and the effect on 263P function (Fig. 3.31). Overexpression of more than one member of the NF-1 family resulted in an increase in the capacity of 263P to repress the activity of the CS-A promoter. If the repressor action of the 263P NF-1 complex was mediated through the COOH-terminal domain, some degree of specificity may have been expected from these experiments. These results instead imply that the mechanism through which NF-1 mediates repressor activity involves structures or domains that are common to the cDNAs that were overexpressed. The common domain in this case is the NH_2-terminal DNA binding domain.

Two general mechanisms were outlined in section 6.1.2b through which NF-1 proteins mediate their functional effects. The first mechanism uses the COOH-terminal domain for specific interactions with, and recruitment of, other factors and complexes. The second general mechanism of NF-1-mediated effects is achieved by blocking the association of other factors due to the presence of NF-1 binding. Involvement of the DNA-binding domain, without specificity for a particular COOH-terminal domain, supports the interpretation that in the case of 263P, NF-1 repression is achieved through a passive mechanism. That is, that mutually exclusive binding of NF-1 with a potential 263P activator occurs. NF-1 was previously demonstrated to down regulate the activity of the rat poly (ADP-ribose) polymerase promoter through mutually exclusive binding with

Sp1 (317). Mutually exclusive binding between NF-1 and RFX1 at PSE-A was certainly observed, and provides additional support for this interpretation.

Repressor activity with 263P in pituitary GC cells was only observed with the intact CS-A promoter that had the capacity to form Pit-1 complexes. This result implied some degree of specificity by the NF-1 complex for inhibition of Pit-1 mediated activation in these cells. The literature does not contain evidence for previously documented NF-1 and Pit-1 interactions. NF-1, however, has been demonstrated to cooperate with the Pit-1 related transcription factor Oct-1 in the replication of adenoviral DNA (220, 318-320). For this process, the DNA binding domains of NF-1 (153, 223) and Oct-1 are both necessary and sufficient (321). In fact, the Pit-1 POU domain can be substituted for the Oct-1 POU domain in this process (321). The DNA binding domain of NF-1 was also the region that was interpreted above to be involved in repression of Pit-1 mediated activation. The adenoviral replication system may support a regulatory mechanism that involves the NF-1 and Pit-1 DNA binding domains.

The data presented in this thesis support protein-protein interactions between NF-1 and Pit-1 in the presence of DNA binding elements (Figs. 5.3 and 5.4), which were not detected in the absence of DNA binding sites (Figs. 5.6 and 5.7). Pit-1 has the capacity to undergo conformational changes when associated with its DNA binding site, and this can alter its interactions with other proteins (131, 132). This has the potential to increase the stability of interactions that did not readily occur in the absence of DNA binding. Therefore, while the NF-1 repressor mechanism may involve mutually exclusive binding with putative 263P activators, the specificity in GC cells for repression of the Pit-1-mediated activity may also involve interactions between NF-1 and Pit-1. These interactions were only observed in the presence of DNA. Thus, it is possible that the inclusion of DNA in the EMSA and nuclease protection experiments either (i) stabilized putative NF-1/Pit-1 and RFX1/Pit-1 interactions or (ii) contributed to a protein conformation that permitted the occurrence of these interactions. To examine the possible contribution of DNA to the stability of the interactions, one possible experiment would be to isolate P sequence factors in the presence of DNA, such as on a DNA affinity column.

The purified complexes could be analyzed by protein blotting for the presence of NF-1, RFX1, and Pit-1. Alternatively, the proteins could be analyzed using mass-spectrometry to identify these and potentially additional proteins as components of the PSF complex.

6.2.2 Applying the *in vitro* analysis to potential mechanisms for *in vivo* repression.

P sequences repress the activity of the CS-A promoter *in vitro*. This ability has been proposed to contribute to the lack of placental GH/CS expression in the pituitary *in vivo*, but this hypothesis has not been tested directly. The data presented in Chapter 4 demonstrated an inverse correlation between hyperacetylation at both P sequences and the CS promoters, with the detection of transcription factor binding at 263P. Specifically, NF-1 and RFX1 associated with P sequences in pituitary chromatin, but were not observed in placental chromatin, and 263P and the CS promoters were hyperacetylated in placental but not pituitary chromatin samples. In the pituitary, the observation of transcription factor association with the P sequences supports the hypothesis that these elements may contribute to the differential gene expression pattern of the GH/CS locus in this tissue.

6.2.2a Regulation of the GH/CS locus in the pituitary *in vivo*.

As discussed in Chapter 1.2.1, the transcription factor Pit-1 is integral to pituitary expression of GH *in vivo* as well as *in vitro*. The current model of GH expression *in vivo* implicates Pit-1 association at the HS I/II region as the key event in regulation of the locus in the pituitary (123-126). It is hypothesised that Pit-1 binding in this region results in a 'peak' of hyperacetylation, that then spreads throughout the LCR and to GH-N, but not the placental GH/CS genes (123, 129). Pit-1 has previously been shown to have the capacity to recruit co-activators, such as the histone acetyltransferases CREB binding protein (CBP) and p/CAF (132). This ability is believed to contribute to hyperacetylation of histones in the locus and LCR.

In Chapter 4, ChIP with the hyperacetylated H4 antibody and human pituitary chromatin samples resulted in several B/I ratios that were significantly higher than background levels (Fig. 4.8). In the LCR, hyperacetylation of HS I/II, HS III, and HS IV regions was observed. In contrast to previous studies, a 'peak' of hyperacetylation in pituitary chromatin was not observed for the HS I/II region (123, 129, 150). One possibility for this discrepancy involves the source of tissue for the experiments. In reports where a 'peak' of hyperacetylation was observed, the tissue samples expressed abnormally high levels of GH, and were enriched for somatotrophs. In contrast, the tissue samples used for the experiments in Chapter 4 were 'normal' pituitary samples. The lack of somatotroph enrichment in these samples may have masked the ability to observe the hyperacetylation 'peak'. Alternatively, these observations may also support a hypothesis where multiple regions of the GH LCR are involved in pituitary regulation of the locus. This alternate hypothesis is consistent with assessment of locus control in transgenic mice, where appropriately regulated expression, as defined by a lack of gigantism, has only been reported in the presence of the entire LCR (120, 122, 123) (See Table 1.1).

6.2.2b Potential mechanisms for repression of placental GH/CS gene expression in the pituitary *in vivo*.

The data in Chapter 4 also extended the correlation between hyperacetylation and gene expression in the pituitary, as the first evidence was presented that the GH but not the CS promoters were hyperacetylated in pituitary chromatin. This difference was observed despite the presence of Pit-1 binding sites in the placental promoters. It suggests that if Pit-1 is involved in hyperacetylation of GH-N and the LCR, then blockage of Pit-1 mediated hyperacetylation may contribute to the lack of placental gene expression in the pituitary. There are at least two methods through which this could be achieved: (i) disruption of Pit-1 binding at the placental promoters or (ii) interference with the ability of associated Pit-1 to recruit co-activators involved in histone hyperacetylation and gene expression. In either scenario, it is proposed that P sequences contribute to this mechanism.

Interference with the ability of Pit-1 to associate with the promoter is an example of passive repression. Pit-1 association at any site in the locus has yet to be accurately determined, and it is unknown whether the association occurs at GH-N but not the placental promoters *in vivo*. It has been reported, however, that RFX1 and NF-1 association with P sequences can be detected in pituitary chromatin samples (Fig. 4.9) (130). It is a possibility that the presence of these factors upstream of the placental, and not the GH-N, promoters, sterically inhibits the association of Pit-1. In the context of chromatin, 263P may actually be 'closer' to the Pit-1 binding site than in the context of 263PCSp.Luc. There is an approximately 5-fold increase in the 2195 base pairs that separate 263P from the proximal Pit-1 binding site in the CS-A promoter in the genome, compared to the 443 base pairs that separate 263P from the proximal Pit-1 binding site in 263PCSp.Luc. The genome, however, is condensed as chromatin, which at the level of the 10 nm fiber has a packaging ratio of 6-7 and at the level of the 30 nm fiber is packaged approximately 40-fold. The relative 'nearness' of the P sequence complex to the placental promoters could contribute to steric inhibition of Pit-1 association. A comparison of Pit-1 binding at the GH and CS promoters, using the ChIP protocol developed here for human pituitary tissue samples, would be the next experimental step in testing support for this possible mechanism.

A second potential mechanism for repression *in vivo* is an example of active repression. Pit-1 has the capacity to recruit histone acetyltransferase (HAT), as well as histone deacetylase (HDAC) complexes (131, 132). Alteration in the recruitment of co-activators/co-repressors was shown to occur based on the spacing of nucleotides in the Pit-1 binding sites (131). The proximal Pit-1 contact region of the rat GH promoter contains a 2 bp insertion when compared to a Pit-1 site from the rat prolactin promoter (Fig. 6.7A). In this study, the 2 bp insertion was demonstrated to alter the structure of the Pit-1/DNA binding site complex (131). That is, due to the flexible nature of POU domain proteins, the conformation of Pit-1 when associated with its DNA site was found to depend on the sequence information in the site itself. Based on this analysis, the homologous GH and CS promoters would be expected to favour the recruitment of HDACs over HATs, as the spacing of nucleotides in their proximal Pit-1 sites is

homologous to the spacing in the rat pituitary GH promoter (104, 131, 216) (Fig. 6.7B). In contrast, the spacing of nucleotides in the HS I/II Pit-1 binding sites, being more like the prolactin than the GH Pit-1 site, would be expected to favour the recruitment of HATs (Fig. 6.7C). There are two pieces of experimental evidence in the literature that would support the hypothesis that the GH-N and HS I/II Pit-1 sites associate with Pit-1 differently, and that this has functional implications. The first example comes from *in vitro* structural assessments of Pit-1 binding to the HS I/II region. When Pit-1 association with HS I/II was first demonstrated, the rat prolactin Pit-1 site, but not the human GH proximal Pit-1 site, was an efficient competitor for Pit-1 binding (124). This observation suggests that Pit-1 binding at the HS I/II region more closely resembles Pit-1 association with the prolactin than the GH promoter site. The second example comes from a study that compared the function of GH promoter and HS I/II Pit-1 sites in transgenic mice (126). It was reported that a tandem array of the three Pit-1 binding sites from HS I conferred somatotroph-specific expression of a GH-N transgene in six of transgenic mouse six lines. In contrast, a comparable array composed of three rat GH promoter Pit-1 sites (homologous in spacing to human GH) did not result in pituitary transgene expression (126).

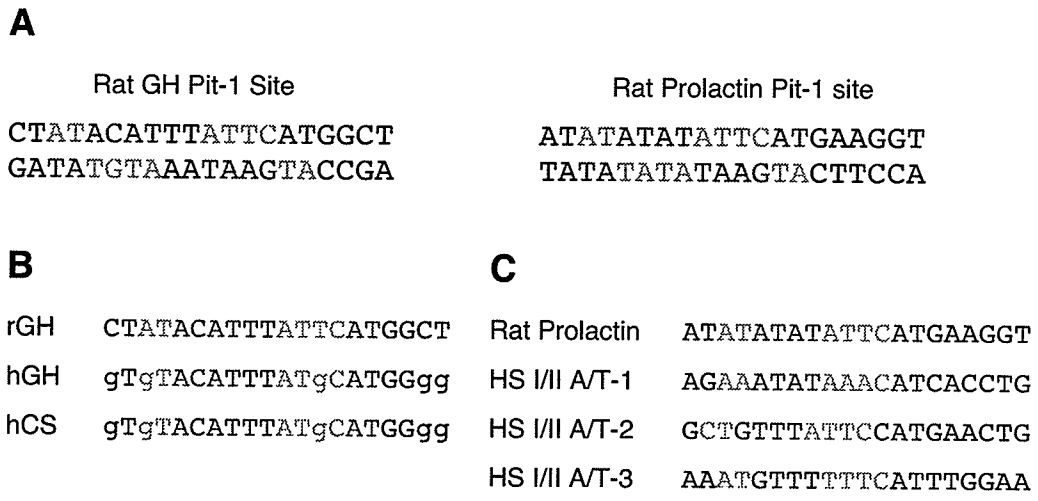


Figure 6.7: Spacing of nucleotides in Pit-1 binding sites.

(A) The spacing of nucleotides in the GH Pit-1 site is believed to assist in the recruitment of HDAC complexes, while the spacing of the prolactin Pit-1 site assists in the recruitment of HAT complexes (131). **(B)** Alignment of the proximal Pit-1 sites in the rat (r) and human (h) GH and CS promoters. Sequence information obtained from (104, 216). Mismatched nucleotides are noted by lower case. **(C)** Comparison of the spacing in the rat prolactin Pit-1 site and the three HS I/II Pit-1 sites. Alignments are based on (124).

From these observations, a model is proposed where the presence of pituitary P sequence complexes, upstream of the placental promoters, assists in the recruitment of HDACs and the resulting lack of hyperacetylation that is observed. This model is supported by the structure of the placental promoter Pit-1 sites, which would favour the recruitment of HDACs over HATs, and the additional observation that NF-1 also has the capacity associate with HDAC complexes (199). The question arises as to what is responsible for the hyperacetylation that is observed at the GH-N promoter, given that its Pit-1 site would also favour HDAC recruitment? P sequences are not found upstream of the GH-N promoter. Instead, it is the HS I/II region, with three potentially HAT-recruiting Pit-1 sites. As hyperacetylation is believed to be the result of an imbalance in the dynamic processes of acetylation over deacetylation, this model would assume that the GH-N promoter is influenced by the proximity of HAT recruitment at the HS I/II region, and results in the hyperacetylation that is observed. Support for this model can be found in studies involving the β-globin locus. An analogous situation has been described, where

active deacetylation is believed to contribute to the silencing of promoters in the locus that are not expressed (322). According to the histone code hypothesis (50), the hyperacetylation pattern of the locus would then be 'read' by the transcriptional machinery and result in differential expression of the GH/CS family members in the pituitary. In an effort to test the viability of this theory, pituitary ChIP experiments with co-activator and co-repressor antibodies should be performed.

The application of this model assists in understanding the failure to observe pituitary repression when 263P was used in transgenic mice (150). The construct that was utilized in this study consisted of a modification of the 1.6 kb HS I/II(F) / 0.5 kb GH-N transgene (described in section 1.2.1), where 0.5 kb GH-N was flanked by 263P (150). The 1.6 kb HS I/II region is a potent enhancer of gene activity in the pituitaries of transgenic mice (120). This enhancer activity localizes to the Pit-1 binding sites (124-126), which are spaced in a manner similar to the rat prolactin Pit-1 sites (Fig. 6.7). In this model, the activity 1.6 kb HS I/II *in vivo* is presumed to be due to the recruitment of co-activator histone acetyltransferase complexes, supported by the observation that deletion of two Pit-1 sites in the P1 transgenic mouse results in both a loss of HS I/II hyperacetylation and GH-N expression in the pituitary (123). Thus, in the 1.6 kb HS I/II(F) / 263P / 0.5 kb GH-N / 263P transgenic mice, HAT complexes should be actively recruited to the HS I/II region. As previously discussed, NF-1 is a context dependant regulator (section 6.1.2c). As well as its ability to interact with HDAC complexes (199), it has been demonstrated to recruit the histone acetyltransferase CBP (231). The proximity of NF-1 to the HS I/II region in this construct could influence the NF-1 family member that associates with 263P, as well as the co-activators versus co-repressors that are ultimately recruited. To assess the ability of P sequences to repress pituitary expression of the placental genes *in vivo*, an approach is suggested using RecA-assisted homologous recombination, which was used to delete Pit-1 binding sites from the HS I/II region (123), to delete one or more 263P sequences from their native position in the GH/CS locus. In addition to investigating the contribution of 263P to mutually exclusive gene expression in the pituitary, these experiments would have the added advantage of assessing the

contribution of P sequences to expression of the locus in the placenta. This approach is discussed further in section 6.4.1a.

6.3 Potential mechanisms for the involvement of P sequences in activated expression of the GH/CS locus in the placenta.

6.3.1 The *in vitro* system did not model P sequence enhancer activity in the placenta.

A third research objective for this thesis was to determine if the placental (choriocarcinoma) cell lines were a suitable *in vitro* model for the study of placental enhancer function, and if so, to use this system to identify candidate PSFs. Although the *in vitro* experiments provided the first evidence of significant 263P function in placental cell lines, this activity was not as an enhancer, but as a repressor element (Fig. 3.2). As discussed in section 6.1.1c, the *in vivo* enhancer activity correlates with the observation of histone hyperacetylation. The *in vitro* gene transfer system may not necessarily reflect these enhancer activity mechanisms. It is possible, however, that the *in vitro* system has the capacity to provide useful information about protein binding events. Based on this potential, the *in vitro* binding events were assessed, and NF-1 and RFX1 became candidate placental PSFs.

A second piece of information provided by the *in vitro* system was the potential for functional cross-talk between the 263P and CS downstream enhancer elements. The 241 bp CS-B element is a potent enhancer of gene expression in placental cells (137-139). The construct used in the initial characterization of 263P function contained this CS-B enhancer fragment, and a significant effect of 263P in placental cells was not observed (133). This contrasts to the significant repression of reporter gene activity by 263P in placental cells when the CS enhancer was not included in the construct (Fig. 3.2). Taken together, these observations suggest that the presence of the CS enhancer had the capacity

to override or mask the repressor function of 263P, implying that functional cross-talk can occur between the 263P and CS enhancer complexes.

6.3.2 Potential mechanisms for P sequence enhancer activity *in vivo.*

The composition of the placental P sequence complex *in situ* was discussed in section 6.1.1c. Briefly, the ChIP assays did not confirm the association of either NF-1 or RFX1 with P sequences in human term placenta chromatin (Fig. 4.10). The interpretation of this result was that either (i) these factors associate and some aspect of the experiment failed to identify this, (ii) other members of the RFX family, which were not assessed, associate with P sequences in the placenta, or that (iii) the placental PSFs are separate factors from NF-1 or RFX1. Each of these interpretations will be considered in terms of the potential mechanisms for P sequence enhancer activity in the placenta.

6.3.2a Regulation of the GH/CS locus in the placenta *in vivo.*

As discussed in Chapter 1.2.2d, both proximal and distal regulatory elements appear to be required for appropriate regulation of the GH/CS locus in the placenta. Proximal elements, such as the P sequences, were proposed to function as enhancer elements that allowed for activated expression levels. Distal elements, such as the HS III-V region were proposed to insulate the locus from the effects of surrounding chromatin, essentially creating an independent chromatin domain for the genes of the locus. These interpretations were based on the data obtained from transgenic mouse lines. Constructs containing the distal HS III-V LCR elements alone gave gene expression levels that correlated to transgene copy number, however, the expression levels in the placenta were low or nonexistent (120). In contrast, constructs containing proximal sequences, but lacking the distal HS III-V region, resulted in variable levels of expression that were deemed to be the result of site of integration effects (120, 150). The observation of hyperacetylation within the HS III-V region and at multiple proximal regions (Fig. 4.7) supports the interpretation that these sequences play a role in regulation of the GH/CS locus in the placenta.

The current theory of LCR function has been based primarily on observations of the β-globin multigene locus. As discussed in Chapter 1.1.2d, this model places the role of the LCR as a potent enhancer of gene activity, with more proximal sequences involved in poising the locus for transcription (55, 68-72). The regulation of GH gene expression in the pituitary appears to agree with this model, with the distal HS I/II region functioning as the primary enhancer element to achieve activated expression levels (120, 123, 124). It is unclear, however, which regions function to poise the GH/CS locus for expression in somatotrophs or syncytiotrophoblasts. Transgenic mouse data for both the pituitary and placenta implicate the HS III-V region in this role. Further studies are needed to determine if this is in fact the function of this region in both tissues. It is clear, however, that in the placenta, the LCR 'enhancer' element may not reside within the defined distal LCR (HS sites) at all. Proximal sequences, in particular the P sequences and downstream CS enhancers have the potential to be involved in a mechanism for locus control *in vivo* that deviates from the pituitary GH-N and β-globin models. That is, in the case of regulating placental expression of the GH/CS locus, these relatively more proximal sequences appear to function as the primary enhancer elements.

The observation of functional cross-talk between 263P and the CS enhancer *in vitro*, may therefore, also play a role in regulating gene expression *in vivo*. In term placenta syncytiotrophoblasts, the expressions of the either the CS-A or CS-B genes are both greater than expression of the GH-V gene (83). The P sequences are located upstream of each of these genes, and therefore, are clearly not the only elements involved in activated placental expression. In contrast, the CS enhancers are only located downstream of the CS genes. The presence of these sequences alone may account for the increased levels of CS relative to GH-V expression. Alternatively, the possibility exists that the CS enhancers and the P sequences work synergistically to result in expression of CS, that at term has been noted to be greater than any other peptide hormone produced in the human body (323).

6.3.2b Proposed P sequence mechanisms.

The P sequences have been proposed to function as enhancers of placental gene expression *in vivo*, and have been observed to contain hyperacetylated histones in human term placenta (Fig. 4.7) and (150). A question arises as to whether the histone hyperacetylation is a critical component of the putative enhancer activity *in vivo*. This question has not been examined here. A second question is also evident, however, as to the mechanisms responsible for the recruitment of histone modifying complexes, specifically, histone acetyltransferases (HATs), to P sequences in the placenta. The attempt to identify placental PSFs *in situ* was seen as a first step in identifying the mechanism, or source, of the placental hyperacetylation. As the *in situ* P sequence complex was not determined, there are several possibilities that remain for the source of placental hyperacetylation.

An investigation of the placental P sequence complex *in situ* did not detect the association of either NF-1 or RFX1 (Fig. 4.10). The involvement of these factors as components of this complex, however, was not ruled out, based on the lack of available positive controls. The mechanistic implications of NF-1 and RFX1 involvement will, therefore, be discussed.

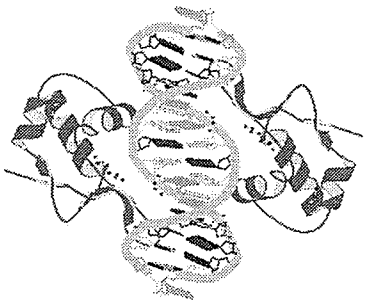
In considering the potential for NF-1 involvement, a case could be made that this association contributes to the generation of and/or assists in maintaining accessible hyperacetylated P sequences in human term placenta. In terms of HAT complex recruitment, a direct interaction between NF-1 and CBP has previously been demonstrated (257). It has also been proposed that NF-1 may have the capacity to replace repressive histone structures, or disrupt the distribution of nucleosomes (324). This hypothesis is supported by observations that NF-1 has the capacity to directly interact with histone H3 (249, 324), and that histone H1 and NF-1 may recognize similar binding sites (325, 326). This second observation was proposed as support of the idea that NF-1 may activate transcription by direct displacement of histone H1 binding, which *in vitro* is supported by the capacity of NF-1 to overcome the repressive effects of histone H1 on promoter function (327). The potential for NF-1 to be involved in chromatin remodeling

processes that relate to gene expression *in vivo*, has perhaps been best demonstrated, however, with the well characterized mouse mammary tumour virus (MMTV) promoter system. In this system, the requirement of NF-1 for hormone-mediated chromatin remodeling has been suggested to be the result of NF-1 involvement in stabilization of the remodeling complex at the promoter (230, 328-331). Taking all of the above information together, there are clearly several ways in which NF-1 has the potential to contribute to the observation of P sequence hyperacetylation in human term placenta.

In contrast to the case of NF-1, there is not a direct link between RFX1 and histone acetyltransferases or chromatin structure that is currently present in the literature. In terms of chromatin studies, the most well characterized system involving the RFX family is the role of RFX5 in regulation of major histocompatibility complex (MHC) class II genes. RFX5 is a component of the protein complex known as the MHC enhancesome (275, 332). This complex associates with X boxes in MHC class II promoters and sets the stage for further recruitment of factors, namely the non-DNA-binding co-activator protein CTIIA, and subsequent chromatin modifications that lead to gene activation *in vivo* (333-335). Although RFX5 is considerably divergent from other members of the RFX family (188, 275), the common feature that is shared is the winged helix DNA binding domain (188, 189). This feature presents an intriguing possibility for the potential involvement of RFX1 in P sequence hyperacetylation and enhancer function *in vivo*. Based on the *in vitro* binding evidence, RFX2 and RFX3 were also candidate PSFs in the placenta, therefore, the following potential mechanism also applies to these factors.

The DNA binding domain of the RFX family is a winged helix (189). The winged helix structure does not 'wrap' around the DNA sequence (Fig. 6.7A), as is seen with many other transcription factors, such as leucine zippers or homeodomains (Fig. 6.7B), but instead associate with only one side of the sequence. This structure is reminiscent of classical steroid receptors such as estrogen receptor (ER), androgen receptor (AR), progesterone receptor (PR), and glucocorticoid receptor (GR) (Fig. 6.7C).

A



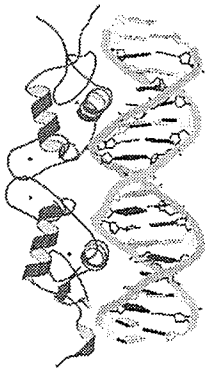
Two RFX monomers on either side of the DNA element.

B

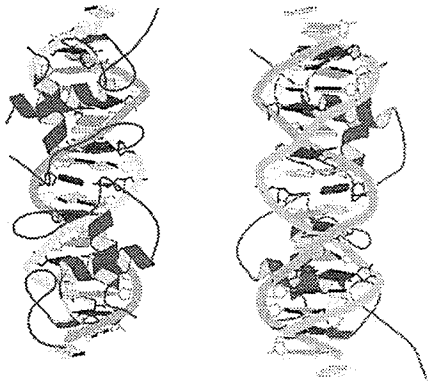
Pit-1

C/EBP

C



GR



ER

Figure 6.8: DNA binding domain structures.
DNA binding domain structures are taken from the Entrez-Structure protein database (PBD) at NCBI (http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Structure). **(A)** The co-crystal structure of the RFX DNA binding domain in complex with its cognate X-box binding site. This schematic shows two RFX molecules on either side of the DNA molecule. The RFX monomers do not contact each other, and each associate with only one side of the DNA molecule. The PBD entry for this structure is 1DP7 and is based on (189). **(B)** The DNA binding domains of homeodomain and leucine zipper proteins 'wrap' around the DNA, making contacts on both sides of the molecule. As representative members of these families, the structures of the POU homeodomain factor Pit-1 and the leucine zipper C/EBPα are shown. The PBD entry for Pit-1 is 1AU7 and is based on (336). The PBD entry for C/EBPα is 1NWQ and is based on (337). **(C)** Steroid receptors make contact with only one side of the DNA. Representative structures shown are the GR and the ER. The PBD entry for GR is 1GLU and is based on (338). The PBD entry for ER is 1HCQ and is based on (339).

Steroid receptors have the capacity to associate with their DNA binding site, even when that site is contained within the nucleosomal structure (340-342). This may be due to the manner in which these factors associate with DNA, in that by requiring only one side of the DNA for contact, remodeling of the nucleosomal structure would not appear to be a necessary prerequisite for recognition of their respective binding site. As such, in many systems, the association of a steroid receptor complex has been demonstrated to be the 'first step' in the initiation of chromatin remodeling events that lead to activation (343). If this ability is due to the manner in which these factors associate with their binding sites, then it is reasonable to speculate that RFX proteins, with their winged-helix DNA binding domains, may have this same capacity. This hypothesis is supported by the 'initiator' role of the RFX enhancesome in the MHC system. Further support comes from studies involving the transcription factor HNF-3, which also is a winged helix protein (291). HNF-3 has been shown to have the ability to associate with its binding site in the albumin promoter even when the chromatin is in a compacted state. This leads to increased accessibility of the chromatin structure for gene expression (344). From these considerations, the role of RFX1 in the P sequence placental system may be early association, which then allows for chromatin remodeling events to occur. If this is the case, it may explain why RFX1 association was not detected in the human term placenta chromatin.

There is also the possibility that the P sequence complex in the placenta does not contain either NF-1 or RFX1. As the *in vitro* system does not model the *in vivo* enhancer activity, candidate PSFs are not limited to the PSE-A and PSE-B regions. In section 6.1.4a, additional putative PSFs were proposed that were based on database analysis of the 263P sequence (Tables 6.1 and 6.2). Several factors are within these tables. It is interesting to note that this analysis included HNF-3 and other related winged helix factors. The above speculation for the potential role of the RFX family could also be applied to these factors.

6.4 Future Directions

6.4.1 The pituitary system.

6.4.1a The contribution of 263P to a lack of placental gene expression in the pituitary.

The placental GH/CS genes are not expressed in the pituitary, despite extensive homologies in their regulatory regions (83), and an accessible chromatin structure (127). The studies described here were centred around a role for P sequence elements in contributing to differential gene expression in the pituitary, and the capacity for these elements to function as repressors of placental gene activity has been investigated in a pituitary cell model system. An outstanding question, however, is whether the P sequences contribute to a lack of placental gene expression in the pituitary *in vivo*.

Previously, it has been demonstrated that an 87 kb fragment of human chromosome 17 (the P1 clone), containing almost the entire GH/CS locus (it lacks the CS-B gene) and LCR, will result in efficient and appropriately regulated expression of the human GH/CS family in the pituitaries of transgenic mice (122). To assess the contribution of individual regions within the LCR towards pituitary expression, the use of targeted deletions within this construct have also been reported (123). It is proposed that this same approach could be taken to test the hypothesis that P sequences repress pituitary expression of the

placental genes in the pituitary *in vivo*. This would require that transgenic mouse lines containing the P1 clone be obtained or generated. In addition, one or more of the 263P elements in the P1 clone could be deleted using a RecA-assisted homologous recombination strategy (123, 345), and the resulting transgenic mice compared to the intact P1 lines. As the P1 clone does not contain CS-B, CS-L is a pseudogene, and GH-V is expressed at reduced levels compared to the CS genes in the placenta, it would be important to ensure that the P sequence upstream of the CS-A gene was deleted in these constructs. The generation of the P1 and P1Δ263P transgenic mice would serve two purposes. They would first provide an opportunity to assess if deletion of the P sequences resulted in expression of the placental genes in transgenic pituitaries. Secondly, these same lines could be used to assess the contribution of P sequences to regulation of the GH/CS locus in the placenta (see section 6.4.2).

6.4.1b The role of Pit-1.

The data presented here supported the hypothesis that the *in vitro* repressor mechanism of 263P in pituitary cells involves interference with Pit-1 transactivation. The application of this to a potential repression mechanism *in vivo* was also discussed (6.2.2b). Briefly, it was speculated that the role of the P sequences *in vivo* may involve either interference with the ability of Pit-1 to associate with the placental promoters, or alternatively, if Pit-1 association occurs, contribution to the lack of histone hyperacetylation in the 263P and placental promoter regions. A question arising from this is whether Pit-1 associates with the placental promoters in the pituitary *in vivo*, which has not been previously assessed.

The assessment of Pit-1 association *in vivo* requires the availability of a 'good' antibody for immunoprecipitation. A current limitation is that the commercial antibodies that are available are not adequate for immunoprecipitation experiments (unpublished data from our laboratory and personal communication with Dr. Harry Elsholtz, University of Toronto). When a 'good' Pit-1 antibody is available, it is proposed that it should be used to assess the association of Pit-1 with the CS promoters, as well as the GH-N promoter and the HS I/II region using ChIP with human pituitary tissue samples. These

experiments could also be extended to assess the recruitment of HAT and HDAC complexes at these regions, as well as at 263P.

6.4.1c Identification of additional PSFs.

In section 6.1.4, a rationale was presented for the hypothesis that additional PSFs may be present in the 263P repressor complex. Candidate factors were proposed in section 6.1.4a, most notably, members of the HNF-3/forkhead and C/EBP families of transcription factors. An approach for the investigation of this hypothesis was outlined in section 6.1.4b.

6.4.2 The placental system.

Clearly, relatively less is understood regarding the regulation of the GH/CS locus in the placenta than in the pituitary model. Perhaps most intriguing, however, is that while the GH-N/pituitary system appears to resemble regulation of the β-globin multigene cluster, the placental system offers a potentially different model for analysis of the mechanisms involved in locus control. The evidence to date indicates that in the placental system, the *in vivo* 'enhancer' sequences, which would qualify as the LCR under the current definition of this activity, may be more proximal than the distal upstream HS sites, namely the P sequences and the CS enhancer regions.

6.4.2a Regulation of the GH/CS locus in the placenta.

In section 6.4.1a, an approach was proposed for investigating the role of P sequences in the pituitary that used targeted deletions in the P1 clone and transgenic mice. As previously mentioned, these deletions would also serve as a model system to assess the contribution of P sequences in regulation of the GH/CS locus in the placenta.

The regulation of the GH/CS locus in the placenta could also be assessed with the use of a transgene that combines the HS III-V region with the 15 kb CS-A HindIII fragment. As outlined in Chapter 1.2.2c (see Table 1.2), the 15 kb CS-A HindIII construct results in placental expression of CS-A, the level of which appears to be subject to the site where

the transgene is integrated (120). In addition, the HS III-V region has been demonstrated to have the capacity to confer site of integration independent expression (see Tables 1.1 and 1.2). A construct that combines these two elements, and assesses CS-A expression, may provide an alternate model system for the analysis of placental expression. This construct could be used in either transgenic mice or stable cell culture lines. Deletions and targeted mutations in this construct could be used to investigate the contributions of individual regulatory elements, such as the P sequences, CS enhancers, and distal upstream HS sites, to gene expression levels.

6.4.3b Identification of the placental PSFs.

The *in vitro* placental cell structural and functional data appear to resemble the pituitary repressor complex. Despite this, two possible approaches are proposed for the identification of candidate placental PSFs. The first approach would be to assess the locations where proteins associate with P sequences in the placenta using *in vivo* footprinting. Using this technique it may be possible to locate regions of protein binding and correlate these regions to candidate PSFs based on database analysis. A second possible approach for the identification of candidate placental PSFs would be to focus on determining the sequence requirements for 263P enhancer activity. Enhancer activity of 263P was observed in both a chromatin context (150) as well as when the 263P fragment was linked to the minimal thymidine kinase (Tk) promoter (Fig. 5.1). Using either the *in vitro* Tk promoter system, or potentially the HS III-V 15 kb CS-A transgene that was described in section 6.4.2a, scanning mutations in the 263P sequence could be tested for their ability to disrupt enhancer activity. Sequences with some contribution to enhancer activity can be assessed by database analysis for potential binding factors. Following the identification of candidate factors using one or more of the above proposed approaches, the association of candidate placental PSFs could be assessed *in situ* using ChIP with human term placenta tissue.

CHAPTER 7

REFERENCES

1. **Arents G, Burlingame RW, Wang BC, Love WE, Moudrianakis EN** 1991 The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. Proc Natl Acad Sci U S A 88:10148-52.
2. **Klug A, Rhodes D, Smith J, Finch JT, Thomas JO** 1980 A low resolution structure for the histone core of the nucleosome. Nature 287:509-16.
3. **Kornberg RD** 1974 Chromatin structure: a repeating unit of histones and DNA. Science 184:868-71.
4. **Kornberg RD, Thomas JO** 1974 Chromatin structure; oligomers of the histones. Science 184:865-8.
5. **Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ** 1997 Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389:251-60.
6. **Noll M, Kornberg RD** 1977 Action of micrococcal nuclease on chromatin and the location of histone H1. J Mol Biol 109:393-404.
7. **Thomas JO, Kornberg RD** 1975 An octamer of histones in chromatin and free in solution. Proc Natl Acad Sci U S A 72:2626-30.
8. **Richmond TJ, Davey CA** 2003 The structure of DNA in the nucleosome core. Nature 423:145-50.
9. **Olins AL, Olins DE** 1974 Spheroid chromatin units (v bodies). Science 183:330-2.
10. **Woodcock CL, Safer JP, Stanchfield JE** 1976 Structural repeating units in chromatin. I. Evidence for their general occurrence. Exp Cell Res 97:101-10.
11. **Woodcock CL, Sweetman HE, Frado LL** 1976 Structural repeating units in chromatin. II. Their isolation and partial characterization. Exp Cell Res 97:111-9.
12. **Spadafora C, Bellard M, Compton JL, Chambon P** 1976 The DNA repeat lengths in chromatins from sea urchin sperm and gastrule cells are markedly different. FEBS Lett 69:281-5.

13. **Thoma F, Koller T, Klug A** 1979 Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. J Cell Biol 83:403-27.

14. **Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD** 1994 Molecular Biology of the Cell 3rd ed. Garland Publishing

15. **Finch JT, Klug A** 1976 Solenoidal model for superstructure in chromatin. Proc Natl Acad Sci U S A 73:1897-901.

16. **Horowitz RA, Agard DA, Sedat JW, Woodcock CL** 1994 The three-dimensional architecture of chromatin in situ: electron tomography reveals fibers composed of a continuously variable zig-zag nucleosomal ribbon. J Cell Biol 125:1-10.

17. **Woodcock CL, Horowitz RA** 1995 Chromatin Organization re-viewed. Trends Cell Biol. 5:272-7

18. **Paulson JR, Laemmli UK** 1977 The structure of histone-depleted metaphase chromosomes. Cell 12:817-28.

19. **Marsden MP, Laemmli UK** 1979 Metaphase chromosome structure: evidence for a radial loop model. Cell 17:849-58.

20. **Cook PR, Brazell IA** 1975 Supercoils in human DNA. J Cell Sci 19:261-79.

21. **Cook PR, Brazell IA** 1976 Conformational constraints in nuclear DNA. J Cell Sci 22:287-302.

22. **Cook PR, Brazell IA, Jost E** 1976 Characterization of nuclear structures containing superhelical DNA. J Cell Sci 22:303-24.

23. **Iarovaia O, Hancock R, Lagarkova M, Miassod R, Razin SV** 1996 Mapping of genomic DNA loop organization in a 500-kilobase region of the Drosophila X chromosome by the topoisomerase II-mediated DNA loop excision protocol. Mol Cell Biol 16:302-8.

24. **Beato M, Eisfeld K** 1997 Transcription factor access to chromatin. Nucleic Acids Res 25:3559-63.

25. **Imbalzano AN, Kwon H, Green MR, Kingston RE** 1994 Facilitated binding of TATA-binding protein to nucleosomal DNA. Nature 370:481-5.

26. **Godde JS, Nakatani Y, Wolffe AP** 1995 The amino-terminal tails of the core histones and the translational position of the TATA box determine TBP/TFIIA association with nucleosomal DNA. Nucleic Acids Res 23:4557-64.

27. **Garel A, Axel R** 1976 Selective digestion of transcriptionally active ovalbumin genes from oviduct nuclei. Proc Natl Acad Sci U S A 73:3966-70.

28. **Weintraub H, Groudine M** 1976 Chromosomal subunits in active genes have an altered conformation. Science 193:848-56.

29. **Delcuve GP, Davie JR** 1989 Chromatin structure of erythroid-specific genes of immature and mature chicken erythrocytes. Biochem J 263:179-86.

30. **Kingston RE, Narlikar GJ** 1999 ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. Genes Dev 13:2339-52.

31. **Lorch Y, Zhang M, Kornberg RD** 1999 Histone octamer transfer by a chromatin-remodeling complex. Cell 96:389-92.

32. **Ito T, Bulger M, Pazin MJ, Kobayashi R, Kadonaga JT** 1997 ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. Cell 90:145-55.

33. **Hamiche A, Sandaltzopoulos R, Gdula DA, Wu C** 1999 ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF. Cell 97:833-42.

34. **Langst G, Bonte EJ, Corona DF, Becker PB** 1999 Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. Cell 97:843-52.

35. **Tsukiyama T, Palmer J, Landel CC, Shiloach J, Wu C** 1999 Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. Genes Dev 13:686-97.

36. **Varga-Weisz PD, Blank TA, Becker PB** 1995 Energy-dependent chromatin accessibility and nucleosome mobility in a cell-free system. Embo J 14:2209-16.

37. **Lorch Y, Cairns BR, Zhang M, Kornberg RD** 1998 Activated RSC-nucleosome complex and persistently altered form of the nucleosome. Cell 94:29-34.

38. **Schnitzler G, Sif S, Kingston RE** 1998 Human SWI/SNF interconverts a nucleosome between its base state and a stable remodeled state. Cell 94:17-27.

39. **Ikeda K, Steger DJ, Eberharter A, Workman JL** 1999 Activation domain-specific and general transcription stimulation by native histone acetyltransferase complexes. Mol Cell Biol 19:855-63.

40. **Nightingale KP, Wellinger RE, Sogo JM, Becker PB** 1998 Histone acetylation facilitates RNA polymerase II transcription of the Drosophila hsp26 gene in chromatin. Embo J 17:2865-76.

41. **Tse C, Sera T, Wolffe AP, Hansen JC** 1998 Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. Mol Cell Biol 18:4629-38.

42. **Noma K, Allis CD, Grewal SI** 2001 Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. Science 293:1150-5.

43. **Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T** 2002 Active genes are tri-methylated at K4 of histone H3. Nature 419:407-11.

44. **Mahadevan LC, Willis AC, Barratt MJ** 1991 Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid, and protein synthesis inhibitors. Cell 65:775-83.

45. **Chadee DN, Hendzel MJ, Tylipski CP, Allis CD, Bazett-Jones DP, Wright JA, Davie JR** 1999 Increased Ser-10 phosphorylation of histone H3 in mitogen-stimulated and oncogene-transformed mouse fibroblasts. J Biol Chem 274:24914-20.

46. **Walia H, Chen HY, Sun JM, Holth LT, Davie JR** 1998 Histone acetylation is required to maintain the unfolded nucleosome structure associated with transcribing DNA. J Biol Chem 273:14516-22.

47. **Garcia-Ramirez M, Rocchini C, Ausio J** 1995 Modulation of chromatin folding by histone acetylation. J Biol Chem 270:17923-8.

48. **Ogryzko VV, Schiltz RL, Russanova V, Howard BH, Nakatani Y** 1996 The transcriptional coactivators p300 and CBP are histone acetyltransferases. Cell 87:953-9.

49. **Spencer TE, Jenster G, Burcin MM, Allis CD, Zhou J, Mizzen CA, McKenna NJ, Onate SA, Tsai SY, Tsai MJ, O'Malley BW** 1997 Steroid receptor coactivator-1 is a histone acetyltransferase. Nature 389:194-8.

50. **Strahl BD, Allis CD** 2000 The language of covalent histone modifications. Nature 403:41-5.

51. **Orphanides G, Lagrange T, Reinberg D** 1996 The general transcription factors of RNA polymerase II. Genes Dev 10:2657-83.

52. **Nikolov DB, Burley SK** 1997 RNA polymerase II transcription initiation: a structural view. Proc Natl Acad Sci U S A 94:15-22.

53. **Baltimore D** 2001 Our genome unveiled. Nature 409:814-6.

54. **Guy LG, Kothary R, Wall L** 1997 Position effects in mice carrying a lacZ transgene in cis with the beta-globin LCR can be explained by a graded model. Nucleic Acids Res 25:4400-7.

55. **Sawado T, Halow J, Bender MA, Groudine M** 2003 The beta -globin locus control region (LCR) functions primarily by enhancing the transition from transcription initiation to elongation. Genes Dev 17:1009-18.

56. **Walters MC, Fiering S, Eidemiller J, Magis W, Groudine M, Martin DI** 1995 Enhancers increase the probability but not the level of gene expression. Proc Natl Acad Sci U S A 92:7125-9.

57. **Walters MC, Magis W, Fiering S, Eidemiller J, Scalzo D, Groudine M, Martin DI** 1996 Transcriptional enhancers act in cis to suppress position-effect variegation. Genes Dev 10:185-95.

58. **Weintraub H** 1988 Formation of stable transcription complexes as assayed by analysis of individual templates. Proc Natl Acad Sci U S A 85:5819-23.

59. **Blackwood EM, Kadonaga JT** 1998 Going the distance: a current view of enhancer action. Science 281:61-3.

60. **Fiering S, Whitelaw E, Martin DI** 2000 To be or not to be active: the stochastic nature of enhancer action. Bioessays 22:381-7.

61. **Li H, Capetanaki Y** 1994 An E box in the desmin promoter cooperates with the E box and MEF-2 sites of a distal enhancer to direct muscle-specific transcription. Embo J 13:3580-9.

62. **Magis W, Fiering S, Groudine M, Martin DI** 1996 An upstream activator of transcription coordinately increases the level and epigenetic stability of gene expression. Proc Natl Acad Sci U S A 93:13914-8.

63. **Swamynathan SK, Piatigorsky J** 2002 Orientation-dependent influence of an intergenic enhancer on the promoter activity of the divergently transcribed mouse Shsp/alpha B-crystallin and Mkbp/HspB2 genes. J Biol Chem 277:49700-6.

64. **Elgin SC** 1988 The formation and function of DNase I hypersensitive sites in the process of gene activation. J Biol Chem 263:19259-62.

65. **Gross DS, Garrard WT** 1988 Nuclease hypersensitive sites in chromatin. Annu Rev Biochem 57:159-97.

66. **Forrester WC, Epner E, Driscoll MC, Enver T, Brice M, Papayannopoulou T, Groudine M** 1990 A deletion of the human beta-globin locus activation region causes a major alteration in chromatin structure and replication across the entire beta-globin locus. Genes Dev 4:1637-49.

67. **Grosveld F, van Assendelft GB, Greaves DR, Kollias G** 1987 Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell 51:975-85.

68. **Bender MA, Bulger M, Close J, Groudine M** 2000 Beta-globin gene switching and DNase I sensitivity of the endogenous beta-globin locus in mice do not require the locus control region. Mol Cell 5:387-93.

69. **Epner E, Reik A, Cimbora D, Telling A, Bender MA, Fiering S, Enver T, Martin DI, Kennedy M, Keller G, Groudine M** 1998 The beta-globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse beta-globin locus. Mol Cell 2:447-55.

70. **Reik A, Telling A, Zitnik G, Cimbora D, Epner E, Groudine M** 1998 The locus control region is necessary for gene expression in the human beta-globin locus but not the maintenance of an open chromatin structure in erythroid cells. Mol Cell Biol 18:5992-6000.

71. **Schubeler D, Francastel C, Cimbora DM, Reik A, Martin DI, Groudine M** 2000 Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human beta-globin locus. Genes Dev 14:940-50.

72. **Schubeler D, Groudine M, Bender MA** 2001 The murine beta-globin locus control region regulates the rate of transcription but not the hyperacetylation of histones at the active genes. Proc Natl Acad Sci U S A 98:11432-7.

73. **Bharadwaj RR, Trainor CD, Pasceri P, Ellis J** 2003 LCR-regulated transgene expression levels depend on the Oct-1 site in the AT-rich region of beta -globin intron-2. Blood 101:1603-10.

74. **Carter D, Chakalova L, Osborne CS, Dai YF, Fraser P** 2002 Long-range chromatin regulatory interactions in vivo. Nat Genet 32:623-6.

75. **Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W** 2002 Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell 10:1453-65.

76. **Johnson KD, Grass JA, Park C, Im H, Choi K, Bresnick EH** 2003 Highly restricted localization of RNA polymerase II within a locus control region of a tissue-specific chromatin domain. Mol Cell Biol 23:6484-93.

77. **Rojo F** 2001 Mechanisms of transcriptional repression. Curr Opin Microbiol 4:145-51.

78. **Jepsen K, Rosenfeld MG** 2002 Biological roles and mechanistic actions of co-repressor complexes. J Cell Sci 115:689-98.

79. **Courey AJ, Jia S** 2001 Transcriptional repression: the long and the short of it. Genes Dev 15:2786-96.

80. **Johnson AD** 1995 The price of repression. Cell 81:655-8.

81. **Gaston K, Jayaraman PS** 2003 Transcriptional repression in eukaryotes: repressors and repression mechanisms. Cell Mol Life Sci 60:721-41.

82. **Barsh GS, Seeburg PH, Gelinas RE** 1983 The human growth hormone gene family: structure and evolution of the chromosomal locus. Nucleic Acids Res 11:3939-58.

83. **Chen EY, Liao YC, Smith DH, Barrera-Saldana HA, Gelinas RE, Seeburg PH** 1989 The human growth hormone locus: nucleotide sequence, biology, and evolution. Genomics 4:479-97.

84. **Hirt H, Kimelman J, Birnbaum MJ, Chen EY, Seeburg PH, Eberhardt NL, Barta A** 1987 The human growth hormone gene locus: structure, evolution, and allelic variations. DNA 6:59-70.

85. **Chopin LK, Veveris-Lowe TL, Philipps AF, Herington AC** 2002 Co-expression of GH and GHR isoforms in prostate cancer cell lines. Growth Horm IGF Res 12:126-36.

86. **Kyle CV, Evans MC, Odell WD** 1981 Growth hormone-like material in normal human tissues. J Clin Endocrinol Metab 53:1138-44.

87. **Mol JA, Henzen-Logmans SC, Hageman P, Misdorp W, Blankenstein MA, Rijnberk A** 1995 Expression of the gene encoding growth hormone in the human mammary gland. J Clin Endocrinol Metab 80:3094-6.

88. **Palmetshofer A, Zechner D, Luger TA, Barta A** 1995 Splicing variants of the human growth hormone mRNA: detection in pituitary, mononuclear cells and dermal fibroblasts. Mol Cell Endocrinol 113:225-34.

89. **Schwarzler P, Untergasser G, Hermann M, Dirnhofer S, Abendstein B, Madersbacher S, Berger P** 1997 Selective growth hormone/placental lactogen gene transcription and hormone production in pre- and postmenopausal human ovaries. J Clin Endocrinol Metab 82:3337-41.

90. **Untergasser G, Hermann M, Rumpold H, Berger P** 1998 Complex alternative splicing of the GH-V gene in the human testis. Eur J Endocrinol 139:424-7.

91. **Untergasser G, Kranewitter W, Schwarzler P, Madersbacher S, Dirnhofer S, Berger P** 1997 Organ-specific expression pattern of the human growth hormone/placental lactogen gene-cluster in the testis. Mol Cell Endocrinol 130:53-60.

92. **Costoya JA, Vidal A, Garcia-Barros M, Arce V, Devesa J** 1996 Expression of the human growth hormone normal gene (hGH-N) in proliferating and differentiated HL-60 cells. Exp Cell Res 228:164-7.

93. **Hattori N, Ikekubo K, Ishihara T, Moridera K, Hino M, Kurahachi H** 1994 Spontaneous growth hormone (GH) secretion by unstimulated human lymphocytes and the effects of GH-releasing hormone and somatostatin. J Clin Endocrinol Metab 79:1678-80.

94. **Kao TL, Harbour DV, Meyer WJ, 3rd** 1992 Immunoreactive growth hormone production by cultured lymphocytes. Ann N Y Acad Sci 650:179-81.

95. **Kao TL, Supowit SC, Thompson EA, Meyer WJ, 3rd** 1992 Immunoreactive growth hormone production by human lymphocyte cell lines. Cell Mol Neurobiol 12:483-98.

96. **Kooijman R, Berus D, Malur A, Delhase M, Hooghe-Peters EL** 1997 Human neutrophils express GH-N gene transcripts and the pituitary transcription factor Pit-1b. Endocrinology 138:4481-4.

97. **Lytras A, Quan N, Vrontakis ME, Shaw JE, Cattini PA, Friesen HG** 1993 Growth hormone expression in human Burkitt lymphoma serum-free Ramos cell line. Endocrinology 132:620-8.

98. **Maggiano N, Piantelli M, Ricci R, Larocca LM, Capelli A, Ranelletti FO** 1994 Detection of growth hormone-producing cells in human thymus by immunohistochemistry and non-radioactive in situ hybridization. J Histochem Cytochem 42:1349-54.

99. **Varma S, Sabharwal P, Sheridan JF, Malarkey WB** 1993 Growth hormone secretion by human peripheral blood mononuclear cells detected by an enzyme-linked immunoplaque assay. J Clin Endocrinol Metab 76:49-53.

100. **Wu H, Devi R, Malarkey WB** 1996 Localization of growth hormone messenger ribonucleic acid in the human immune system--a Clinical Research Center study. J Clin Endocrinol Metab 81:1278-82.

101. **Bulger M, Groudine M** 1999 Looping versus linking: toward a model for long-distance gene activation. Genes Dev 13:2465-77.

102. **Higgs DR** 1998 Do LCRs open chromatin domains? Cell 95:299-302.

103. **Cattini PA, Peritz LN, Anderson TR, Baxter JD, Eberhardt NL** 1986 The 5'-flanking sequences of the human growth hormone gene contain a cell-specific control element. DNA 5:503-9.

104. **Lefevre C, Imagawa M, Dana S, Grindlay J, Bodner M, Karin M** 1987 Tissue-specific expression of the human growth hormone gene is conferred in part by the binding of a specific trans-acting factor. Embo J 6:971-81.

105. **Bodner M, Karin M** 1987 A pituitary-specific trans-acting factor can stimulate transcription from the growth hormone promoter in extracts of nonexpressing cells. Cell 50:267-75.

106. **Bodner M, Castrillo JL, Theill LE, Deerinck T, Ellisman M, Karin M** 1988 The pituitary-specific transcription factor GHF-1 is a homeobox-containing protein. Cell 55:505-18.

107. **Ingraham HA, Chen RP, Mangalam HJ, Elsholtz HP, Flynn SE, Lin CR, Simmons DM, Swanson L, Rosenfeld MG** 1988 A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. Cell 55:519-29.

108. **Radovick S, Nations M, Du Y, Berg LA, Weintraub BD, Wondisford FE** 1992 A mutation in the POU-homeodomain of Pit-1 responsible for combined pituitary hormone deficiency. Science 257:1115-8.

109. **Li S, Crenshaw EB, 3rd, Rawson EJ, Simmons DM, Swanson LW, Rosenfeld MG** 1990 Dwarf locus mutants lacking three pituitary cell types result from mutations in the POU-domain gene pit-1. Nature 347:528-33.

110. **Chen RP, Ingraham HA, Treacy MN, Albert VR, Wilson L, Rosenfeld MG** 1990 Autoregulation of pit-1 gene expression mediated by two cis-active promoter elements. Nature 346:583-6.

111. **Peers B, Voz ML, Monget P, Mathy-Hartert M, Berwaer M, Belayew A, Martial JA** 1990 Regulatory elements controlling pituitary-specific expression of the human prolactin gene. Mol Cell Biol 10:4690-700.

112. **Fox SR, Jong MT, Casanova J, Ye ZS, Stanley F, Samuels HH** 1990 The homeodomain protein, Pit-1/GHF-1, is capable of binding to and activating cell-specific elements of both the growth hormone and prolactin gene promoters. Mol Endocrinol 4:1069-80.

113. **Steinfelder HJ, Hauser P, Nakayama Y, Radovick S, McClaskey JH, Taylor T, Weintraub BD, Wondisford FE** 1991 Thyrotropin-releasing hormone regulation of human TSHB expression: role of a pituitary-specific transcription

factor (Pit-1/GHF-1) and potential interaction with a thyroid hormone-inhibitory element. Proc Natl Acad Sci U S A 88:3130-4.

114. **Cattini PA, Eberhardt NL** 1987 Regulated expression of chimaeric genes containing the 5'-flanking regions of human growth hormone-related genes in transiently transfected rat anterior pituitary tumor cells. Nucleic Acids Res 15:1297-309.

115. **Cattini PA, Anderson TR, Baxter JD, Mellon P, Eberhardt NL** 1986 The human growth hormone gene is negatively regulated by triiodothyronine when transfected into rat pituitary tumor cells. J Biol Chem 261:13367-72.

116. **Cattini PA, Klassen M, Nachtigal M** 1988 Regulation of human chorionic somatomammotropin gene expression in rat pituitary tumour cells. Mol Cell Endocrinol 60:217-24.

117. **Nachtigal MW, Nickel BE, Klassen ME, Zhang WG, Eberhardt NL, Cattini PA** 1989 Human chorionic somatomammotropin and growth hormone gene expression in rat pituitary tumour cells is dependent on proximal promoter sequences. Nucleic Acids Res 17:4327-37.

118. **Lemaigre FP, Peers B, Lafontaine DA, Mathy-Hartert M, Rousseau GG, Belayew A, Martial JA** 1989 Pituitary-specific factor binding to the human prolactin, growth hormone, and placental lactogen genes. DNA 8:149-59.

119. **Nickel BE, Nachtigal MW, Bock ME, Cattini PA** 1991 Differential binding of rat pituitary-specific nuclear factors to the 5'-flanking region of pituitary and placental members of the human growth hormone gene family. Mol Cell Biochem 106:181-7.

120. **Jones BK, Monks BR, Liebhaber SA, Cooke NE** 1995 The human growth hormone gene is regulated by a multicomponent locus control region. Mol Cell Biol 15:7010-21.

121. **Bennani-Baiti IM, Asa SL, Song D, Iratni R, Liebhaber SA, Cooke NE** 1998 DNase I-hypersensitive sites I and II of the human growth hormone locus control region are a major developmental activator of somatotrope gene expression. Proc Natl Acad Sci U S A 95:10655-60.

122. **Su Y, Liebhaber SA, Cooke NE** 2000 The human growth hormone gene cluster locus control region supports position-independent pituitary- and placenta-specific expression in the transgenic mouse. J Biol Chem 275:7902-9.

123. **Ho Y, Elefant F, Cooke N, Liebhaber S** 2002 A defined locus control region determinant links chromatin domain acetylation with long-range gene activation. Mol Cell 9:291-302.

124. **Jin Y, Surabhi RM, Fresnoza A, Lytras A, Cattini PA** 1999 A role for A/T-rich sequences and Pit-1/GHF-1 in a distal enhancer located in the human growth hormone locus control region with preferential pituitary activity in culture and transgenic mice. Mol Endocrinol 13:1249-66.

125. **Shewchuk BM, Asa SL, Cooke NE, Liebhaber SA** 1999 Pit-1 binding sites at the somatotrope-specific DNase I hypersensitive sites I, II of the human growth hormone locus control region are essential for in vivo hGH-N gene activation. J Biol Chem 274:35725-33.

126. **Shewchuk BM, Liebhaber SA, Cooke NE** 2002 Specification of unique Pit-1 activity in the hGH locus control region. Proc Natl Acad Sci U S A 99:11784-9.

127. **Nickel BE, Cattini PA** 1996 Nuclease sensitivity of the human growth hormone-chorionic somatomammotropin locus in pituitary and placenta suggest different mechanisms for tissue-specific regulation. Mol Cell Endocrinol 118:155-62.

128. **Hjelle BL, Phillips JA, 3rd, Seeburg PH** 1982 Relative levels of methylation in human growth hormone and chorionic somatomammotropin genes in expressing and non-expressing tissues. Nucleic Acids Res 10:3459-74.

129. **Elefant F, Cooke NE, Liebhaber SA** 2000 Targeted recruitment of histone acetyltransferase activity to a locus control region. J Biol Chem 275:13827-34.

130. **Norquay LD, Yang X, Sheppard P, Gregoire S, Dodd JG, Reith W, Cattini PA** 2003 RFX1 and NF-1 Associate with P Sequences of the Human Growth Hormone Locus in Pituitary Chromatin. Mol Endocrinol 17:1027-38.

131. **Scully KM, Jacobson EM, Jepsen K, Lunyak V, Viadiu H, Carriere C, Rose DW, Hooshmand F, Aggarwal AK, Rosenfeld MG** 2000 Allosteric effects of Pit-1 DNA sites on long-term repression in cell type specification. Science 290:1127-31.

132. **Xu L, Lavinsky RM, Dasen JS, Flynn SE, McInerney EM, Mullen TM, Heinzel T, Szeto D, Korzus E, Kurokawa R, Aggarwal AK, Rose DW, Glass CK, Rosenfeld MG** 1998 Signal-specific co-activator domain requirements for Pit-1 activation. Nature 395:301-6.

133. **Nachtigal MW, Nickel BE, Cattini PA** 1993 Pituitary-specific repression of placental members of the human growth hormone gene family. A possible mechanism for locus regulation. J Biol Chem 268:8473-9.

134. **Richardson BD, Langland RA, Bachurski CJ, Richards RG, Kessler CA, Cheng YH, Handwerger S** 2000 Activator protein-2 regulates human placental lactogen gene expression. Mol Cell Endocrinol 160:183-92.

135. **Fitzpatrick SL, Walker WH, Saunders GF** 1990 DNA sequences involved in the transcriptional activation of a human placental lactogen gene. Mol Endocrinol 4:1815-26.

136. **Rogers BL, Sobnosky MG, Saunders GF** 1986 Transcriptional enhancer within the human placental lactogen and growth hormone multigene cluster. Nucleic Acids Res 14:7647-59.

137. **Jacquemin P, Oury C, Peers B, Morin A, Belayew A, Martial JA** 1994 Characterization of a single strong tissue-specific enhancer downstream from the three human genes encoding placental lactogen. Mol Cell Biol 14:93-103.

138. **Lytras A, Cattini PA** 1994 Human chorionic somatomammotropin gene enhancer activity is dependent on the blockade of a repressor mechanism. Mol Endocrinol 8:478-89.

139. **Walker WH, Fitzpatrick SL, Saunders GF** 1990 Human placental lactogen transcriptional enhancer. Tissue specificity and binding with specific proteins. J Biol Chem 265:12940-8.

140. **Lytras A, Surabhi RM, Zhang JF, Jin Y, Cattini PA** 1996 "Repair' of the chorionic somatomammotropin-A "enhancer' region reveals a novel functional element in the chorionic somatomammotropin-B enhancer. Mol Cell Endocrinol 119:1-10.

141. **Jacquemin P, Alsat E, Oury C, Belayew A, Muller M, Evain-Brion D, Martial JA** 1996 The enhancers of the human placental lactogen B, A, and L

genes: progressive activation during in vitro trophoblast differentiation and importance of the DF-3 element in determining their respective activities. DNA Cell Biol 15:845-54.

142. **Jacquemin P, Oury C, Belayew A, Martial JA** 1994 A TEF-1 binding motif that interacts with a placental protein is important for the transcriptional activity of the hCS-B enhancer. DNA Cell Biol 13:1037-45.

143. **Jiang SW, Eberhardt NL** 1994 The human chorionic somatomammotropin gene enhancer is composed of multiple DNA elements that are homologous to several SV40 enhansons. J Biol Chem 269:10384-92.

144. **Jiang SW, Eberhardt NL** 1995 Involvement of a protein distinct from transcription enhancer factor-1 (TEF-1) in mediating human chorionic somatomammotropin gene enhancer function through the GT-IIC enhanson in choriocarcinoma and COS cells. J Biol Chem 270:13906-15.

145. **Jiang SW, Trujillo MA, Eberhardt NL** 1997 Human chorionic somatomammotropin enhancer function is mediated by cooperative binding of TEF-1 and CSEF-1 to multiple, low-affinity binding sites. Mol Endocrinol 11:1223-32.

146. **Jacquemin P, Martial JA, Davidson I** 1997 Human TEF-5 is preferentially expressed in placenta and binds to multiple functional elements of the human chorionic somatomammotropin-B gene enhancer. J Biol Chem 272:12928-37.

147. **Jiang SW, Wu K, Eberhardt NL** 1999 Human placental TEF-5 transactivates the human chorionic somatomammotropin gene enhancer. Mol Endocrinol 13:879-89.

148. **Brar AK, Richards RG, Cheng YH, Richardson B, Kanda Y, Handwerger S** 2000 Mitogen-activated protein kinase activates human placental lactogen-B enhancer by an NF-IL6-dependent pathway. Endocrine 12:47-52.

149. **Jiang SW, Shepard AR, Eberhardt NL** 1995 An initiator element is required for maximal human chorionic somatomammotropin gene promoter and enhancer function. J Biol Chem 270:3683-92.

150. **Elefant F, Su Y, Liebhaber SA, Cooke NE** 2000 Patterns of histone acetylation suggest dual pathways for gene activation by a bifunctional locus control region. Embo J 19:6814-22.

151. **Norquay LD, Jin Y, Surabhi RM, Gietz RD, Tanese N, Cattini PA** 2001 A member of the nuclear factor-1 family is involved in the pituitary repression of the human placental growth hormone genes. Biochem J 354:387-95.

152. **Donjacour AA, Rosales A, Higgins SJ, Cunha GR** 1990 Characterization of antibodies to androgen-dependent secretory proteins of the mouse dorsolateral prostate. Endocrinology 126:1343-54.

153. **Mermod N, O'Neill EA, Kelly TJ, Tjian R** 1989 The proline-rich transcriptional activator of CTF/NF-I is distinct from the replication and DNA binding domain. Cell 58:741-53.

154. **Herrero Sanchez C, Reith W, Silacci P, Mach B** 1992 The DNA-binding defect observed in major histocompatibility complex class II regulatory mutants concerns only one member of a family of complexes binding to the X boxes of class II promoters. Mol Cell Biol 12:4076-83.

155. **Reith W, Ucla C, Barras E, Gaud A, Durand B, Herrero-Sanchez C, Kobr M, Mach B** 1994 RFX1, a transactivator of hepatitis B virus enhancer I, belongs to a novel family of homodimeric and heterodimeric DNA-binding proteins. Mol Cell Biol 14:1230-44.

156. **Tashjian AH, Jr., Yasumura Y, Levine L, Sato GH, Parker ML** 1968 Establishment of clonal strains of rat pituitary tumor cells that secrete growth hormone. Endocrinology 82:342-52.

157. **Leite V, Cardoso EA, Bock ME, Sobrinho LG, Cattini PA** 1996 A possible role for D8/PSF-A-like sequences in lactotroph versus somatotroph expression of the human prolactin gene. J Endocrinol 149:473-83.

158. **Pattillo RA, Gey GO, Delfs E, Huang WY, Hause L, Garancis DJ, Knoth M, Amatruda J, Bertino J, Friesen HG, Mattingly RF** 1971 The hormone-synthesizing trophoblastic cell in vitro: a model for cancer research and placental hormone synthesis. Ann N Y Acad Sci 172:288-98.

159. **Kohler PO, Bridson WE** 1971 Isolation of hormone-producing clonal lines of human choriocarcinoma. J Clin Endocrinol Metab 32:683-7.

160. **Nickel BE, Cattini PA** 1991 Tissue-specific expression and thyroid hormone regulation of the endogenous placental growth hormone variant and chorionic somatomammotropin genes in a human choriocarcinoma cell line. Endocrinology 128:2353-9.

161. **Cornille F, Emery P, Schuler W, Lenoir C, Mach B, Roques BP, Reith W** 1998 DNA binding properties of a chemically synthesized DNA binding domain of hRFX1. Nucleic Acids Res 26:2143-9.

162. **Osada S, Daimon S, Nishihara T, Imagawa M** 1996 Identification of DNA binding-site preferences for nuclear factor I-A. FEBS Lett 390:44-6.

163. **Morency CA, Neumann JR, Russian KO** 1987 A novel rapid assay for chloramphenicol acetyltransferase gene expression. Biotechniques 5:444-448

164. **Nickel BE, Kardami E, Cattini PA** 1990 Differential expression of human placental growth-hormone variant and chorionic somatomammotropin in culture. Biochem J 267:653-8.

165. **Quandt K, Frech K, Karas H, Wingender E, Werner T** 1995 MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res 23:4878-84.

166. **Wingender E, Chen X, Hehl R, Karas H, Liebich I, Matys V, Meinhardt T, Pruss M, Reuter I, Schacherer F** 2000 TRANSFAC: an integrated system for gene expression regulation. Nucleic Acids Res 28:316-9.

167. **Dignam JD, Lebovitz RM, Roeder RG** 1983 Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res 11:1475-89.

168. **Nordeen SK** 1988 Luciferase reporter gene vectors for analysis of promoters and enhancers. Biotechniques 6:454-8.

169. **Gorman CM, Merlino GT, Willingham MC, Pastan I, Howard BH** 1982 The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eukaryotic cells by DNA-mediated transfection. Proc Natl Acad Sci U S A 79:6777-81.

170. **Osada S, Matsubara T, Daimon S, Terazu Y, Xu M, Nishihara T, Imagawa M** 1999 Expression, DNA-binding specificity and transcriptional regulation of nuclear factor 1 family proteins from rat. Biochem J 342:189-98.

171. **Siegrist CA, Durand B, Emery P, David E, Hearing P, Mach B, Reith W** 1993 RFX1 is identical to enhancer factor C and functions as a transactivator of the hepatitis B virus enhancer. Mol Cell Biol 13:6375-84.

172. **Jiang SW, Eberhardt NL** 1997 The human chorionic somatomammotropin enhancers form a composite silencer in pituitary cells in vitro. Mol Endocrinol 11:1233-44.

173. **Surabhi RM** 1999 Studies on the role of the distal sequences involved in the expression of the human growth hormone / chorionic somatomammotropin gene family. Physiology. University of Manitoba, Winnipeg

174. **Clevidence DE, Overdier DG, Tao W, Qian X, Pani L, Lai E, Costa RH** 1993 Identification of nine tissue-specific transcription factors of the hepatocyte nuclear factor 3/forkhead DNA-binding-domain family. Proc Natl Acad Sci U S A 90:3948-52.

175. **Sutton J, Costa R, Klug M, Field L, Xu D, Largaespada DA, Fletcher CF, Jenkins NA, Copeland NG, Klemsz M, Hromas R** 1996 Genesis, a winged helix transcriptional repressor with expression restricted to embryonic stem cells. J Biol Chem 271:23126-33.

176. **Hromas R, Ye H, Spinella M, Dmitrovsky E, Xu D, Costa RH** 1999 Genesis, a Winged Helix transcriptional repressor, has embryonic expression limited to the neural crest, and stimulates proliferation in vitro in a neural development model. Cell Tissue Res 297:371-82.

177. **Labosky PA, Kaestner KH** 1998 The winged helix transcription factor Hfh2 is expressed in neural crest and spinal cord during mouse development. Mech Dev 76:185-90.

178. **Lints TJ, Parsons LM, Hartley L, Lyons I, Harvey RP** 1993 Nkx-2.5: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants. Development 119:419-31.

179. **Komuro I, Izumo S** 1993 Csx: a murine homeobox-containing gene specifically expressed in the developing heart. Proc Natl Acad Sci U S A 90:8145-9.

180. **Bodmer R** 1993 The gene tinman is required for specification of the heart and visceral muscles in Drosophila. Development 118:719-29.

181. **Gronostajski RM** 2000 Roles of the NFI/CTF gene family in transcription and development. Gene 249:31-45.

182. **Rajas F, Delhase M, De La Hoya M, Verdood P, Castrillo JL, Hooghe-Peters EL** 1998 Nuclear factor 1 regulates the distal silencer of the human PIT1/GHF1 gene. Biochem J 333:77-84.

183. **Chou SJ, Lai KN, Chung B** 1996 Characterization of the upstream sequence of the human CYP11A1 gene for cell type-specific expression. J Biol Chem 271:22125-9.

184. **Gronostajski RM, Adhya S, Nagata K, Guggenheimer RA, Hurwitz J** 1985 Site-specific DNA binding of nuclear factor I: analyses of cellular binding sites. Mol Cell Biol 5:964-71.

185. **Meisterernst M, Gander I, Rogge L, Winnacker EL** 1988 A quantitative analysis of nuclear factor I/DNA interactions. Nucleic Acids Res 16:4419-35.

186. **Gronostajski RM** 1986 Analysis of nuclear factor I binding to DNA using degenerate oligonucleotides. Nucleic Acids Res 14:9117-32.

187. **Gronostajski RM** 1987 Site-specific DNA binding of nuclear factor I: effect of the spacer region. Nucleic Acids Res 15:5545-59.

188. **Emery P, Durand B, Mach B, Reith W** 1996 RFX proteins, a novel family of DNA binding proteins conserved in the eukaryotic kingdom. Nucleic Acids Res 24:803-7.

189. **Gajiwala KS, Chen H, Cornille F, Roques BP, Reith W, Mach B, Burley SK** 2000 Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding. Nature 403:916-21.

190. **Emery P, Strubin M, Hofmann K, Bucher P, Mach B, Reith W** 1996 A consensus motif in the RFX DNA binding domain and binding domain mutants with altered specificity. Mol Cell Biol 16:4486-94.

191. **Iwama A, Pan J, Zhang P, Reith W, Mach B, Tenen DG, Sun Z** 1999 Dimeric RFX proteins contribute to the activity and lineage specificity of the interleukin-5 receptor alpha promoter through activation and repression domains. Mol Cell Biol 19:3940-50.

192. **Jones KA, Kadonaga JT, Rosenfeld PJ, Kelly TJ, Tjian R** 1987 A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell 48:79-89.

193. **Kruse U, Qian F, Sippel AE** 1991 Identification of a fourth nuclear factor I gene in chicken by cDNA cloning: NFI-X. Nucleic Acids Res 19:6641.

194. **Rupp RA, Kruse U, Multhaup G, Gobel U, Beyreuther K, Sippel AE** 1990 Chicken NFI/TGGCA proteins are encoded by at least three independent genes: NFI-A, NFI-B and NFI-C with homologues in mammalian genomes. Nucleic Acids Res 18:2607-16.

195. **Qian F, Kruse U, Lichter P, Sippel AE** 1995 Chromosomal localization of the four genes (NFIA, B, C, and X) for the human transcription factor nuclear factor I by FISH. Genomics 28:66-73.

196. **Cooke DW, Lane MD** 1999 Transcription factor NF1 mediates repression of the GLUT4 promoter by cyclic-AMP. Biochem Biophys Res Commun 260:600-4.

197. **Cooke DW, Lane MD** 1999 The transcription factor nuclear factor I mediates repression of the GLUT4 promoter by insulin. J Biol Chem 274:12917-24.

198. **Osada S, Ikeda T, Xu M, Nishihara T, Imagawa M** 1997 Identification of the transcriptional repression domain of nuclear factor 1-A. Biochem Biophys Res Commun 238:744-7.

199. **Singh J, Murata K, Itahana Y, Desprez PY** 2002 Constitutive expression of the Id-1 promoter in human metastatic breast cancer cells is linked with the loss of NF-1/Rb/HDAC-1 transcription repressor complex. Oncogene 21:1812-22.

200. **Courtois SJ, Lafontaine DA, Lemaigre FP, Durviaux SM, Rousseau GG** 1990 Nuclear factor-I and activator protein-2 bind in a mutually exclusive way to overlapping promoter sequences and trans-activate the human growth hormone gene. Nucleic Acids Res 18:57-64.

201. **Degterev A, Foster JA** 1999 The role of NF-1 factors in regulation of elastin gene transcription. Matrix Biol 18:295-307.

202. **Roulet E, Armentero MT, Krey G, Corthesy B, Dreyer C, Mermod N, Wahli W** 1995 Regulation of the DNA-binding and transcriptional activities of Xenopus laevis NFI-X by a novel C-terminal domain. Mol Cell Biol 15:5552-62.

203. **Liu Y, Bernard HU, Apt D** 1997 NFI-B3, a novel transcriptional repressor of the nuclear factor I family, is generated by alternative RNA processing. J Biol Chem 272:10739-45.

204. **Blomquist P, Li Q, Wrange O** 1996 The affinity of nuclear factor 1 for its DNA site is drastically reduced by nucleosome organization irrespective of its rotational or translational position. J Biol Chem 271:153-9.

205. **Dedon PC, Soults JA, Allis CD, Gorovsky MA** 1991 A simplified formaldehyde fixation and immunoprecipitation technique for studying protein-DNA interactions. Anal Biochem 197:83-90.

206. **Orlando V** 2000 Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. Trends Biochem Sci 25:99-104.

207. **Orlando V, Strutt H, Paro R** 1997 Analysis of chromatin structure by in vivo formaldehyde cross-linking. Methods 11:205-14.

208. **Kuo MH, Allis CD** 1999 In vivo cross-linking and immunoprecipitation for studying dynamic Protein:DNA associations in a chromatin environment. Methods 19:425-33.

209. **Spencer VAaD, J.R.** 2002 Isolation of proteins Cross-linked to DNA by Formaldehyde. In: Walker JM (ed) The Protein Protocols Handbook 2nd ed. Humana Press Inc., Totowa, NJ:753-757

210. **Reinhold MI, McEwen DG, Naski MC** 2003 Fibroblast Growth Factor Receptor 3 Gene: Regulation by Serum Response Factor. Mol Endocrinol 16:16

211. **Ulaner GA, Yang Y, Hu JF, Li T, Vu TH, Hoffman AR** 2003 CTCF binding at the insulin-like growth factor-II (IGF2)/H19 imprinting control region is insufficient to regulate IGF2/H19 expression in human tissues. Endocrinology 144:4420-6.

212. **Woelfle J, Chia DJ, Rotwein P** 2003 Mechanisms of growth hormone action: identification of conserved Stat5 binding sites that mediate GH-induced insulin-like growth factor-I gene activation. J Biol Chem 7:7

213. **Follows GA, Tagoh H, Lefevre P, Morgan GJ, Bonifer C** 2003 Differential transcription factor occupancy but evolutionarily conserved chromatin features at the human and mouse M-CSF (CSF-1) receptor loci. Nucleic Acids Res 31:5805-16.

214. **Seo JB, Noh MJ, Yoo EJ, Park SY, Park J, Lee IK, Park SD, Kim JB** 2003 Functional characterization of the human resistin promoter with adipocyte determination- and differentiation-dependent factor 1/sterol regulatory element binding protein 1c and CCAAT enhancer binding protein-alpha. Mol Endocrinol 17:1522-33.

215. **Miyake A, Konishi M, Martin FH, Hernday NA, Ozaki K, Yamamoto S, Mikami T, Arakawa T, Itoh N** 1998 Structure and expression of a novel member, FGF-16, on the fibroblast growth factor family. Biochem Biophys Res Commun 243:148-52.

216. **Lemaigre FP, Courtois SJ, Durviaux SM, Egan CJ, LaFontaine DA, Rousseau GG** 1989 Analysis of cis- and trans-acting elements in the hormone-sensitive human somatotropin gene promoter. J Steroid Biochem 34:79-83.

217. **Lemaigre FP, Lafontaine DA, Courtois SJ, Durviaux SM, Rousseau GG** 1990 Sp1 can displace GHF-1 from its distal binding site and stimulate transcription from the growth hormone gene promoter. Mol Cell Biol 10:1811-4.

218. **Kamat A, Graves KH, Smith ME, Richardson JA, Mendelson CR** 1999 A 500-bp region, approximately 40 kb upstream of the human CYP19 (aromatase) gene, mediates placenta-specific expression in transgenic mice. Proc Natl Acad Sci U S A 96:4575-80.

219. **Kamat A, Mendelson CR** 2001 Identification of the regulatory regions of the human aromatase P450 (CYP19) gene involved in placenta-specific expression. J Steroid Biochem Mol Biol 79:173-80.

220. **Nagata K, Guggenheimer RA, Enomoto T, Lichy JH, Hurwitz J** 1982 Adenovirus DNA replication in vitro: identification of a host factor that stimulates

synthesis of the preterminal protein-dCMP complex. Proc Natl Acad Sci U S A 79:6438-42.

221. **Nagata K, Guggenheimer RA, Hurwitz J** 1983 Adenovirus DNA replication in vitro: synthesis of full-length DNA with purified proteins. Proc Natl Acad Sci U S A 80:4266-70.

222. **Armentero MT, Horwitz M, Mermod N** 1994 Targeting of DNA polymerase to the adenovirus origin of DNA replication by interaction with nuclear factor I. Proc Natl Acad Sci U S A 91:11537-41.

223. **Chen M, Mermod N, Horwitz MS** 1990 Protein-protein interactions between adenovirus DNA polymerase and nuclear factor I mediate formation of the DNA replication preinitiation complex. J Biol Chem 265:18634-42.

224. **Bosher J, Robinson EC, Hay RT** 1990 Interactions between the adenovirus type 2 DNA polymerase and the DNA binding domain of nuclear factor I. New Biol 2:1083-90.

225. **Mul YM, Van der Vliet PC** 1992 Nuclear factor I enhances adenovirus DNA replication by increasing the stability of a preinitiation complex. Embo J 11:751-60.

226. **Zorbas H, Rogge L, Meisterernst M, Winnacker EL** 1989 Hydroxyl radical footprints reveal novel structural features around the NF I binding site in adenovirus DNA. Nucleic Acids Res 17:7735-48.

227. **Kim JL, Nikolov DB, Burley SK** 1993 Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature 365:520-7.

228. **Osada S, Daimon S, Ikeda T, Nishihara T, Yano K, Yamasaki M, Imagawa M** 1997 Nuclear factor 1 family proteins bind to the silencer element in the rat glutathione transferase P gene. J Biochem (Tokyo) 121:355-63.

229. **Bois-Joyeux B, Danan JL** 1994 Members of the CAAT/enhancer-binding protein, hepatocyte nuclear factor-1 and nuclear factor-1 families can differentially modulate the activities of the rat alpha-fetoprotein promoter and enhancer. Biochem J 301:49-55.

230. **Hebbar PB, Archer TK** 2003 Nuclear factor 1 is required for both hormone-dependent chromatin remodeling and transcriptional activation of the mouse mammary tumor virus promoter. Mol Cell Biol 23:887-98.

231. **Lefevre P, Melnik S, Wilson N, Riggs AD, Bonifer C** 2003 Developmentally regulated recruitment of transcription factors and chromatin modification activities to chicken lysozyme cis-regulatory elements in vivo. Mol Cell Biol 23:4386-400.

232. **Fletcher CF, Jenkins NA, Copeland NG, Chaudhry AZ, Gronostajski RM** 1999 Exon structure of the nuclear factor I DNA-binding domain from C. elegans to mammals. Mamm Genome 10:390-6.

233. **Gounari F, De Francesco R, Schmitt J, van der Vliet P, Cortese R, Stunnenberg H** 1990 Amino-terminal domain of NF1 binds to DNA as a dimer and activates adenovirus DNA replication. Embo J 9:559-66.

234. **Imagawa M, Sakaue R, Tanabe A, Osada S, Nishihara T** 2000 Two nuclear localization signals are required for nuclear translocation of nuclear factor 1-A. FEBS Lett 484:118-24.

235. **Prado F, Vicent G, Cardalda C, Beato M** 2002 Differential role of the proline-rich domain of nuclear factor 1-C splice variants in DNA binding and transactivation. J Biol Chem 277:16383-90.

236. **Apt D, Liu Y, Bernard HU** 1994 Cloning and functional analysis of spliced isoforms of human nuclear factor I-X: interference with transcriptional activation by NFI/CTF in a cell-type specific manner. Nucleic Acids Res 22:3825-33.

237. **Kruse U, Sippel AE** 1994 The genes for transcription factor nuclear factor I give rise to corresponding splice variants between vertebrate species. J Mol Biol 238:860-5.

238. **Santoro C, Mermod N, Andrews PC, Tjian R** 1988 A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. Nature 334:218-24.

239. **Monaci P, Nuzzo M, Stampfli S, Tollervey D, De Simone V, Nicosia A** 1995 A complex interplay of positive and negative elements is responsible for the different transcriptional activity of liver NF1 variants. Mol Biol Rep 21:147-58.

240. **Nebl G, Cato AC** 1995 NFI/X proteins: a class of NFI family of transcription factors with positive and negative regulatory domains. Cell Mol Biol Res 41:85-95.

241. **Kim TK, Roeder RG** 1994 CTD-like sequences are important for transcriptional activation by the proline-rich activation domain of CTF1. Nucleic Acids Res 22:251.

242. **Meisterernst M, Rogge L, Foeckler R, Karaghiosoff M, Winnacker EL** 1989 Structural and functional organization of a porcine gene coding for nuclear factor I. Biochemistry 28:8191-200.

243. **Wendler W, Altmann H, Ludwig-Winnacker E** 1994 Transcriptional activation of NFI/CTF1 depends on a sequence motif strongly related to the carboxyterminal domain of RNA polymerase II. Nucleic Acids Res 22:2601-3.

244. **Xiao H, Lis JT, Greenblatt J, Friesen JD** 1994 The upstream activator CTF/NF1 and RNA polymerase II share a common element involved in transcriptional activation. Nucleic Acids Res 22:1966-73.

245. **Kruse U, Sippel AE** 1994 Transcription factor nuclear factor I proteins form stable homo- and heterodimers. FEBS Lett 348:46-50.

246. **Goyal N, Knox J, Gronostajski RM** 1990 Analysis of multiple forms of nuclear factor I in human and murine cell lines. Mol Cell Biol 10:1041-8.

247. **Jethanandani P, Goldberg E** 2001 ldhc expression in non-germ cell nuclei is repressed by NF-I binding. J Biol Chem 276:35414-21.

248. **Alevizopoulos A, Dusserre Y, Ruegg U, Mermod N** 1997 Regulation of the transforming growth factor beta-responsive transcription factor CTF-1 by calcineurin and calcium/calmodulin-dependent protein kinase IV. J Biol Chem 272:23597-605.

249. **Alevizopoulos A, Dusserre Y, Tsai-Pflugfelder M, von der Weid T, Wahli W, Mermod N** 1995 A proline-rich TGF-beta-responsive transcriptional activator interacts with histone H3. Genes Dev 9:3051-66.

250. **Kulkarni S, Gronostajski RM** 1996 Altered expression of the developmentally regulated NFI gene family during phorbol ester-induced differentiation of human leukemic cells. Cell Growth Differ 7:501-10.

251. **Chaudhry AZ, Lyons GE, Gronostajski RM** 1997 Expression patterns of the four nuclear factor I genes during mouse embryogenesis indicate a potential role in development. Dev Dyn 208:313-25.

252. **das Neves L, Duchala CS, Tolentino-Silva F, Haxhiu MA, Colmenares C, Macklin WB, Campbell CE, Butz KG, Gronostajski RM, Godinho F** 1999 Disruption of the murine nuclear factor I-A gene (Nfia) results in perinatal lethality, hydrocephalus, and agenesis of the corpus callosum. Proc Natl Acad Sci U S A 96:11946-51.

253. **Grunder A, Ebel TT, Mallo M, Schwarzkopf G, Shimizu T, Sippel AE, Schrewe H** 2002 Nuclear factor I-B (Nfib) deficient mice have severe lung hypoplasia. Mech Dev 112:69-77.

254. **Chaudhry AZ, Vitullo AD, Gronostajski RM** 1998 Nuclear factor I (NFI) isoforms differentially activate simple versus complex NFI-responsive promoters. J Biol Chem 273:18538-46.

255. **Chiang CM, Roeder RG** 1995 Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators. Science 267:531-6.

256. **Kim TK, Roeder RG** 1994 Proline-rich activator CTF1 targets the TFIIB assembly step during transcriptional activation. Proc Natl Acad Sci U S A 91:4170-4.

257. **Leahy P, Crawford DR, Grossman G, Gronostajski RM, Hanson RW** 1999 CREB binding protein coordinates the function of multiple transcription factors including nuclear factor I to regulate phosphoenolpyruvate carboxykinase (GTP) gene transcription. J Biol Chem 274:8813-22.

258. **Bernier D, Thomassin H, Allard D, Guertin M, Hamel D, Blaquiere M, Beauchemin M, LaRue H, Estable-Puig M, Belanger L** 1993 Functional analysis of developmentally regulated chromatin-hypersensitive domains carrying the alpha 1-fetoprotein gene promoter and the albumin/alpha 1-fetoprotein intergenic enhancer. Mol Cell Biol 13:1619-33.

259. **Nehls MC, Grapilon ML, Brenner DA** 1992 NF-I/Sp1 switch elements regulate collagen alpha 1(I) gene expression. DNA Cell Biol 11:443-52.

260. **Nehls MC, Rippe RA, Veloz L, Brenner DA** 1991 Transcription factors nuclear factor I and Sp1 interact with the murine collagen alpha 1 (I) promoter. Mol Cell Biol 11:4065-73.

261. **Yamada K, Tanaka T, Noguchi T** 1997 Members of the nuclear factor 1 family and hepatocyte nuclear factor 4 bind to overlapping sequences of the L-II element on the rat pyruvate kinase L gene promoter and regulate its expression. Biochem J 324:917-25.

262. **McCormick A, Brady H, Theill LE, Karin M** 1990 Regulation of the pituitary-specific homeobox gene GHF1 by cell-autonomous and environmental cues. Nature 345:829-32.

263. **O'Connor M, Bernard HU** 1995 Oct-1 activates the epithelial-specific enhancer of human papillomavirus type 16 via a synergistic interaction with NFI at a conserved composite regulatory element. Virology 207:77-88.

264. **Bandyopadhyay S, Gronostajski RM** 1994 Identification of a conserved oxidation-sensitive cysteine residue in the NFI family of DNA-binding proteins. J Biol Chem 269:29949-55.

265. **Morel Y, Barouki R** 2000 The repression of nuclear factor I/CCAAT transcription factor (NFI/CTF) transactivating domain by oxidative stress is mediated by a critical cysteine (Cys-427). Biochem J 348:235-40.

266. **Morel Y, Mermod N, Barouki R** 1999 An autoregulatory loop controlling CYP1A1 gene expression: role of H(2)O(2) and NFI. Mol Cell Biol 19:6825-32.

267. **Morel Y, Barouki R** 1998 Down-regulation of cytochrome P450 1A1 gene promoter by oxidative stress. Critical contribution of nuclear factor 1. J Biol Chem 273:26969-76.

268. **Morel Y, Coumoul X, Nalpas A, Barouki R** 2000 Nuclear factor I/CCAAT box transcription factor trans-activating domain is a negative sensor of cellular stress. Mol Pharmacol 58:1239-46.

269. **Morel Y, Barouki R** 1999 Repression of gene expression by oxidative stress. Biochem J 342:481-96.

270. **Dotzlaw H, Alkhalaf M, Murphy LC** 1992 Characterization of estrogen receptor variant mRNAs from human breast cancers. Mol Endocrinol 6:773-85.

271. **Doyle J, Hoffman S, Ucla C, Reith W, Mach B, Stubbs L** 1996 Locations of human and mouse genes encoding the RFX1 and RFX2 transcription factor proteins. Genomics 35:227-30.

272. **Morotomi-Yano K, Yano K, Saito H, Sun Z, Iwama A, Miki Y** 2002 Human regulatory factor X 4 (RFX4) is a testis-specific dimeric DNA-binding protein that cooperates with other human RFX members. J Biol Chem 277:836-42.

273. **Pugliatti L, Derre J, Berger R, Ucla C, Reith W, Mach B** 1992 The genes for MHC class II regulatory factors RFX1 and RFX2 are located on the short arm of chromosome 19. Genomics 13:1307-10.

274. **Reith W, Herrero-Sanchez C, Kobr M, Silacci P, Berte C, Barras E, Fey S, Mach B** 1990 MHC class II regulatory factor RFX has a novel DNA-binding domain and a functionally independent dimerization domain. Genes Dev 4:1528-40.

275. **Steimle V, Durand B, Barras E, Zufferey M, Hadam MR, Mach B, Reith W** 1995 A novel DNA-binding regulatory factor is mutated in primary MHC class II deficiency (bare lymphocyte syndrome). Genes Dev 9:1021-32.

276. **Huang M, Zhou Z, Elledge SJ** 1998 The DNA replication and damage checkpoint pathways induce transcription by inhibition of the Crt1 repressor. Cell 94:595-605.

277. **Wu SY, McLeod M** 1995 The sak1+ gene of Schizosaccharomyces pombe encodes an RFX family DNA-binding protein that positively regulates cyclic AMP-dependent protein kinase-mediated exit from the mitotic cell cycle. Mol Cell Biol 15:1479-88.

278. **Schmitt EK, Kuck U** 2000 The fungal CPCR1 protein, which binds specifically to beta-lactam biosynthesis genes, is related to human regulatory factor X transcription factors. J Biol Chem 275:9348-57.

279. **Swoboda P, Adler HT, Thomas JH** 2000 The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in C. elegans. Mol Cell 5:411-21.

280. **Durand B, Vandaele C, Spencer D, Pantalacci S, Couble P** 2000 Cloning and characterization of dRFX, the Drosophila member of the RFX family of transcription factors. Gene 246:285-93.

281. **Durand B, Kobr M, Reith W, Mach B** 1994 Functional complementation of major histocompatibility complex class II regulatory mutants by the purified X-box-binding protein RFX. Mol Cell Biol 14:6839-47.

282. **Durand B, Sperisen P, Emery P, Barras E, Zufferey M, Mach B, Reith W** 1997 RFXAP, a novel subunit of the RFX DNA binding complex is mutated in MHC class II deficiency. Embo J 16:1045-55.

283. **Reith W, Kobr M, Emery P, Durand B, Siegrist CA, Mach B** 1994 Cooperative binding between factors RFX and X2bp to the X and X2 boxes of MHC class II promoters. J Biol Chem 269:20020-5.

284. **Reith W, Siegrist CA, Durand B, Barras E, Mach B** 1994 Function of major histocompatibility complex class II promoters requires cooperative binding between factors RFX and NF-Y. Proc Natl Acad Sci U S A 91:554-8.

285. **Moreno CS, Emery P, West JE, Durand B, Reith W, Mach B, Boss JM** 1995 Purified X2 binding protein (X2BP) cooperatively binds the class II MHC X box region in the presence of purified RFX, the X box factor deficient in the bare lymphocyte syndrome. J Immunol 155:4313-21.

286. **Masternak K, Barras E, Zufferey M, Conrad B, Corthals G, Aebersold R, Sanchez JC, Hochstrasser DF, Mach B, Reith W** 1998 A gene encoding a novel RFX-associated transactivator is mutated in the majority of MHC class II deficiency patients. Nat Genet 20:273-7.

287. **Villard J, Peretti M, Masternak K, Barras E, Caretti G, Mantovani R, Reith W** 2000 A functionally essential domain of RFX5 mediates activation of major histocompatibility complex class II promoters by promoting cooperative binding between RFX and NF-Y. Mol Cell Biol 20:3364-76.

288. **David E, Garcia AD, Hearing P** 1995 Interaction of EF-C/RFX-1 with the inverted repeat of viral enhancer regions is required for transactivation. J Biol Chem 270:8353-60.

289. **Katan-Khaykovich Y, Shaul Y** 1998 RFX1, a single DNA-binding protein with a split dimerization domain, generates alternative complexes. J Biol Chem 273:24504-12.

290. **Tierney R, Kirby H, Nagra J, Rickinson A, Bell A** 2000 The Epstein-Barr virus promoter initiating B-cell transformation is activated by RFX proteins and the B-cell-specific activator protein BSAP/Pax5. J Virol 74:10458-67.

291. **Clark KL, Halay ED, Lai E, Burley SK** 1993 Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. Nature 364:412-20.

292. **Overdier DG, Porcella A, Costa RH** 1994 The DNA-binding specificity of the hepatocyte nuclear factor 3/forkhead domain is influenced by amino-acid residues adjacent to the recognition helix. Mol Cell Biol 14:2755-66.

293. **Sheng W, Rance M, Liao X** 2002 Structure comparison of two conserved HNF-3/fkh proteins HFH-1 and genesis indicates the existence of folding differences in their complexes with a DNA binding sequence. Biochemistry 41:3286-93.

294. **Nakayama A, Odajima T, Murakami H, Mori N, Takahashi M** 2001 Characterization of two promoters that regulate alternative transcripts in the microtubule-associated protein (MAP) 1A gene. Biochim Biophys Acta 1518:260-6.

295. **Lietard J, Theret N, Rehn M, Musso O, Dargere D, Pihlajaniemi T, Clement B** 2000 The promoter of the long variant of collagen XVIII, the precursor of endostatin, contains liver-specific regulatory elements. Hepatology 32:1377-85.

296. **Alvarez L, Sanchez-Gongora E, Mingorance J, Pajares MA, Mato JM** 1997 Characterization of rat liver-specific methionine adenosyltransferase gene promoter. Role of distal upstream cis-acting elements in the regulation of the transcriptional activity. J Biol Chem 272:22875-83.

297. **Tsay W, Lee YM, Lee SC, Shen MC, Chen PJ** 1997 Synergistic transactivation of HNF-1alpha, HNF-3, and NF-I contributes to the activation of the liver-specific protein C gene. DNA Cell Biol 16:569-77.

298. **Arenzana N, Rodriguez de Cordoba S** 1996 Promoter region of the human gene coding for beta-chain of C4b binding protein. Hepatocyte nuclear factor-3 and

nuclear factor-I/CTF transcription factors are required for efficient expression of C4BPB in HepG2 cells. J Immunol 156:168-75.

299. **Auge-Gouillou C, Petropoulos I, Zakin MM** 1993 Liver-enriched HNF-3 alpha and ubiquitous factors interact with the human transferrin gene enhancer. FEBS Lett 323:4-10.

300. **Ramji DP, Foka P** 2002 CCAAT/enhancer-binding proteins: structure, function and regulation. Biochem J 365:561-75.

301. **Webster NJ, Kong Y, Cameron KE, Resnik JL** 1994 An upstream element from the human insulin receptor gene promoter contains binding sites for C/EBP beta and NF-1. Diabetes 43:305-12.

302. **Ji C, Chen Y, Centrella M, McCarthy TL** 1999 Activation of the insulin-like growth factor-binding protein-5 promoter in osteoblasts by cooperative E box, CCAAT enhancer-binding protein, and nuclear factor-1 deoxyribonucleic acid-binding sequences. Endocrinology 140:4564-72.

303. **Aigueperse C, Val P, Pacot C, Darne C, Lalli E, Sassone-Corsi P, Veyssiere G, Jean C, Martinez A** 2001 SF-1 (steroidogenic factor-1), C/EBPbeta (CCAAT/enhancer binding protein), and ubiquitous transcription factors NF1 (nuclear factor 1) and Sp1 (selective promoter factor 1) are required for regulation of the mouse aldose reductase-like gene (AKR1B7) expression in adrenocortical cells. Mol Endocrinol 15:93-111.

304. **Thomassin H, Bois-Joyeux B, Delille R, Ikonomova R, Danan JL** 1996 Chicken ovalbumin upstream promoter-transcription factor, hepatocyte nuclear factor 3, and CCAAT/enhancer binding protein control the far-upstream enhancer of the rat alpha-fetoprotein gene. DNA Cell Biol 15:1063-74.

305. **Antes TJ, Levy-Wilson B** 2001 HNF-3 beta, C/EBP beta, and HNF-4 act in synergy to enhance transcription of the human apolipoprotein B gene in intestinal cells. DNA Cell Biol 20:67-74.

306. **Antes TJ, Goodart SA, Huynh C, Sullivan M, Young SG, Levy-Wilson B** 2000 Identification and characterization of a 315-base pair enhancer, located more than 55 kilobases 5' of the apolipoprotein B gene, that confers expression in the intestine. J Biol Chem 275:26637-48.

307. **Rodriguez-Antona C, Bort R, Jover R, Tindberg N, Ingelman-Sundberg M, Gomez-Lechon MJ, Castell JV** 2003 Transcriptional regulation of human CYP3A4 basal expression by CCAAT enhancer-binding protein alpha and hepatocyte nuclear factor-3 gamma. Mol Pharmacol 63:1180-9.

308. **Costa RH, Grayson DR, Darnell JE, Jr.** 1989 Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and alpha 1-antitrypsin genes. Mol Cell Biol 9:1415-25.

309. **Pierreux CE, Urso B, De Meyts P, Rousseau GG, Lemaigre FP** 1998 Inhibition by insulin of glucocorticoid-induced gene transcription: involvement of the ligand-binding domain of the glucocorticoid receptor and independence from the phosphatidylinositol 3-kinase and mitogen-activated protein kinase pathways. Mol Endocrinol 12:1343-54.

310. **Guo Y, Costa R, Ramsey H, Starnes T, Vance G, Robertson K, Kelley M, Reinbold R, Scholer H, Hromas R** 2002 The embryonic stem cell transcription factors Oct-4 and FoxD3 interact to regulate endodermal-specific promoter expression. Proc Natl Acad Sci U S A 99:3663-7.

311. **Raynal JF, Dugast C, Le Van Thai A, Weber MJ** 1998 Winged helix hepatocyte nuclear factor 3 and POU-domain protein brn-2/N-oct-3 bind overlapping sites on the neuronal promoter of human aromatic L-amino acid decarboxylase gene. Brain Res Mol Brain Res 56:227-37.

312. **Millevoi S, Thion L, Joseph G, Vossen C, Ghisolfi-Nieto L, Erard M** 2001 Atypical binding of the neuronal POU protein N-Oct3 to noncanonical DNA targets. Implications for heterodimerization with HNF-3 beta. Eur J Biochem 268:781-91.

313. **Belsham DD, Mellon PL** 2000 Transcription factors Oct-1 and C/EBPbeta (CCAAT/enhancer-binding protein-beta) are involved in the glutamate/nitric oxide/cyclic-guanosine 5'-monophosphate-mediated repression of mediated repression of gonadotropin-releasing hormone gene expression. Mol Endocrinol 14:212-28.

314. **Schaufele F** 1996 CCAAT/enhancer-binding protein alpha activation of the rat growth hormone promoter in pituitary progenitor GHFT1-5 cells. J Biol Chem 271:21484-9.

315. **Jacob KK, Stanley FM** 1999 CCAAT/enhancer-binding protein alpha is a physiological regulator of prolactin gene expression. Endocrinology 140:4542-50.

316. **Enwright JF, 3rd, Kawecki-Crook MA, Voss TC, Schaufele F, Day RN** 2003 A PIT-1 Homeodomain Mutant Blocks the Intranuclear Recruitment Of the CCAAT/Enhancer Binding Protein alpha Required for Prolactin Gene Transcription. Mol Endocrinol 17:209-22.

317. **Laniel MA, Poirier GG, Guerin SL** 2001 Nuclear factor 1 interferes with Sp1 binding through a composite element on the rat poly(ADP-ribose) polymerase promoter to modulate its activity in vitro. J Biol Chem 276:20766-73.

318. **O'Neill EA, Fletcher C, Burrow CR, Heintz N, Roeder RG, Kelly TJ** 1988 Transcription factor OTF-1 is functionally identical to the DNA replication factor NF-III. Science 241:1210-3.

319. **Pruijn GJ, van Driel W, van der Vliet PC** 1986 Nuclear factor III, a novel sequence-specific DNA-binding protein from HeLa cells stimulating adenovirus DNA replication. Nature 322:656-9.

320. **Rosenfeld PJ, O'Neill EA, Wides RJ, Kelly TJ** 1987 Sequence-specific interactions between cellular DNA-binding proteins and the adenovirus origin of DNA replication. Mol Cell Biol 7:875-86.

321. **Verrijzer CP, Strating M, Mul YM, van der Vliet PC** 1992 POU domain transcription factors from different subclasses stimulate adenovirus DNA replication. Nucleic Acids Res 20:6369-75.

322. **Forsberg EC, Downs KM, Christensen HM, Im H, Nuzzi PA, Bresnick EH** 2000 Developmentally dynamic histone acetylation pattern of a tissue-specific chromatin domain. Proc Natl Acad Sci U S A 97:14494-9.

323. **Cattini PA** 2002 Chorionic somatomammotropin. In Wiley Encyclopedia of Molecular Medicine. John Wiley and Sons Inc.:768-770

324. **Muller K, Mermod N** 2000 The histone-interacting domain of nuclear factor I activates simian virus 40 DNA replication in vivo. J Biol Chem 275:1645-50.

325. **Gao B, Jaffe H, Kunos G** 1998 Histone H1 isoforms purified from rat liver bind nonspecifically to the nuclear factor 1 recognition sequence and serve as generalized transcriptional repressors. Mol Cell Biochem 178:187-96.

326. **Ristiniemi J, Oikarinen J** 1989 Histone H1 binds to the putative nuclear factor I recognition sequence in the mouse alpha 2(I) collagen promoter. J Biol Chem 264:2164-74.

327. **Dusserre Y, Mermod N** 1992 Purified cofactors and histone H1 mediate transcriptional regulation by CTF/NF-I. Mol Cell Biol 12:5228-37.

328. **Hebbar PB, Archer TK** 2003 Chromatin remodeling by nuclear receptors. Chromosoma 111:495-504.

329. **Cordingley MG, Riegel AT, Hager GL** 1987 Steroid-dependent interaction of transcription factors with the inducible promoter of mouse mammary tumor virus in vivo. Cell 48:261-70.

330. **Archer TK, Cordingley MG, Wolford RG, Hager GL** 1991 Transcription factor access is mediated by accurately positioned nucleosomes on the mouse mammary tumor virus promoter. Mol Cell Biol 11:688-98.

331. **Prado F, Koop R, Beato M** 2002 Accurate chromatin organization of the mouse mammary tumor virus promoter determines the nature of the synergism between transcription factors. J Biol Chem 277:4911-7.

332. **Reith W, Satola S, Sanchez CH, Amaldi I, Lisowska-Grospierre B, Griscelli C, Hadam MR, Mach B** 1988 Congenital immunodeficiency with a regulatory defect in MHC class II gene expression lacks a specific HLA-DR promoter binding protein, RF-X. Cell 53:897-906.

333. **Masternak K, Muhlethaler-Mottet A, Villard J, Zufferey M, Steimle V, Reith W** 2000 CIITA is a transcriptional coactivator that is recruited to MHC class II promoters by multiple synergistic interactions with an enhanceosome complex. Genes Dev 14:1156-66.

334. **Fontes JD, Kanazawa S, Jean D, Peterlin BM** 1999 Interactions between the class II transactivator and CREB binding protein increase transcription of major histocompatibility complex class II genes. Mol Cell Biol 19:941-7.

335. **Kretsovali A, Agalioti T, Spilianakis C, Tzortzakaki E, Merika M, Papamatheakis J** 1998 Involvement of CREB binding protein in expression of major histocompatibility complex class II genes via interaction with the class II transactivator. Mol Cell Biol 18:6777-83.

336. **Jacobson EM, Li P, Leon-del-Rio A, Rosenfeld MG, Aggarwal AK** 1997 Structure of Pit-1 POU domain bound to DNA as a dimer: unexpected arrangement and flexibility. Genes Dev 11:198-212.

337. **Miller M, Shuman JD, Sebastian T, Dauter Z, Johnson PF** 2003 Structural basis for DNA recognition by the basic region leucine zipper transcription factor CCAAT/enhancer-binding protein alpha. J Biol Chem 278:15178-84.

338. **Luisi BF, Xu WX, Otwinowski Z, Freedman LP, Yamamoto KR, Sigler PB** 1991 Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. Nature 352:497-505.

339. **Schwabe JW, Chapman L, Finch JT, Rhodes D** 1993 The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. Cell 75:567-78.

340. **Collingwood TN, Urnov FD, Wolffe AP** 1999 Nuclear receptors: coactivators, corepressors and chromatin remodeling in the control of transcription. J Mol Endocrinol 23:255-75.

341. **Deroo BJ, Archer TK** 2001 Glucocorticoid receptor-mediated chromatin remodeling in vivo. Oncogene 20:3039-46.

342. **Hsiao PW, Deroo BJ, Archer TK** 2002 Chromatin remodeling and tissue-selective responses of nuclear hormone receptors. Biochem Cell Biol 80:343-51.

343. **Dilworth FJ, Chambon P** 2001 Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription. Oncogene 20:3047-54.

344. **Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS** 2002 Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. Mol Cell 9:279-89.

345. **Yang XW, Model P, Heintz N** 1997 Homologous recombination based modification in Escherichia coli and germline transmission in transgenic mice of a bacterial artificial chromosome. Nat Biotechnol 15:859-65.