Characterization of a Repressor Element and Purification of its Cognate DNA-Binding Protein for the Transcription of the Genes for the Antifreeze Proteins in Wolffish (Anarhichas lupus)

by

Stefán Einar Stefánsson

A Thesis submitted in conformity with the requirements for the Degree of Doctor of Philosophy, Department of Biochemistry, in the University of Toronto

Copyright © by Stefán Einar Stefánsson 1997

National Library of Canada

Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services

Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

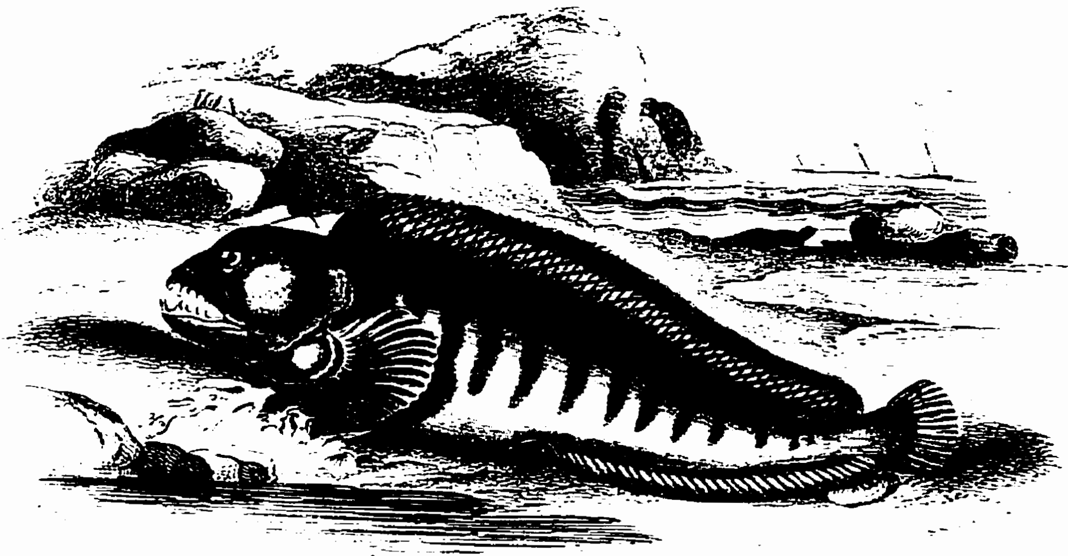
L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-27734-8

Canada

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Dr. C. L. Hew. for his support and guidance over the past several years. Without his help and encouragement, this work would not have been possible. I also thank Dr. Z. Gong for technical help and the gift of a DNA-construct. Furthermore, I would like to thank Dr. S. L. Chan, for the help with nuclear extract and general help in the laboratory, Dr. K. Suzuki for the assistance with the protein purification and Dr. Y. Le Drean for assistance with cell culture work. Ms. L. Mark. I am indebted for many various assistance and help. Other members of Dr. Hew's laboratory. past and present. I owe my gratitude for various assistance. advices and kindness, during my study. Mr. S. Joshi and employees at the Biotech Service Center. past and present, I wish to thank for many various assistance and advices. I would like to thank Dr. D. Pulleyblank. for his help with the design of various oligonucleotides and Dr. D. Templeton for access to his equipment. My parents, I owe many thanks for their support during this time and my brothers, I am indebted for generous encouragements all these years. My co-supervisor. Dr. H. P.Elsholtz, I am thankful for many suggestions. discussions and access of equipment in his laboratory. Finally, I would like to thank Dr. R. A. Anwar. also a co-supervisor. for his generous support, encouragement and many suggestions. and for the time spent in his laboratory, during the start of my studies.



Wolffish (*Anarhichas lupus*). by William Home Lizars. Etched 1843.

Table of Contents

ACKNOWLEDGMENTS ii

TABLE OF CONTENTS iv

ABSTRACT ix

LIST OF FIGURES xi

TABLE xiv

ABBREVIATIONS xv

CHAPTER 1 1

INTRODUCTION 1

1. ANTIFREEZE PROTEINS 2

1.1. AFGP 4

1.2. AFP Type I 4

1.3. AFP Type II 5

1.4. AFP Type III 5

2.0. GENE EXPRESSION AND CONTROL OF AP SYNTHESIS 6

3.0. INITIATION AND REPRESSION OF TRANSCRIPTION 8

3.1. Assembly of the Initiation Complex 8

3.2. Repression of Transcription in Eukaryotes 10

3.3. Inhibition of Translocation of *Trans*-Acting Factors into the Nucleus 13

3.4. Competition and Inhibition of DNA-Binding 14

Competition for DNA-binding 14

Redox controlled DNA-binding 16

Phosphorylation of Transcription Factors 17

3.5. Inhibition of DNA-Binding by Titration 18

3.6. Blocking the Activation of Transcription Factors 21
Retinoblastoma gene product 21
GAL4 21
N-CoR 22
3.7. Inhibition of the Assembly and Activation of the Transcriptional Machinery 23
HMG1 23
Drl 24
Motl 25
HMG2 26
DNA Topoisomerase I 26
IE86 27
Krüppel 28
3.8 Repression of Transcription Mediated by Proteases 29
4. RESEARCH PROPOSAL 30

CHAPTER 2 32
SEQUENCE COMPARISON OF THE GENES FOR THE TYPE III AFP AND THEIR FLANKING REGIONS 32
ABSTRACT 33
INTRODUCTION 34
MATERIALS AND METHODS 37
RESULTS 38
DISCUSSION 47

CHAPTER 3 51
CHARACTERIZATION OF A REPRESSOR ELEMENT FOR THE GENE FOR THE AFP IN WOLFFISH 51

ABSTRACT 52
INTRODUCTION 53
MATERIALS AND METHODS 56
Preparation of Nuclear Extract from Liver from Rainbow Trout 56
Subcloning 57
Radiolabelling of DNA 58
Experiments with Gel Mobility Shift 59
Measurements of Protein Concentration 60
Cell Cultures 61
Transfection of Cells 62
Estimation of Promoter Strength 63
Chloramphenicol Acetyltransferase Assay 63
Measurement of β-Galactosidase Activity 64
Measurement of Luciferase Activity 64
DNA-Footprinting 64
Maxam and Gilbert Reaction 65
Large Scale Preparation of DNA Plasmids and Sequencing of DNA 66
RESULTS 67
Experiments with Mobility Shift Assay 67
Studies with DNA-Transfection 68
Studies with DNA-Footprinting 69
DISCUSSION 78

CHAPTER 4 81
PURIFCATION AND CHARACTERIZATION OF A REPRESSOR PROTEIN
FOR THE TRANSCRIPTION OF TYPE III AFP GENE 81
ABSTRACT 82

INTRODUCTION 83
MATERIALS AND METHODS 86
Preparation of pBLXb Construct 86
Preparation of (Xba I)$_n$ DNA 86
Cross-linking Proteins to DNA 87
Mobility Shift Experiments 88
Construction of (Xba I)$_n$ Affinity Column 89
Oligomer "b" Affinity Column 89
Electrophoresis of Proteins 90
Gel Staining 91
N-terminal Sequencing and Amino Acid Composition Analysis 91
Bio-Gel P-60 Chromatography 91
Concentration of Samples from the Chromatography Columns 92
Chromatography with (Xba I)$_n$ Affinity Column 92
Chromatography with Oligomer "b" Affinity Column 93
Electroblotting and Staining of Membrane 93
Large Scale Preparation of Nuclei 93
RESULTS 95
Cross-Linking the "b" Element to the Putative Repressor 95
Purification of the Repressor Protein 95
Chemical Analysis of Purified Proteins 97
DISCUSSION 106

CHAPTER 5 109
CONCLUSION 109
CONCLUSION AND GENERAL DISCUSSION 110
Characterization of a Repressor Element in the AFP Promoter in Wolffish ... 110

Purification of the Repressor Binding Protein 114

Future Experiments and Prospects 115

REFERENCES.. 117

Characterization of a Repressor Element and Purification of its Cognate DNA-Binding Protein for the Transcription of the Genes for the Antifreeze Proteins in Wolffish

(Anarhichas lupus)

Stefán Einar Stefánsson

Ph.D. Thesis, 1997

Department of Biochemistry, University of Toronto

ABSTRACT

Many marine fish secrete antifreeze proteins (AFP) into their sera to protect them selves from freezing in icy sea water. Four different types of AFPs have been characterized. antifreeze glycoprotein and AFP type I, II and III. The transcriptional control for the Type III Antifreeze protein (AFP) genes is poorly understood. In this present study, all known Type III genes and flanking DNA sequences were compared, revealing that the Type III AFP genes from wolffish (*Anarhichas lupus*), ocean pout (*Macrozoarces americanus*) and Antarctic eel pout (*Lycodichthys dearborni*) have common core sequences. These core segments are intervened by unique regions, which are present in some, but not in all of the genes aligned. A region from wolffish AFP genes, located between -220 to -90 bp upstream from the transcription start site, was previously characterized to repress transcription. This segment falls into two different core regions, with a unique segment between. Two out of the three characterized wolffish sequences, have two 12 base pair repeated elements in tandem, called "b" elements within this described repressor region. The first repeated "b" element, is within a core region, and therefore is present in all Type III AFP flanking sequences. The second repeated "b" element is unique to the two wolffish sequences. The -220 to -90 bp fragment is shown to interact with proteins from nuclear extract of various sources with strongest interaction in liver tissues. Several proteins appear with the -220 to -90 bp DNA segment. The "b" elements also interact with nuclear proteins from the liver. Furthermore.

the repressor region is shown here to be due to the two "b" elements, which mediate a strong repression of transcription, presumably by interacting with the general transcription factors. Several additional DNA-binding sites are also identified within this segment.

To further distinguish the nature of the repressor element, its putative binding protein(s) was purified and characterized. Three proteins of different size appear to interact with the "b" element in liver tissue. These proteins, termed pBB for protein binding to "b" element, are between 34 and 38.5 kDa in size. The size of the purified proteins is in good agreement with cross-linking experiment performed on the crude nuclear extract. All three proteins appear to be rich in glutamine and glutamic acid residues, which are commonly found in repression and activation domains of *trans*-acting factors. Finally, all three proteins have identical N-terminal sequences, raising the possibility that the difference in molecular weight might be due to micro-heterogeneity, alternative splicing of the mRN. Both the elements and the DNA-binding proteins described here appear to be novel, and have not been previously characterized.

List of Figures

Wolffish (*Anarhichas lupus*)........ iii

Figure 1-1. Antifreeze activity is highest during the winter months in fish off the coast of Newfoundland 7

Figure 1-2. Several factors are needed for efficient transcription. 11

Figure 1-3. Titration prevents DNA-binding 19

Figure 2-1. Sequence alignment of the genes and flanking region for Type III AFP from ocean pout, wolffish and Antarctic eel pout........ 41

Figure 2-2. Schematic representation of sequence alignment of the AFP genes and flanking sequences........ 45

Figure 2-3. Schematic figure of core segments in the flanking region and the gene for the Type III AFP........ 46

Figure 3-1. Binding of nuclear proteins with the -220 to -90 DNA-fragment from the wolffish promoter. 70

Figure 3-2. Multiple bands appear in a mobility shift assay with a fragment from wolffish promoter........ 71

Figure 3-3. Oligomer with two "b" elements is specific and does not bind to putative "a" binding proteins 72

Figure 3-4. DNA constructs for transfections 73

Figure 3-5. A repressor element is located to the two "b" elements 74

Figure 3-6. The "b" element has strong repressor activity in fish cell lines 75

Figure 3-7. The "b" element is protected from DNase I digestion 76

Figure 3-8. Sequence protected from DNase I digestion 77

Figure 4-1. The "b" element cross-links to a protein of around 36 kDa in size 98

Figure 4-2. Schematic outline for purification of the putative repressor protein for the "b" element 99

Figure 4-3. Protein components of around 60 kDa or larger binds to the upstream region of the AFP promoter in wolffish 100

Figure 4-4. Chromatography with the $(Xba\ I)_n$ DNA-affinity column 101

Figure 4-5. Positive fractions from the $(Xba\ I)_n$ column interact with the "a, b and S" elements 102

Figure 4-6. Positive fraction from the "b" affinity column interacts with the"b" element in mobility shift assay 103

Figure 4-7. Positive fraction from the "b" affinity column shows three distinct bands 104

Figure 4-8. Amino acid composition of the three proteins isolated with the "b" affinity column 105

Figure 5-1. Model for the action of the pBB on the transcription complex 112

Table

Table 1. Comparison of fish AFPs 3

ABBREVIATIONS

AEP	Antarctic eel pout
AFP	Antifreeze protein
bp	Base pair
BIS	N,N'-Methylene-bis-acrylamide
BSA	Bovine serum albumin
°C	Degrees on Celsius
cAMP	Cyclic adenosine monophosphate
CAT	Chloramphenicol acetyl transferase
C/EBP	CAAT/ enhancer binding protein
CHH	Chum salmon heart cell line
CHOP	C/EBP homologous protein
CoA	Coenzyme A
CRE	Cyclic AMP responsive element
CTD	Carboxy terminal domain
Da	Daltons
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanine triphosphate
dNTP	Deoxynucleotide triphosphate
dTTP	Deoxythymidine triphosphate
DNA	Deoxyribonucleotide acid
dsDNA	Double stranded DNA
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
eve	Even-skipped

FBS	Fetal bovine serum
G/C	Deoxyguanine and deoxycytidine
GTF	General transcription factors
HBS	Hepes buffer saline
HEPES	(N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid])
HMG	High mobility group
IE	Immediately early
mRNA	Messenger RNA
nM	Nanometer
OP	Ocean pout
OPA	One for all
PBS	Phosphate buffer saline
PMSF	Phenylmethylsulfonyl fluoride
PVDF	Polyvinylidene difluoride
pWON	Plasmid with wolffish negative elements
redox	Oxidation and reduction
Rb	Retinoblastoma
RAR	Retinoic acid receptor
RXR	Retinoic X receptor
RNA	Ribonucleic acid
RNA Pol	RNA polymerase
RTH	Rainbow trout hepatoma cell line
SDS	Sodium dodecyl sulfate
TAF	TATA binding associated factors
TBE	Tris borate EDTA
TBP	TATA binding protein
TEMED	NNN'N' Tetramethylethylenediamine

TF	Transcription factor
TK	Thymidine kinase
Tris	Tris[hydroxymethyl]-aminomethane hydrochloride
UV	Ultraviolet
WO	Wolffish

Chapter 1

INTRODUCTION

1. ANTIFREEZE PROTEINS

For survival in ice laden seas or polar waters, several species of fish synthesize proteins that prevent ice crystal growth in their plasma. The need for this antifreeze protein in marine teleosts arises because their plasma composition is hypotonic in comparison with sea water and will therefore freeze at higher temperatures than sea water. The plasma for most fish will freeze around -0.7 °C, whereas sea water freezes around -1.9 °C (Holmes & Donaldson, 1969). Fish inhabiting these cold waters can therefore encounter temperatures that are more than one degree below the freezing point of their plasma. In an expedition to northern Canada, Scholander and coworkers observed that the plasma of inshore Arctic fish had an unusually low freezing point of around -1.4 °C (Scholander *et al.*, 1957). In 1969 DeVries and Wohlschlag reported that the antifreeze agent in fish from the Antarctic shelf was a proteinaceous molecule (DeVries & Wohlschlag, 1969).

Subsequently, freezing resistance has been observed in many other organisms. In woody plants during the over-wintering period, the cold-hardiness is attributed to structural features of the tissue (George & Burke, 1984). In winter rye (*Secale cereal L.*), unique antifreeze proteins have been isolated from the apoplast (Hon *et al.*, 1994), and plants collected during autumn and winter from northern Indiana have antifreeze activity (Urrutia *et al.*, 1992). Many insects remain frost-tolerant in a super cooled state, to temperatures as low as -61.6 °C. These insects contain a high concentration of polyhydric alcohols such as glycerol which combined with ice nucleating agents may allow them to reach these low temperatures (Sømme, 1982). Antifreeze proteins have been detected in insects as well. These proteins range in size from 14 to 20 kDa, lack carbohydrate, have a high percentage of hydrophilic amino acids and some contain

	AFGP	AFP Type I	AFP Type II	AFP Type III
Carbohydrate	Yes	No	No	No
Mass	2,600 to 33,000	3,300 to 4,500	11,000 to 24,000	6500 to 12000
Primary structure	(Ala Ala Thr)n disaccaride	alanine-rich multiple of eleven aa repeats	cystine-rich disulfide linked	general
Secondary structure	expanded	α helical amphiphilic	ß strand	ß strand
Tertiary structure	N.D.	100% helix	N.D.	ß sheet
Biosynthesis	multiprotein	prepro AFP	prepro AFP	pre AFP
Protein components	8	7	2-6	12
Gene copies	N.D.	40-50	15	43-150
Fish species	Antarctic nototheniids (cods)	right eyed flounders (winter flounder) Sculpins (shorthorn)	sea raven smelt, herring,	eel pouts (ocean pout, wolffish, Antarctic eel pout)

Table 1. Comparison of fish AFPs. N.D. not done, revised from Chan, 1995.

cysteine residues (Duman & Olsen, 1993). Certain species of terrestrial frogs can survive frozen at -6 °C for several days, due to a high glycerol content (Schmid, 1982).

In fish, four unique types of antifreeze have been identified (Table 1). These are antifreeze glycoprotein (AFGP), and AFP Type I, II and III, each with a distinct composition and structure.

1.1. AFGP

The first antifreeze protein to be isolated was from a nototheniid, caught off the coast of Antarctica (DeVries *et al.*, 1970). AFGPs have also been identified from cod fishes which are an unrelated species (Van Voorhies *et al.*, 1978). These proteins are made up of repeats of Ala-Ala-Thr residues, and have a disaccharide attached to the hydroxyl oxygen of the threonine residue. Loss of the disaccharide unit reduces the antifreeze activity (Komatsu *et al.*, 1970). These proteins are found in a relatively high concentration in the fish sera (10-25 mg/ml) and range in size from 2,600 to 33,000 Daltons (Davies & Hew, 1990). The proteins are synthesized as multiproteins, which are subsequently cleaved to form a mature protein (Hsiao *et al.*, 1990).

1.2. AFP Type I

The Type I AFP is the most extensively studied of all the AFPs. These proteins are found among right eyed flounders and sculpins (Duman & DeVries, 1974; Hew *et al.*, 1980). The proteins range in size from 3,000 to 5,000 Daltons and are rich in alanine residues (>60% mol Ala); they form α-helices with threonine residues protruding on one side of the helix, and it is the threonines that are postulated to bind to the ice surface(Yang *et al.*, 1988; Sicheri & Yang, 1995).

The protein is transcribed from around 40-50 genes per genome. It is synthesized as a preproprotein, with a typical 23 amino acid residue signal peptide, which is cleaved

during processing of the protein. The proprotein is further processed by the removal of 21 residues from the amino end (Hew *et al.*, 1986). The two major forms of winter flounder (*Pleuronectus americanus*) contain three 11-amino acid tandem repeats of the sequence $ThrX_2AsxX_7$, where X is usually alanine or other residues that favor α-helix formation. The effect of the repeat is to generate an α-helix, amphiphilic in nature, which is stabilized with internal and terminal ion dipolar interactions.

1.3. AFP Type II

Of the various AFPs, Type II has the widest distribution in fish. It is found in unrelated species such as herring (*Clupea harangeus harangeus*), smelt (*Osmerus mordax*) and sea raven (*Hemitipterus americanus*) (Slaughter *et al.*, 1981; Ewart *et al.*, 1992). These proteins range in size from 11,300 to 24,000 Daltons and are transcribed from around 15 genes per genome (Hayes *et al.*, 1989). The uniqueness arises from the fact that they have a high content of cysteine residues, which form disulfide bridges. The Type II proteins have predominantly a β-sheet structure (Sonnichsen *et al.*, 1995) and are synthesized as preproproteins (Hayes *et al.*, 1989).

The Type II AFPs have homology to the carbohydrate-recognition domain of Ca^{2+}-dependent (C-type) lectins (Ewart & Fletcher, 1993). Furthermore, the smelt and herring Type II AFP are unique among known fish AFPs, in their requirement of Ca^{2+} as a cofactor for antifreeze activity (Ewart *et al.*, 1996).

1.4. AFP Type III

The Type III AFP has the most limited distribution among various groups of AFPs, found only in fishes of the suborder of Zoarcoide, which live in the Antarctic and Arctic regions. These proteins have neither a bias towards any amino acid residue, nor do they contain any cysteine residue. These proteins are around 6,000 Daltons in size

and are transcribed from around 150 genes in the ocean pout (*Macrozoares americanus*) (Hew *et al.*, 1988). The genes are linked and arranged in tandem repeats. The structure of the protein has been solved with nuclear magnetic resonance spectroscopy and crystallography, and was shown to be a compact globular structure with eight β strands arranged into three β-sheets (Sonnichsen *et al.*, 1993; Jia *et al.*, 1996). The protein has a flat ice-binding surface, made of five residues, which bind to the ice crystal, thereby preventing its growth (Jia *et al.*, 1996).

2.0. GENE EXPRESSION AND CONTROL OF AFP SYNTHESIS

The expression of the AFPs in fish off the coast of Newfoundland is under seasonal control (Figure 1-1). Concomitant with the coldest temperature of the sea, these fish have the highest concentration of AFPs circulating in their plasma. The genes for different groups of AFPs, have no common *cis*-acting elements, which would have suggested a common transcriptional control. Furthermore, the tissue distribution of the various AFPs differs. Expression of the genes for Type II is restricted to the liver, whereas the genes for Type I and III are ubiquitously expressed, but mostly in the liver.

Interestingly, in winter flounder the rate of transcription for the AFP genes have not affected by ambient sea water temperature (Vaisius *et al.*, 1989). The temperature instead affects the clearance rate of the AFP from the plasma (Fletcher, 1981). Exposure to 10-11 °C during the winter months results in low levels of AFP in the plasma, presumably due to accelerated clearance rate of the protein from the plasma (Fletcher, 1981). Degradation of the mRNA for AFP is also affected by the sea water temperature. Flounders kept at 18 °C during the fall, have low levels of mRNA for AFP, presumably due to higher degradation rate (Price *et al.*, 1986).

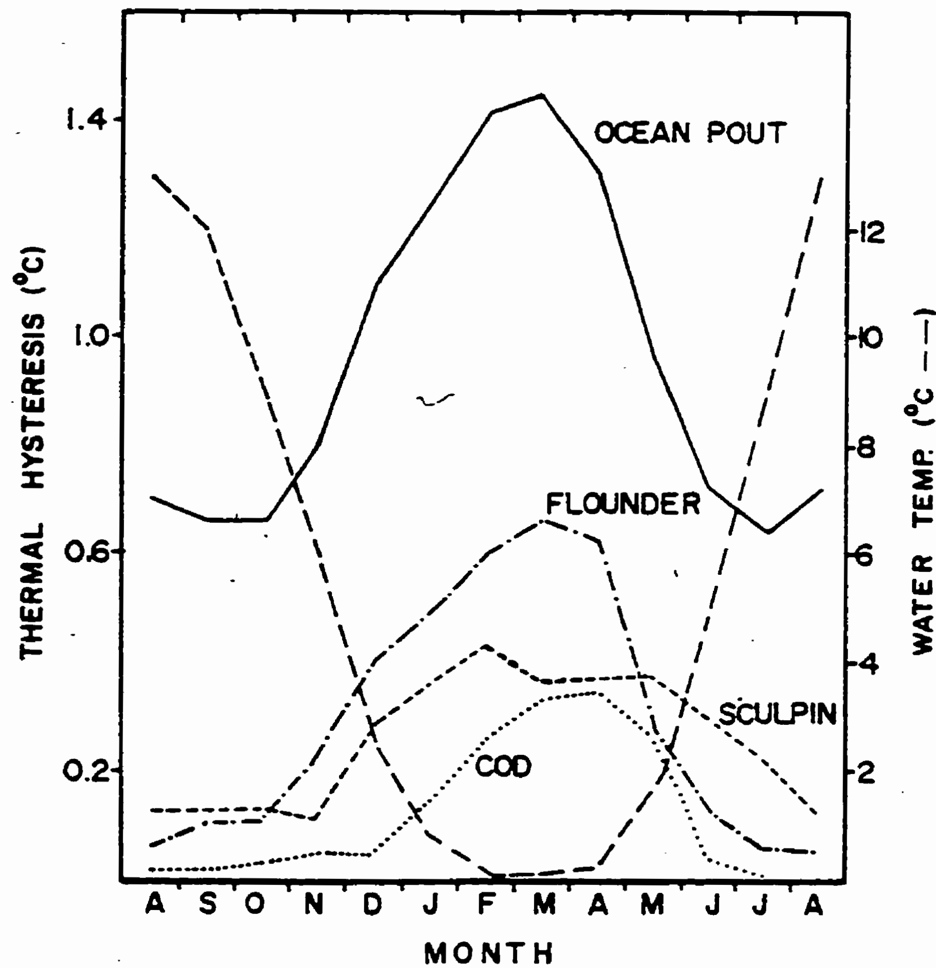


Figure 1-1. Antifreeze activity is highest during the coldest months in the sea water in fish off the coast of Newfoundland (Davies *et al*., 1988).

The expression of AFP from flounder is under the control of photoperiod. Long day length (≥ 14 hours) suppresses synthesis of AFP. In contrast, short day length did not induce synthesis of AFP (Fletcher, 1981; Fourney *et al.*, 1984b).

Implants of pituitary tissue or injections of pituitary extract from tissues collected in the spring, resulted in inhibition of AFP synthesis (Fletcher, 1979; Fletcher *et al.*, 1984). Subsequently it was discovered that growth hormone from the pituitary gland mediates inhibition of AFP synthesis. During the summer season, the presence of growth hormone represses synthesis of AFP, whereas in the winter season, when growth hormone is absent, AFP accumulates in the plasma (Idler *et al.*, 1989).

The AFGP from Atlantic cod (*Gadus morhua*) is not affected by photoperiod, but synthesis seems to be accelerated by lower temperatures (Fletcher *et al.*, 1987).

Promoter analysis of the genes for AFP Type I, II and III have revealed regions containing both positive and negative elements for transcription but when tested in various cell lines or in medaka embryos (*Oryzaias latipes*) no tissue specific regions were identified (Gong *et al.*, 1991; Gong & Hew, 1993).

3.0. INITIATION AND REPRESSION OF TRANSCRIPTION

3.1. Assembly of the Initiation Complex

Eukaryotic cells contain three different RNA polymerases (RNA pol), each interacting with a different set of genes; RNA pol I transcribes only genes for ribosomal RNAs; RNA pol II transcribes genes coding for proteins and snRNAs (small nuclear RNA); RNA pol III transcribes snRNAs, tRNAs, 5S RNA. The scope of this introduction will be limited to a discussion of initiation and control of transcription by RNA pol II.

The RNA pol II consists of approximately 10 subunits. The two largest proteins in yeast. *Saccharomyces cerivisiae*, are 220 and 140 kDa in size and are related to the β and β' subunits of the RNA polymerase from *Escherichia coli,* respectively. The largest fragment contains a repeated heptad peptide, in the carboxy terminal domain (CTD) (Young, 1991). The enzyme binds only to the initiation complex with the non-phosphorylated form of CTD. Once transcription has started the heptad repeat is heavily phosphorylated (Usheva *et al.*, 1992). The RNA pol II can not initiate transcription on its own, but needs in addition other factors termed general transcription factors (GTF) (Segall *et al.*, 1986).

The initiation complex assembles around two distinct promoter elements. that work either independently or collectively, the TATA box and the initiator element. For most eukaryotic genes coding for proteins, the first step in gene expression is a template commitment in which of TFIID (transcription factor IID) binds to the minor groove of the TATA box: which is usually located around 30 base pairs upstream of the start site for transcription. The TFIID is a heterogeneous multisubunit complex with a core component. TATA binding protein (TBP), that binds to the DNA element. TBP is a relatively small protein. with a molecular size around 27 kDa in humans. It has a conserved carboxyl-terminus with repeated segments and a variable N-terminal end (Horikoshi *et al.*, 1989; Hoeijmakers, 1990; Kao *et al.*, 1990: Yamamoto *et al.*. 1992). Co-crystals of TBP with DNA show that the protein binds as a saddle over the minor groove by bending the DNA (Kim *et al.*, 1993a; Kim *et al.*. 1993b). Attached to the TBP, are TBP associated factors (TAF). The exact role of the TAFs is unknown. TBP supports basal transcription. but for activated transcription, addition of TAFs were thought to be needed. even though TBP has been shown to interact directly with activating factors (Hoey *et al.*, 1990; Reese *et al.*, 1994). The TBP and TAFs mediate interaction with other DNA-bound *trans*-acting factors. the RNA Pol II and GTFs

(Usheva *et al.*, 1992; Chiang & Roeder, 1995; Dikstein *et al.*, 1996). Recently. transcription in yeast was shown not to be affected by mutating most of the TAFs. even though the mutation was lethal. This suggested that the TAFs might be involved in cell cycle control (Moqtaderi *et al.*, 1996; Walker *et al.*, 1996).The largest factor associated with TBP, $TAF_{II}250$ nucleates the assembly of other TAFs on TBP (Weinzierl *et al.*. 1993).

The next component to be integrated into the complex is debated. TFIIA stabilizes the binding of TFIID, but can be omitted in many systems. The role of TFIIA might be to counteract repressors associated with the initiation complex. TFIIA is coded by two genes in yeast, and deletion of either one renders the cell non-viable (Ranish *et al.*. 1992). The other possible factor to enter after TFIID, is TFIIB, with subsequent incorporation of the RNA pol II, followed by TFIIF, TFIIE, TFIIH and TFIIJ. The absolute requirement of GTFs for transcription depends on promoter content and DNA topology (Parvin & Sharp, 1993). Recently, the RNA pol II was found to exist as a holocomplex, with all the aforementioned factors, excluding TFIIA. The TBP was in sub-molar concentrations in this complex and had to be supplemented for maximal activity (Figure 1-2) (Koleske & Young, 1994; Ossipow *et al.*, 1995).

Repression of transcription can in principle affect any stage of the assembly of the initiation complex, processing of the reaction or the activity of the activators. This spans from blocking the nuclear localization signal (Wulczyn *et al.*, 1992) to halting the assembly of the initiation complex (Auble *et al.*, 1994).

3.2. Repression of Transcription in Eukaryotes

Control of transcription is mediated by sequence-specific DNA-binding factors. which interact either directly or indirectly with the general transcription machinery.

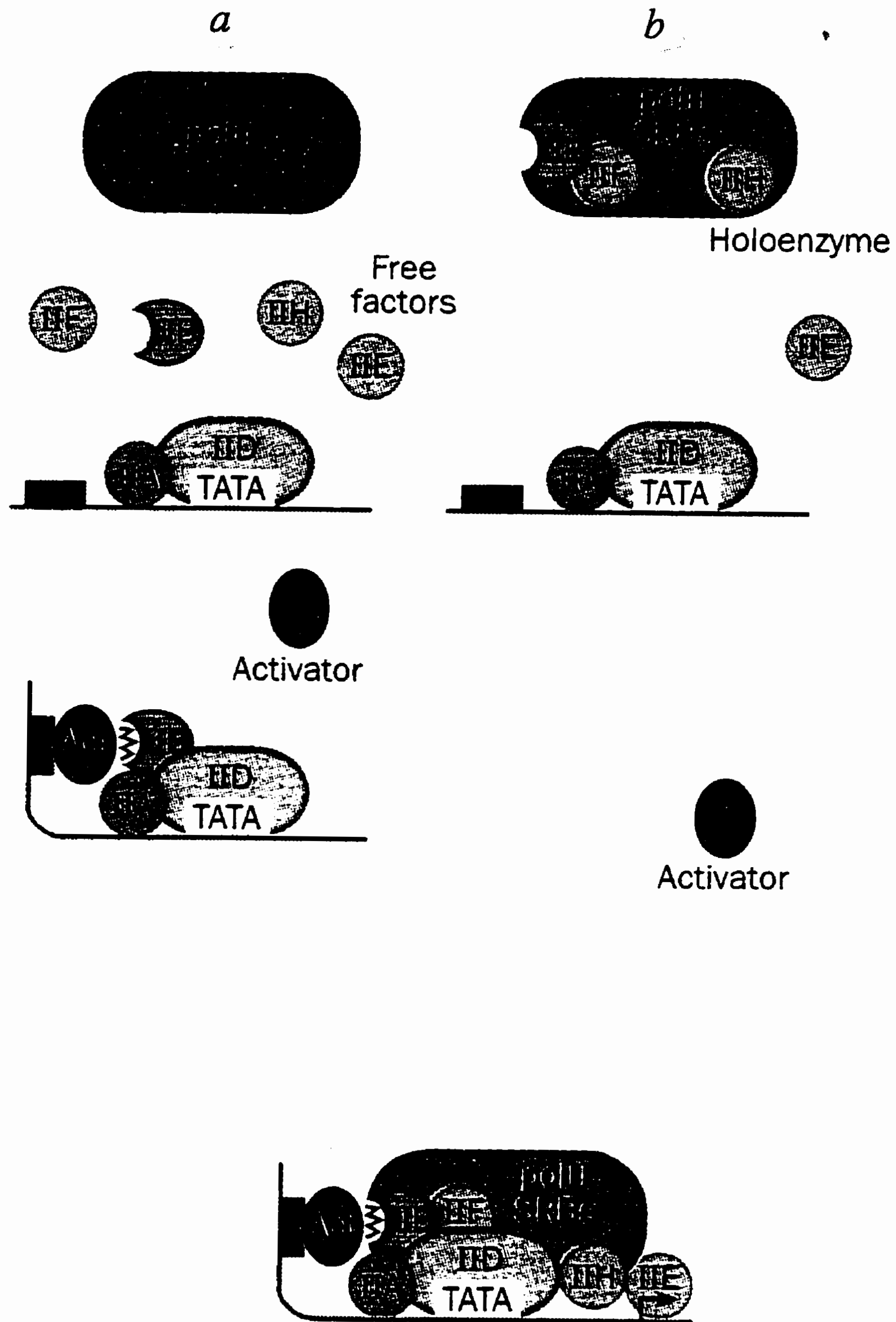


Figure 1-2. Several factors are needed for efficient transcription. Two theories are used to explain the assembly of the initiation complex. A) Ordered stepwise assembly is needed for initiation or B) RNA pol II binds to the DNA as a holoenzyme (Carey, 1994).

Presently, a great number of *trans*-acting factors have been purified and characterized which have diverse effects on transcription (Faisst & Meyer, 1992). Along with characterizing their cognate DNA-elements, protein-protein interaction and *trans*-action domains have been identified. This has led to the characterization of several families of *trans*-acting factors (Harrison, 1991). Several common motifs of activation domains have been identified, based on the prevalence of amino acids residue content. These are 1) proline rich domains, found in factors such as CTF/NF-I and AP2, 2) acidic regions in GAL4 and GCN4 3) glutamine enriched domain in Sp1 and OCT-1 and OCT-2 and finally 4) isoleucine rich domains in NTF-1 in *Drosophilia*. Other activating domains have been found, but no clear structural or amino acid residues have emerged (Mitchell & Tjian, 1989; Harrison, 1991; Attardi & Tjian, 1993; Kunzler *et al.*, 1994)

Research of gene control has largely focused on activating factors with less attention diverted towards repressors; the latter were sometimes were considered not as relevant compared with activators as was expressed in a popular textbook Molecular Biology of the Cell (Alberts *et al.*, 1983): "Common sense suggests that in higher cells.... most gene specific regulatory proteins act as gene activators, serving to turn on particular genes for transcription". And in the same chapter "In a typical eukaryotic cell, only about 7% of the DNA sequence are ever transcribed into RNA. It seems very unlikely that transcription is specifically blocked on the remaining 93% of the DNA by tens of thousands of different repressor proteins". Contrary to these earlier views, transcriptional repressors have now been shown to play a major role in gene control and life cycle of organisms. The yeast mating system is centered around silencers (Dillin & Rine, 1995): the tumor suppressors of Wilms tumor gene product (WT1), retinoblastoma susceptibility gene product (Rb) and p53 have all been implicated in transcriptional repressing activity (Fattaey *et al.*, 1993; Liu *et al.*, 1993; Wang *et al.*, 1993). The life

cycle of the bacteriophage λ is under the control of the its own repressor, cI, directing the phage into a lysogenic or lytic pathway (Ptashne, 1986).

The DNA is effectively prevented from transcription by chromosome condensation. The maintenance of the silent state involves methylation of the cytosine residue, primarily at CpG sites. Access to the DNA for transcription is also affected by chromatin structure, were compact structure prevents the assembly of transcription complexes (Herschbach & Johnson, 1993). The accessibility of the DNA for transcription, due to either methylation or chromatin structure will not be covered in this review.

Several different mechanisms of repression have been identified (Levine & Manley, 1989; Renkawitz, 1990; Cowell, 1994; Koleske & Young, 1994). The examples of repression can be divided into several different classes based on their mode of action. These are competition or inhibition of translocation of transacting factors to the nucleus, steric hindrance for DNA-binding, blocking of activation, titration, redox regulation, protease digestion and direct repression which affects the assembly of the general transcription machinery.

Inhibition of Translocation of Trans-Acting Factors into the Nucleus

One of the earliest steps which a repressor could inhibit activation, is by preventing the translocation of the activators to the nucleus. The IκB (Inhibitor of NFκB) exemplifies this, by retaining Rel activators in the cytoplasm. Members of the Rel family respond to signal transduction stimuli, and are involved in immune responses in mammals and dorso-ventral polarity during the embryonic development in *Drosophila* (Geisler *et al.*, 1992; Drew *et al.*, 1995). The proteins in the Rel family contain a 300 amino acid conserved region in their amino terminus, which harbors the nuclear localization signal, phosphorylation site for protein kinase A, dimerisation domain, and a

region responsible for interaction with IκB (Verma *et al.*, 1995). The counteracting domain in IκB has several so-called ankyrin-like repeats. This motif, commonly found in a number of different proteins, is responsible for the retention of the Rel proteins in the cytoplasm. The NFκB (Nuclear Factor for the Ig κ gene in B cells) is a complex composed of two proteins; p65 and p50. Only p65 interacts with the IκB (Matthews & Hay, 1995). Upon stimulation, the IκB is phosphorylated and releases the NFκB , which then translocates into the nucleus and activates transcription. The IκB is believed to mask the nuclear localization signal on the NFκB through its ankyrin repeat, thereby retaining the complex in the cytoplasm (Hoffman, 1991).

Interestingly, the p50 protein is synthesized as a p105 preprotein. The p50 is proteolytically cleaved from the p105 protein upon activation. The carboxy terminus of p105 contains seven ankyrin-like repeats which are presumed to overlap the nuclear localization signal and prevent migration to the nucleus (Blank *et al.*, 1992; Beg & Baldwin, 1993; Grimm & Baeuerle, 1993; Wasserman, 1993).

3.4. Competition and Inhibition of DNA-Binding

Competition for DNA binding

Competition for a DNA-binding site is common in bacteria (Renkawitz, 1993). One of the first examples of repression of transcription in eukaryotes comes from the large T antigen from simian virus 40 (SV40) which binds to its own promoter and thereby prevents the binding or initiation of the transcription machinery (Fanning & Knippers, 1992). This mode of repression is difficult to achieve in most eukaryotic genes, as they are under the control of many transcriptional elements, and therefore many repressor proteins would be needed for efficient silencing (Johnson, 1995). As described below, several examples are known where the repressor prevents the binding of the TFIID complex to the TATA box.

Engrailed, a homeodomain protein, which is involved in the development of the *Drosophila* embryo, competes with TFIID for binding to the TATA box. Once the engrailed protein binds the TATA box, the assembly of the pre-initiation complex is thwarted. In contrast, if the engrailed protein is preincubated with TFIID, the GTFs are able to assemble and transcription can proceed (Ohkuma *et al.*, 1990).

Even-skipped (eve), another homeodomain protein in *Drosophila*, participates in segmentation during embryonic development. In the promoter for the Ultrabithorax gene, eve binds to an overlapping site for the activator, zeste, and prevents its DNA-mediated activation (TenHarmsel *et al.*, 1993). The same promoter has strong affinity binding sites for eve, located between +50 and +100, downstream of transcriptional start site. Low affinity binding sites are located between +18 to +43, and -140 to -170. Other possible low affinity binding sites are located around the TATA box. The eve protein is suggested to have a cooperative blocking mechanism on transcription, where the high affinity binding sites, enhance eve binding to the lower affinity site, resulting in DNA loop formation and thereby preventing TFIID from making template commitment (Austin & Biggin, 1995). Interestingly, eve also represses transcription by preventing the assembly of the preinitiation complex on the basal promoter for the alcohol dehydrogenase promoter (Johnson & Krasnow, 1992).

The human immunodeficiency virus-1, bovine papilloma virus promoter P_{105} and the P-element in *Drosophila*, all contain a TATA box and an alternative overlapping binding site for TMF, E2 and transposase, respectively. These proteins repress transcription by preventing the TFIID from DNA-binding and template commitment (Dostatni *et al.*, 1991; Kaufman & Rio, 1991; Garcia *et al.*, 1992).

Repression of transcription can result from competition between a repressor and an activator for a DNA-binding site. A GC-rich sequence, found in many promoters, binds to a ubiquitous factor, Sp1 (Kadonaga *et al.*, 1987). The same DNA-binding

element also interacts with a repressor Sp3, competing for DNA-binding with Sp1 factor (Hagen *et al.*, 1994).

Several examples are known where two activators compete for overlapping elements, causing steric hindrance and resulting in reduced gene expression. This is common in the steroid receptor super family. An enhancer for the glycoprotein hormone α–subunit gene is activated by cAMP through cAMP responsive element (CRE). With the addition of glucocorticoids, this activation is strongly reduced. The overlapping of glucocorticoid element with the CRE is believed to prevent the activation of cAMP (Akerblom *et al.*, 1988). The transcription factor Ap1, a heterodimer prototypal from c-jun and Fos represses the expression of human osteocalsin gene, by competing with retinoic acid receptor for an overlapping response element (Schule, 1990).

Redox controlled DNA-binding

Oxidation-reduction (redox) reactions involve the transfer of hydrogen or electrons. The redox potential within the cell is controlled by several different chemicals, such as: thioredoxin (Buchanan *et al.*, 1994), glutathione (Dröge *et al.*, 1994) or redox factor 1 (Ref-1) (Xanthoudakis & Curran, 1992). Sulfur atoms in cysteine residues can be oxidized to yield sulfur bridges or possibly to sulfeinic (RSOH) or sulfinic acid (RSO_2H), depending on nearby amino acid composition (Abate *et al.*, 1990). Spontaneous oxidation of the cysteine residues is enhanced by surrounding basic residues (Ng *et al.*, 1993). Selenium can function as an antioxidant, and rats fed on selenium depleted diets showed altered DNA-binding activity of many *trans*-acting factors (Christensen & Pusey, 1994). The c-Fos–c-Jun heterodimer has a single cysteine residue in each DNA-binding domain, which mediates DNA interaction (Glover & Harrison, 1995). Depending on the redox state, this residue dictates DNA-binding (Abate *et al.*, 1990). In oxidized form, DNA-binding is prevented, but through a cellular

factor, Ref-1, it is reduced and DNA-binding takes place (Abate *et al.*, 1990; Xanthoudakis & Curran, 1992). Ref-1 also stimulates DNA-binding of Jun homodimers, NF-κB, Myb and members of ATF/CREB family (Guehmann *et al.*, 1992; Xanthoudakis *et al.*, 1992). The reductant, thioreducin, also activates NF-κB and augments the activity of human immunodeficiency virus promoter (Matthews *et al.*, 1992; Hayashi *et al.*, 1993). Unlike Fos-Jun, the oxidized state of NF-κB involves disulfide bonds (Matthews *et al.*, 1992).

DNA-binding of several other factors, including Ku protein (Zhang & Yaneva, 1993), p53 (Hainaut & Milner, 1993), USF (Pognonec *et al.*, 1992) and thyroid transcription factor I (Arnone *et al.*, 1995) are also inhibited by oxidation. In contrast, DNA-binding of the human homeobox protein HoxB5, is not affected by the redox state, but instead, demonstrates cooperative binding, but only in its oxidative form (Galang & Hauser, 1993).

Phosphorylation of Transcription Factors

Phosphorylation of transcription factors is often a initial step in their activation (Hunter & Karin, 1992; Hill & Treisman, 1995). Theoretically, dephosphorylation of these factors would then act as a repressor. A few examples have emerged where the phosphorylation plays a direct role in repressing the transcription.

The c-Jun protein has five major phosphorylation sites, with three sites located adjacent to the DNA-binding domain (Binétruy *et al.*, 1991). Phosphorylation of these three sites, prevents DNA-binding. These sites are phosphorylated in non-stimulated cells, but are dephosphorylated by protein kinase C stimuli, with subsequent DNA-interaction (Boyle *et al.*, 1991). One or more of these three sites can be phosphorylated by several different protein kinases, including glycogen synthase kinase 3, shaggy, and casein kinase II (Li *et al.*, 1861; de Groot *et al.*, 1992; Lin *et al.*, 1992).

The Pit-1 protein is expressed in the pituitary gland and activates expression of genes for growth hormone and prolactin during ontogeny. This protein contains the POU-domain, which is involved in DNA-binding and dimerisation (Rosenfeld, 1991). Pit-1 is phosphorylated in the POU -domain, with the result that its conformation on the DNA element is altered. Furthermore, the same phosphorylation can increase or decrease the activation of transcription, depending on the sequences flanking the core element (Kapiloff *et al.*, 1991).

C-Myb, a nuclear oncoprotein, is phosphorylated at its amino-terminus, near its DNA-binding domain. This phosphorylation, by casein kinase II, inhibits sequence specific interaction of c-Myb. Interestingly, this phosphorylation site is deleted in nearly all oncogenically activated Myb proteins (Lüscher *et al.*, 1990).

Phosphorylation of Max, a b/HLH/LZ (basic/helix-loop-helix/leucine zipper) activator, prevents its DNA-binding as homodimer, but does not affect the binding of Myc/Max heterodimer. Both these dimers bind the same element, with opposite effects. Max homodimers repress transcription, whereas the heterodimers are activators. The phosphorylation will therefore have net results of activation (Berberich & Cole. 1992).

3.5. Inhibition of DNA-Binding by Titration

Several examples are known where a repressor titrates out the activity of an activator (Figure 1-3).

The helix-loop-helix (HLH) proteins compromise a family of *trans*-acting factors found in species ranging from plants to mammals. They all have in common a basic region, which binds the DNA element, and two helices intervened by a loop, which mediates dimerisation. Factors in this family are involved in diverse biological functions such as initiation of tumorigenesis (myc) or muscle development (MyoD family). MyoD

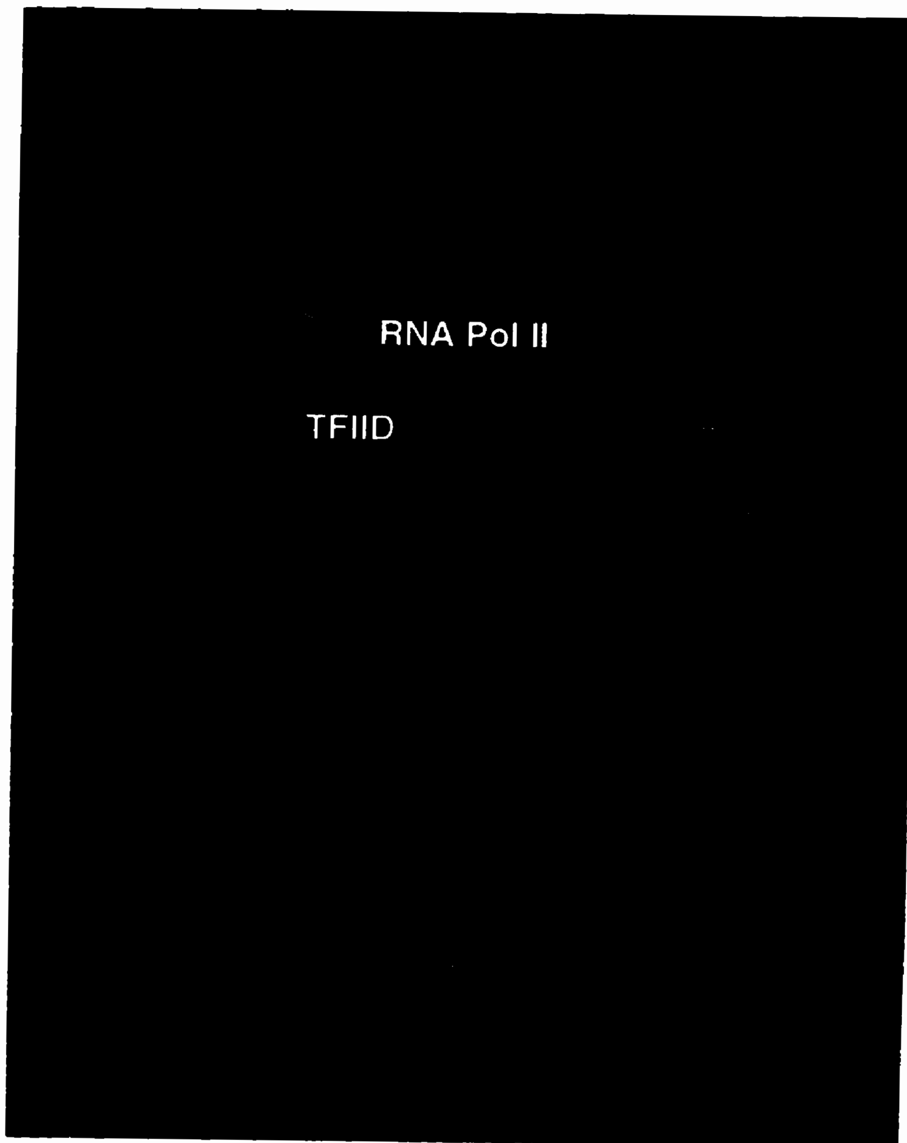


Figure 1-3. Titration prevents DNA-binding. Several factors are repressed by titration, when a repressor protein binds to the activator and prevents DNA-binding. The RNA pol II is labeled as red, TFIID as blue, transcriptional activator as red and repressor as magenta.

forms heterodimers with ubiquitous proteins of the E-family, which also are bHLH proteins (Barinaza, 1991; Visvader & Begley, 1991). The Id protein (Inhibitor of DNA binding), has helix-loop-helix sequence but lacks the basic region that is essential for DNA binding. It is able to form a heterodimer with MyoD, E12 and E47, thereby dissociating the active heterodimer. By overexpressing the Id protein, the activation of muscle creatine kinase is down regulated, due to heterodimer formation with MyoD. Furthermore, during differentiation of myoblast cells, the Id protein is down regulated.
This is believed to allow proteins of the MyoD family to activate muscle specific genes and the differentiation of the cell (Benezra *et al.*, 1990).
The family of steroid receptors have two binding domains for zinc atoms, each coordinated by four cysteine residues, termed zinc fingers (Evans, 1988). The sequence intervening the zinc fingers has high homology within the steroid receptors. The protein calreticulin, a multifunctional protein that acts as a major Ca^{2+} binding and storage protein in the cell, binds to the intervening domain of the zinc fingers. This interaction prevents the steroid receptors from binding to the DNA, and thereby blocks activation of transcription (Burns *et al.*, 1994; Dedhar *et al.*, 1994).

Members of the C/EBP (CAAT/ Enhancer Binding Protein) family, share similar sequences in their b/ZIP region (basic-leucine zipper), which also mediates DNA-binding and dimerisation. The ubiquitously expressed protein C/EBPγ lacks the transcriptional activation domain. By forming heterodimers with C/EBPα, it prevents the latter from activating transcription (Cooper *et al.*, 1995).

Another member of the C/EBP family, CHOP-10, (C/EBP-homologous protein), has two proline residues in the basic region, which normally mediates DNA-binding. Through the leucine zipper region, it forms heterodimers with C/EBP and other members of the family. The heterodimers fail to bind to the DNA, due to the proline

residues in the basic region of the CHOP-10, thereby preventing activation of transcription (Ron & Habener, 1992).

3.6. Blocking the Activation of Transcription Factors

The transmission of activation is mediated through the activation domain of the *trans*-acting factor. Blocking of the activation domain, results in a repression of the activator.

Retinoblastoma gene product

In the G1 phase of the cell cycle, the retinoblastoma gene product (Rb) plays a central role in directing the fate of the cycle. The Rb is present in two isoforms during the cell cycle: under- and hyper-phosphorylated. In the G1 phase the Rb is underphosphorylated, but as the cell progresses into the S phase, the protein becomes hyperphosphorylated (DeCaprio *et al.*, 1989). In the G1 phase, the Rb interacts with transcription factor E2F, which is released upon phosphorylation of Rb (Chellappan *et al.*, 1991). The E2F which forms a heterodimer with an unrelated protein DP-1, induces quiescent cells to enter into the S phase (Girling *et al.*, 1993; Johnson *et al.*, 1993). The Rb interacts and blocks the activation site on E2F, and thereby prevents it from inducing activation (Hagemeier *et al.*, 1993).

GAL4

In yeast, galactose induces transcription of genes required for metabolism of galactose and melibiose, whereas glucose represses this transcription. The central protein in this metabolic control of transcription, is GAL4, a DNA-binding protein, with six cysteine residues, coordinating two zinc ions in its amino terminal end. GAL4 binds to the major groove of the DNA as a dimer to a symmetrical 17-base-pair sequence.

where the zinc cluster and adjoining residues make base-specific contacts (Marmorstein *et al.*, 1992). The GAL4 has two activation domains, rich in acidic residues (AR1 and AR2), the first in the amino terminal half and the second at the carboxyl terminus. The major activation domain AR2 is suggested to be an antiparallel β sheet, where one face is the actual activation domain, and the other side is a binding site for a repressor, GAL80 (Leuther *et al.*, 1993). Upon activation by galactose, the protein GAL3 is believed to mediate the signal to the GAL4-GAL80 complex causing conformational changes in the complex, resulting in activation. (Bhat & Hopper, 1992). After galactose activation, the GAL4 and GAL80 still remain bound to the DNA (Leuther & Johnston, 1992).

N-CoR

The members of the non-steroid receptor superfamily bind to the consensus sequence, AGGTCA that can be configured in a variety of structural motifs. These proteins can bind either as a homodimers, heterodimers or as a single protein, depending on the organization of the elements and receptors. A critical component in heterodimer formation is the retinoic X receptor (RXR), which binds with higher affinity to the DNA element than the homodimers (Mangelsdorf & Evans, 1995). Depending on the context of the elements, the same two heterodimers can either activate or repress transcription. When RXR forms heterodimers with a retinoic acid receptor (RAR) bound to two direct repeats and spaced by five nucleotides (DR5), transcription is activated (Durand *et al.*, 1992). Conversely, the same receptors repress transcription when bound to DR1 (Mangelsdorf *et al.*, 1991). Subsequently it was found that the polarity of receptors binding to the elements played a crucial role. On a DR5 element, the RXR binds to the 5' element, but in DR1, it binds on the 3' element, with opposite binding for the RAR (Kurokawa *et al.*, 1994). This difference in the activity was recently found to stem from a new factor, nuclear receptor co-repressor (N-CoR). In a non-activated state, N-CoR

binds to RAR, when bound to DR1 or DR5, in complex with RXR. Upon activation with retinoic acid, the DR5 bound RXR/RAR becomes an activator, by releasing N-CoR. In contrast, on DR1, with N-CoR bound to the RAR/RXR complex, activation with retinoic acid has no effect on dissociating the co-repressor. Instead, both are concomitantly bound and the complex represses transcription. The N-CoR was further found to interact with thyroid hormone receptor. The N-CoR interacts with the hinge between the DNA and ligand binding domain of the receptor. This domain is well conserved in all non-steroid binding receptors (Chen & Evans, 1995; Hörlein *et al.*, 1995; Kurokawa *et al.*, 1995).

3.7. Inhibition of the Assembly and Activation of the Transcriptional Machinery

Several factors have been characterized which can block the assembly of the initiation complex. This is a diverse group of proteins that exercise different modes of action to repress transcription.

HMG1

The high mobility group 1 protein (HMG1) is a non-histone chromatin associated protein. HMG1 has been shown to selectively bind cruciform DNA, rather than specific DNA sequence. Recently the HMG1 was found to repress transcription by binding to DNA-bound TBP (Ge & Roeder, 1994). This interaction prevented further assembly of GTF. In contrast, when TFIIA was first allowed to incorporate into the TBP-DNA complex, with or without TFIIB, the HMG1 did not exert any inhibitory effects. Furthermore the HMG1 repressed both basal and activated transcription. The HMG1 has an acidic domain and this region might interact with the basic domain of TBP.

The net effect of HMG1 seems to prevent the assembly of general transcription factors onto template committed promoter and this can be counteracted by TFIIA (Ge & Roeder, 1994).

Dr1

Dr1, a 19 kDa protein from HeLa cells, similarly to HMG1 prevents the incorporation of GTFs. It also interacts with the TBP, but unlike HMG1, TFIIA can not relieve this repression.

Dr1 is isolated as a phosphoprotein *in vivo*. The phosphorylated protein, replaces both TFIIA and TFIIB from the TBP-DNA complex. The absence of TFIIA and TFIIB prevents further assembly of the initiation complex. In the absence of the phosphate groups on Dr1, the protein binds to the TFIIA-TBP-DNA complex without any replacements of proteins, and blocks further assembly of GTFs. The phosphorylation therefore regulates the specificity of the Dr1 (Inostroza *et al.*, 1992).

Dr1 contains an acidic region, which interacts with TBP and is proposed to form an amphipathic helix. The basic region of TBP is believed to be the counteracting segment of Dr1. A separate region, rich in alanine residues, mediates the repression of the Dr1 on most promoters, but a different region within the protein also inhibits transcription. The Dr1 protein was shown to exist in a complex with TFIID in HeLa cells.

Repression mediated by Dr1 can be alleviated by DNA-bound activator, depending on their activation domains. A hybrid DNA-specific activator, containing activation domains from E1A and VP16 completely restored the promoter activity, whereas hybrid proteins containing the activation domain from Sp1 were only partially able to alleviate Dr1-mediated repression. A hybrid protein with the CTF (CAAT

binding transcriptional factor) activation domain had no effects on the repression from Dr1 (Yeung *et al.*, 1994).

Dr1 also represses transcription from genes transcribed by RNA pol III. presumably by preventing BRF, a TFIIB related factor, from being incorporated into the initiation complex. RNA pol I was not affected by Dr1 (White *et al.*, 1994).

Motl

Another repressor, Motl/ADA, depends on the hydrolysis of ATP or dATP for its function. In the absence of nucleotides, Motl binds to TBP complexed with the TATA box. Upon addition of ATP the Motl and TBP dissociate from the DNA, but the proteins still remain in complex. By adding bacterial alkaline phosphatase. the TBP will rebind to its cognate element, indicating that the repression is reversible and phosphatase inhibits the Motl repression. The Motl has further been suggested to be one of the TAFs, but in substoichometric ratio to TBP. Only around 3-5% of the TBP is in complex with Motl(Auble & Hahn, 1993; Kunzler *et al.*, 1994; Poon *et al.*, 1994). The action of Motl is blocked by association of TFIIA to TBP-DNA complex, but TFIIB shows less effect in relieving the repressor effects.

The site of action for Motl is within the basic region in TBP, but the complex is also supported by DNA interactions. The repressor acts mainly on basal and weakly activated transcription, in contrast to strongly activated transcription. where it has minimal effects.

By sequence comparison, Motl was found to have similarities to several proteins involved in transcription, DNA repair and putative helicases. Yeast cells lacking *Motl*. are viable, but can not tolerate elevated levels of TBP. This suggests that one function of Motl is to prevent undiscriminating activation of genes which should be silent. Another possible function might be to accelerate dissociation of TFIID from the DNA. after

transcription has started. In vitro studies have shown that TFIID dissociates as RNA pol II elongates the RNA chain (Kadonaga, 1990). Then the dissociated Mot1-TFIID complex, would allow template commitment by new TFIID to be regulated (Auble *et al.*, 1994).

HMG2

Another member of the high mobility group proteins, HMG2, has been identified as a general repressor of basal transcription. Like HMG1, HMG2 acts through DNA interactions but with a unique mechanism. After 3-5 nucleotides have been synthesized. HMG2 prevents chain elongation of the RNA. Addition of highly purified TFIIH and hydrolysis of ATP or dATP relieves this inhibition and transcription resumes (Stelzer *et al.*, 1994). Generally, the dependence of TFIIH in transcription seems to rely on DNA topology. Transcription from a negatively supercoiled gene for the immunoglobin heavy chain needs only TBP. TFIIB and RNA pol II. Relaxed template needs in addition TFIIF, TFIIE, TFIIH, and a fraction containing TFIIA and TFIIJ (Parvin & Sharp, 1993). In the presence of HMG1, transcription can only proceed with the addition of TFIIH, on both relaxed and supercoiled templates. The effects of TFIIH in this system is not clear, but the complex might promote local supercoils. an activity supported by intrinsic helicase (Stelzer *et al.*, 1994).

DNA Topoisomerase I

When growth of yeast cells reaches stationary phase. the amount of most mRNA decreases, due to repression of transcription. This general repression is mediated by DNA Topoisomerase I (Topo I) (Choder, 1991). This was later supported by using HeLa cell fractions in a reconstituted assay for transcription (Merino *et al.*, 1993). Topo I relaxes supercoiled DNA a through nicking-closing mechanism. and this causes the

release of energy (Caserta *et al.*, 1994). During repression of transcription the Topo I selectively interacts with TBP and can also be isolated in a complex with TFIID. In reconstituted assay for transcription, the TBP mediated Topo I effects on transcription. which could be prevented by the addition of TFIIA. However, Topo I dependent inhibition was not possible after formation of TBP-TFIIA DNA bound complex. Furthermore TFIIB did not have any effects on the activity of Topo I. This repression is alleviated by activation from activators containing acidic activation domains and the presence of Topo I even increased the activity of the promoter.

It was suggested that the presence of the Topo I in TFIID complex would inhibit transcription. After activation, the Topo I would be transferred to the elongation complex, and facilitate the elongation by removing helical tension in the DNA (Merino *et al.*, 1993).

IE86

Among the first genes to be expressed after infection by the human cytomegalo virus, are the immediate early (IE) genes, which control the expression of the viral genome. One of the products is IE86, a *trans*-acting factor which both activates and represses transcription, depending on promoter content. For repression of its own gene. IE86 binds to a *cis* repressor signal (crs) element, located between the TATA box and the transcriptional start site. The proximity of these two elements results in steric hindrance, which impairs the DNA interaction of these proteins, but does not prevent DNA binding. Instead, higher concentrations were needed for the proteins to bind to their cognate elements (Jupp *et al.*, 1993). Bound to the crs, the IE86 affects the assembly of the preinitiation complex (Wu *et al.*, 1993). The binding of TFIID. TFIIA and TFIIB is precluded by IE86 repression activity, which then specifically blocked the RNA pol II from being incorporated into the complex (Lee *et al.*, 1996).

Krüppel

Embryonic development is controlled by temporally and spatially regulated gene expression. *Krüppel* (Kr), a zygotically active gene in *Drosophila* during the blastoderm stage of embryogenesis is involved in the development of the thoraic and abdominal segments of the fly. The Krüppel protein forms a bell shaped gradient in the center of the blastoderm (Pankratz *et al.*, 1990). The protein was found to be a repressor of transcription when transformed into HeLa cells as a fusion protein with the DNA-binding domain from the Lactose repressor. The repressor domain of Kr was further localized to an alanine-rich region in the amino terminus (Licht *et al.*, 1990). Later it was reported that Kr had both activator and repressor activity in transcription using the Schneider S2/S3 cell line from *Drosophila* (Sauer & Jäckle, 1991). Furthermore, this activity was dependent on the concentration of the Krüppel protein. Low amounts of Kr led to activation of transcription, whereas repression was observed at a high Kr concentration. In contrast to a previous report, the amino terminus harbored the activation domain, whereas the carboxy terminus contained the repression domain (Sauer & Jäckle, 1991). Subsequent experiments indicated that the C-terminus was responsible for dimerisation of the Kr protein, and this was also concentration-dependent. For activation Kr is a monomer, but at higher concentration the protein forms a homodimer and represses transcription (Sauer & Jäckel, 1993). Furthermore, a concentration-dependent interaction was reported, between Kr and the GTF. For activation at low concentration, Kr interacts with TFIIB, whereas at high concentration after dimer formation, Kr interacts with TFIIEβ, but not with TFIIB and transcription is repressed (Sauer *et al.*, 1995). It is noteworthy that TFIIE stimulates phosphorylation of the CTD on the largest subunits of RNA pol II, through TFIIH, which is concomitant to promoter clearance. The Kr protein might prevent this step, resulting in a repression of transcription (Ohkuma & Roeder, 1994).

3.8. Repression of Transcription Mediated by Proteases

During *in vivo* infection by the poliovirus in HeLa cells, an inhibition of all three RNA polymerases is observed (Kääriäinen & Ranki, 1984). The decrease of RNA pol II-dependent transcription correlates with reduced appearance of TFIID (Kliewer & Dasgupta, 1988). The poliovirus synthesizes a protease, protease 3C, which recognizes and cleaves the glutamine-glycine peptide bonds (Lawson & Semler, 1990). The non-conserved amino-terminal domain on the TBP is cleaved off by the 3C protease, leaving the DNA-binding domain intact. Furthermore, transforming the cells with the protease alone, reduces the TFIID and decreases transcription by RNA pol II (Clark *et al.*, 1993). Poliovirus infection does not degrade many cellular proteins leaving the disappearance of TFIID as the main candidate for reduced transcription (Urzainqui & Carrasco, 1989; Clark *et al.*, 1993).

A thiol protease, m-calpein, is ubiquitously expressed across all vertebrate species (Goll *et al.*, 1992). The activity of the protease is regulated by calpstatin, an endogenous inhibitor and activated by intracellular Ca^{2+}, which is released through signal transduction cascade. The protease cleaves protein between two separate domains, rather than recognizing specific amino acid residues. The m-calpein cleaves several transcription factors, leaving the DNA-binding domain intact, but releasing the activation domain. Diverse factors, such as Pit-1, USF and ATF, are all cleaved by m-calpein, leaving the DNA-binding activity intact, which then acts as a repressor of transcription (Watt & Molly, 1993).

The preadipocyte factor, AEBP1, represses transcription of the gene for adipose P2, whose product is involved in triglyceride metabolism. During adipocyte differentiation, the AEBP1 disappears, concomitant with activation of several triglyceride genes. The repressor domain of the AEBP1, has carboxypeptidase activity,

that is necessary for the repression. Which proteins are targeted for digestion remains unknown (He *et al.*, 1995).

4. RESEARCH PROPOSAL

The control of transcription is modular in nature, balanced by counteracting factors and cues. The leading interest in gene expression has been on activators and enhancers, with less emphasis on repressors for transcription. Of crucial interest for molecular-marine biologists is to understand the control of the genes for AFP, which have maximum expression during the lowest metabolic period in the fish. The genes for AFP Type III are of special interest in this regard, due to the fact that they show the largest difference between summer and winter of all the various AFPs, and also have the highest activity during the summer (Davies *et al.*, 1988). This makes the Type III AFP a good model for studying seasonal gene control along with elucidating a new insight on specific regulation for the genes for AFP. The availability of functional AFP promoter from the ocean pout has cleared the way for making transgenic Atlantic salmon (*Salmon salar*). A chimeric gene construct in which the growth hormone gene from Chinook salmon (*Onchorynchus tshawytscha*) was placed under the control of the AFP promoter from ocean pout was used to produce the transgenic salmon that grew up to 15 times faster than normal (Du *et al.*, 1992a; Du *et al.*, 1992b).

Furthermore, the promoters for the genes for AFP have been shown to contain a strong repressor element (Gong *et al.*, 1991; Gong & Hew, 1993), but the elements were not localized. In the wolffish AFP gene, the region that harbors the repressor element has two 12 base pair repeated elements in two out of three characterized wolffish sequences. These 12 bp sequences are termed the "b" elements. Further upstream are three direct repeats of eight base pair sequences, called the "a" elements (Scott *et al.*,

1988). Interestingly, the ocean pout, Antarctic eel pout, and the one wolffish AFP sequence have all three "a" elements repeated, but they harbor only one copy of the "b" element (Hew *et al.*, 1988).

The objectives of the thesis can be summarized as follows:

1) To identify the negative element(s) upstream of the gene for AFP in wolffish.

2) To purify and characterize the putative binding factor(s) for the negative element

This region in the gene for AFP in wolffish has no similar sequences to known *cis*-acting elements, suggesting that the repressor might be interacting through a novel sequence. The repressor, based on DNA sequences, appears to be a novel factor, not described before. Combined, these objectives aim at elucidating a new factor involved in the transcriptional control of the gene for the AFP in wolffish.

Chapter 2

SEQUENCE COMPARISON OF THE GENES FOR THE TYPE III AFP AND THEIR FLANKING REGIONS

Abstract

Comparison of the Type III AFP genes and their flanking regions was undertaken by aligning all available DNA sequences from three different species; ocean pout, wolffish and Antarctic eel pout, all belonging to the same suborder: Zoarcoidei.

The sequence alignment reveals common core sequence for all the genes of Type III AFP. The core sequences are dispersed into at least five separate regions, intervened by segments unique for one or more of the AFP genes. The first core sequence includes an eight bp element repeated upstream of the promoter. These sequences, termed the "a" elements, have a consensus sequence TGTCATTG. Two wolffish sequences, out of three characterized, have two 12 bp direct repeats, between the "a" elements and the promoter. These sequences, termed the "b" elements, contain the sequence ATTGGCAGATGT. The first core sequence terminates after the first "b" element, which is present in all Type III AFP genes. The second core segment starts 43 bp upstream of the CAAT box and extends 21 bp downstream. The third core region starts with a putative Sp1 and includes the TATA box and terminates at the start site of transcription. The fourth core segment begins with a "c" element and terminates in repetitive TAAT sequences, termed "f" elements. The fifth core sector begins in the only intron of the genes for the Type III AFP and terminates coincidentally at its end. The sixth core segment is exon 2.

The region containing the putative repressor in wolffish is centered over two core segments which are separated by a unique sequence. Furthermore, the importance of a common transcription control and unique gene control is discussed.

Introduction

DNA sequences are available for several Type III AFP genes, which are from three different fish species. To date, no DNA sequence comparison of the Type III AFP genes has been undertaken from these species. These genes are present in various arrangement and copy numbers within the genome. In ocean pout, the gene copy number has been estimated to be around 150 copies per genome, whereas in wolffish they were estimated to be around 80 to 85 copies per genome. The Antarctic eel pout, has around 43 copies of the AFP genes per genome. Interestingly, ocean pout caught off the coast from New Brunswick, which is located more southerly than Newfoundland, has only one quarter of the AFP genes, compared to its northern counterpart (Hew *et al.*, 1988; Scott *et al.*, 1988; Wang *et al.*, 1995b). The Antarctic eel pout has been suggested to increase its efficiency of the AFP gene expression by splicing two exons or more. Both exons contain the ice-binding part of the AFP. The exons are intervened with a connector sequence, to form multiple exon coding regions. This putative protein then has multiple AFP domains, linked together with the connector sequences. Proteins with two antifreeze domains and a connector sequence have been purified from the Antarctic eel pout (Wang *et al.*, 1995a). In contrast, the genomic structure suggested that a gene with five or more antifreeze domains could exist (Wang *et al.*, 1995b).

The genes for the Type III AFPs are organized differently, depending on the fish species. Typically, the genes contain two exons, the first encoding the signal sequence, whereas the second exon contains the major component of the AFP. The ocean pout has a majority of the genes closely linked, but irregularly spaced (Hew *et al.*, 1988). In the wolffish, two thirds of the AFP genes are organized as tandem repeats, with each repeat containing two genes in inverted orientation. These repeats are clustered in

groups of six or more. The remaining one third of the genes are linked, but irregularly spaced (Scott *et al.*, 1988). The Antarctic eel pout has the genes in at least two different arrangements. The first group is linked with each gene in multiple tandem arrays. The other group has up to five repeated exons, equivalent with the 2nd exon in other Type III AFP genes. These exons are intervened with a connector sequence. An AFP component from Antarctic eel pout has been characterized with two antifreeze domains, intervened with a connector sequence (Wang *et al.*, 1995a). Only one gene with the latter structure has been characterized and it lacks the 5' end (Wang *et al.*, 1995b).

The AFP from Antarctic eel pout is not under seasonal control, and maintains high concentrations of AFP year around, as it inhabits sea waters which rarely rises above 2°C (DeVries, 1982). This is in contrast with fishes caught off the coast of Newfoundland that have high AFP concentration during the winter months and much lower AFP during the summer, reflecting the ambient temperatures which fluctuate from 0°C to 15°C. During the period with the lowest activity of AFP, which is in the summer months, the ocean pout AFP is still higher than the AFP during the winter peaks in other fish species off the coast of Newfoundland (Figure 1-1) (Davies *et al.*, 1988). This raises the question regarding the control of the expression for Type III AFP. The contrast between Arctic and Antarctic fishes in terms of seasonal expression of the Type III AFP eliminates photoperiod as a common factor in gene control. In contrast, sea water temperature could contribute to a common mechanism for the synthesis of the AFPs. If the photoperiod is a determining factor for the gene control of the Type III AFP in North Atlantic fishes, corresponding DNA elements, which mediate the light response, would be absent in the Antarctic eel pout which do not have seasonal changes of the AFP.

Prior to characterizing the repressor element, it is critical to gain information on weather 1) the flanking sequence is conserved, 2) the elements is partly or completely repeated elsewhere in the flanking sequences and 3) if known *cis*-acting sequences are present. Functionally important sequences are expected to be better conserved and should thereby be easier to identify than surrounding non conserved regions (Lewin, 1990).

A comparison of the sequences from various fish species living under different conditions, makes it feasible to delineate DNA regions involved in common control mechanism as opposed to unique DNA sequences, which would likely mediate more specific roles.

This alignment elucidates several core segments which are flanked by unique sequences for one or several of the genes for Type III AFP. A previously identified repressor region in wolffish, located between -220 to -90, (Gong & Hew, 1993) extends over the first and second core segment, and is intervened by unique sequence for wolffish AFP genes. Furthermore, several new repetitive sequence elements and putative regulatory elements are identified.

Materials and methods

Alignment of DNA sequences was done with Mac-DNASIS™ version 2.0 for Macintosh computer ™ (1992). DNA sequences were compared with Harr Plot with subsequent manual alignment of identical sequences (Collins & Coulson, 1987). Identical bases were labeled using Claris Works ™ (Holdaway *et al.*, 1994). The ocean pout 3 and 5 (OP3 and OP5) are from genomic clones λOP3 and 5, respectively (Hew *et al.*, 1988) (accession number J03924 and J03923, respectively, in GenBank®). The sequence of ocean pout 5 was combined with the overlapping sequences from Du *et al.* (1992) (accession number S65567 in GenBank®). All other sequences were from GenBank®, with the following accession numbers: Wolffish 1 (WO1): M22125: wolffish 3: L29011: Antarctic eel pout: U20437 to U20443 (AEP37 to 43). The sequences U2040 and U2042 are overlapping, and are here combined to one continuous sequence and referred to as AEP40. The second half of the AEP40 and AEP42 sequences have repeated sequences, which are renamed either as "AEP40 end 3494" or "AEP42 end 3494", reflecting the endpoint of identical sequences. These sequences are then realigned with identical sequences. The wolffish 2 is the complementary sequence of wolffish 1. The TATA and CAAT boxes, polyadenylation signal, transcriptional start site, coding regions, the connector sequences and the repeated elements "a" (Consensus sequence: TGTCATTG), "b" (ATTGGCAGATGT) and "c" (AAGTCTCAGCc/tAg/cAGCTT) were localized previously (Hew *et al.*, 1988; Scott *et al.*, 1988: Wang *et al.*, 1995b). A sequence corresponding to a second poly A signal, AATATA, was characterized by Sheets, *et al.* (1990).

Results

DNA sequences for the Type III AFP genes were compared. Sequences were from the ocean pout (OP), wolffish (WO) and Antarctic eel pout (AEP) AFP genes. This comparison elucidates common core sequences in the genes for the Type III AFP and their flanking sequences. These core regions are separated by unique sequences, present in one or more sequences for the AFP gene (Figure 2-1 and 2-2). Common core regions, starts with the available sequences from WO1 and 2 and OP3 sequences: ending with one "b" element, which contains the sequence ATTGGCAGATGT (position 2196 to 2376 in Figure 2-1). WO1 and 2, which have two repeated "b" elements, have the second "b" element within their unique region. The second core region begins 43 base pairs upstream of the CAAT box and ends 21 bp later (position 2808 to 2876 in Figure 2-1). The third core segment begins before a putative Sp1 element, includes the TATA box and extends to the transcriptional start site, which is a deoxyguanine nucleotide (position 2899 to 22951 in Figure 2-1). A small insert of four nucleotides in three Antarctic eel pout sequences is located between the third and fourth core segment which spans the next 68 bp (position 2955 to 3013 in Figure 2-1). The fourth core region includes one "c" element and six short "f" elements, which are direct repeats of TAAT sequences. The unique sequence starts immediately prior to the second "c" element. The first exon is not conserved as noted earlier (Wang *et al.*, 1995b). The fifth core region begins in the intron and terminates at the beginning of the second exon (position 3206 to 3288 in Figure 2-1). The sixth and final core sequence begins with the second exon and is terminated a few bp before the stop codons for translation (position 3301 to 3455 in Figure 2-1). Immediately after the second exon, the remaining sequences are unique for one or more clone (Figure 2-1 and 2-2).

Several new elements are deduced based on this sequence comparison. These elements, labeled in an alphabetical order, continuing from previous identified elements ("a, b and c"), with the exception of Sp1 and S element which both have similarities to the consensus sequence of the Sp1 element. The "a" element, which has the consensus sequence TGTCATTG in this alignment (position 2261 to 2285 in Figure 2-1), has two mismatches to a previously published element for the ribosomal proteins; **TGTGACTG** from the yeast *Schizosaccharomyces pombe* (boldfaced letters represent identical nucleotides) (Witt *et al.*, 1993; Witt *et al.*, 1995). A putative Sp1 element is located five bases upstream of the TATA box. This sequence. **CCTCCCACA** has seven out of nine nucleotides identical to the Sp1 consensus sequence: **c**/a **c**/t c/**t** **c**/a **GCC** c/t c/**a** (Faisst & Meyer, 1992). Another putative Sp1 element, termed "S" is located further upstream, between the "a" and "b" elements. The "S" element, has the consensus sequence **AACAGCCCC**, and has 9 out of 10 nucleotides identical to the consensus sequence of Sp1. Functional studies are needed to confirm if these elements are binding sites for the members of the Sp1 family or DNA-binding factors of the genes for the ribosomal proteins.

By extending the "c" element, which had previously been identified by Scott *et al.* (1989) by three residues, it becomes apparent that it is tripartite in nature (starting at positions 2955 and 3014 in Figure 2-1). The first triplet being AAGTCTC and the remaining two having the consensus sequence of AGC t/c t/a t/c. Other repeated sequences are labeled as "d. e, f, g, h and j". The four bp long "f" element. located between the "c" elements is repeated several times as direct repeates in all the AFP genes. In OP5 which harbors the highest number of these elements, the "f" element is repeated 7 times. All the "c and f" elements are within the transcriptional unit. The 15 base pair long "i" element has the consensus sequence ATTCAG t/g ATA a/c TAT a/g, and is intercepted by stretches of pTpA, and is located in the non conserved

intervening region of the wolffish 1 and 2 sequences (starts at position 4216 in Figure 2-1).

All Type III AFP genes have TATA, CAAT and Sp1 elements conserved. The distance of the CAAAT box varies due to an insert in AEP42 and 43. The sequences of AEP40 and AEP42 along with AEP40, have been suggested to belong to one and the same gene, with the 5' sequence missing (Wang *et al.*, 1995b). This alignment shows that AEP40, 42 and 43, have a complete promoter in front of the sequence, corresponding to the 2nd exon.

The connector sequence in the Antarctic eel pout sequences continues directly after the second exon. The connector and subsequent sequences have no similar sequences to either WO1 and 2 or OP3 and 5, to the sequences following the second exon. Located 14 bp after the end of the connector sequence, is a second poly A signal, with a AATATA sequence conserved. This sequence has been shown to act as a poly A site in point mutation experiments (Sheets *et al.*, 1990). Around 200 bp after the poly A signal, the sequences match to the 5' sequence of the WO3, starting around a thousand bp upstream of its transcription start site. This sequence similarity extends over 120 bp, and is interrupted by a unique sequence for the OP5. The sequence similarity between all the OP5, WO3 and AEP sequences starts again coincidentally with the first repeated "d" element, TATGAAA (position 2151 in Figure 2-1). This is followed as described above, with the core region that begins with OP3, WO1 and 2 sequences, spans over the "a" elements and terminates in the first "b" element, thereby closing the circle of identical sequences. The span of the core regions are therefore limited to sequences beginning around the "d" elements, and have a sharp termination coinciding with the end of the second exon. These core regions are flanked and intervened with several unique regions (Figure 2-3).

Figure 2-1. Sequence alignment of the genes and flanking region for Type III AFP from ocean pout, wolffish and Antarctic eel pout. The numbers refer to the first base of the sequence from ocean pout. The decreasing hierarchy of identical sequences is indicated with red, green or blue letters, respectively. Bases not conserved or not aligned are labeled with black. Inserts are indicated with an hyphen. Putative *cis*-elements, repeated sequences, poly adenylation signal and transcriptional stop site are boxed. Transcriptional start site is marked with an arrow and exons are overlined and transcriptional stop signals are underlined.

Connector sequence

Poly A

CAAT

PD1

TATA

E-elements

- 1st Exon -

- 2nd Exon -

Poly A

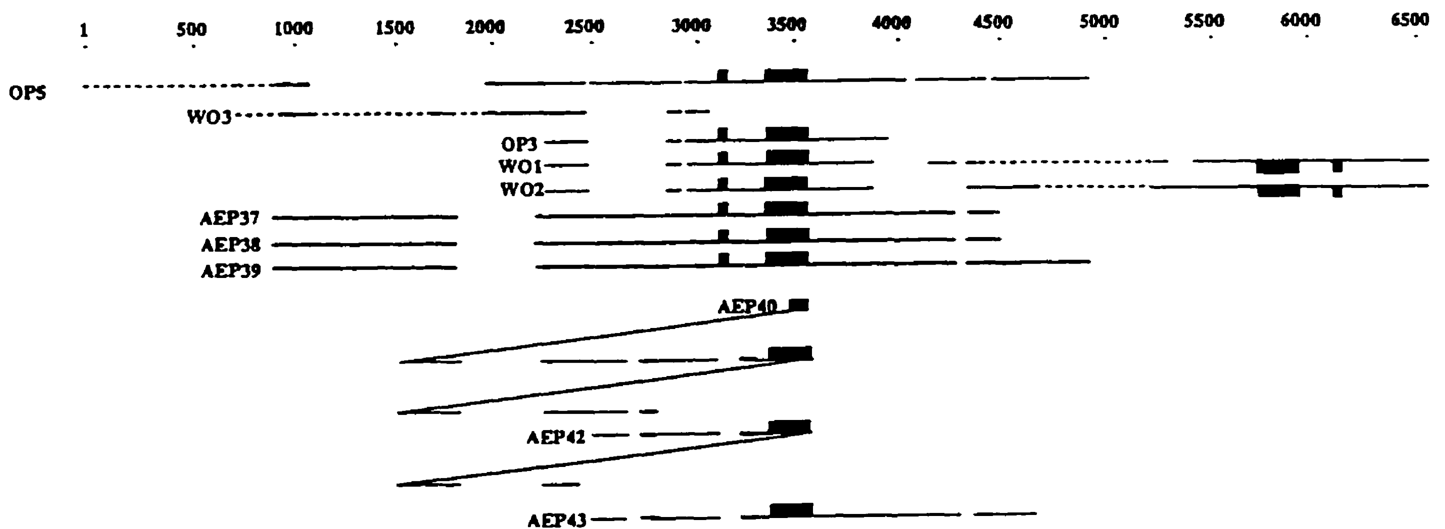


Figure 2-2. Schematic representation of sequence alignment of the Type III AFP genes and flanking sequences. The first and second exon are indicated as small and large filled boxes, respectively. Segments with similar sequences are indicated with a line. Thicker lines signify unique matching segments present in a sub-population of the sequence. Non matching regions are indicated as dotted lines. Numbers correspond to Figure 2-1.

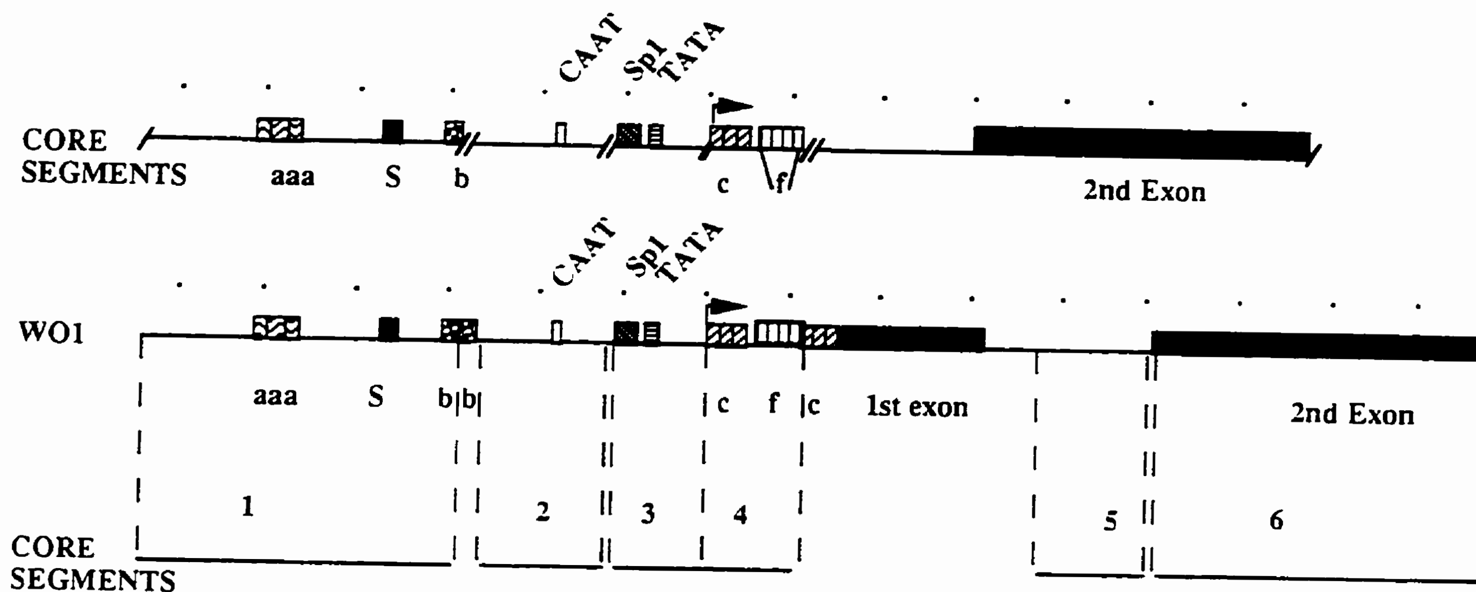


Figure 2-3. Schematic figure of core segments in the flanking region and the gene for the Type III AFP. Conserved region is indicated with a continuous line, and with the unique regions deleted. The filled boxes signify the exons and putative elements are boxed. The start site of transcription is indicated with an arrow. For comparison a schematic figure of wolffish sequence is aligned underneath. Dots mark 50 bp interval. Note that the core sequence contains only one "b" box sequence.

Discussion

The Type III AFP, which is limited to the suborder of Zoarcidae, occurs as a heterogeneous peptide population; this is a manifestation of a multiple gene family. These genes are linked, which is commonly seen for large gene families (Schimke, 1984). The wolffish AFP has two genes arranged in an inverted repeat, combined they make up one unit in a tandem array (Scott *et al.*, 1988). The Antarctic eel pout has the AFP genes linked and in direct repeats (Wang *et al.*, 1995b). In ocean pout the AFP genes are irregularly spaced (Hew *et al.*, 1988). No information is available about the organization of the gene clusters for the Type III AFP. It is noteworthy that the genes are always arranged in tandem repeats and clusters. In clustered structure, the genes are irregularly spaced and in non-repeated order as opposed to repeated structure, where the genes are regularly spaced and in tandem arrays. Recently, Graham (1995) has suggested that tandem arrays of the Type III AFP genes are unstable in the genome, whereas clustered genes are more stable during relaxed selection. The clustered genes were suggested to act as a genetic memory by giving rise to new tandem repeats during adaptation to colder environment. The clustered genes, through unequal recombination, get amplified into direct repeats. As the tandem arrays are thought to be easier to form, they are also believed to disappear faster during relaxed selection, through unequal recombination. Small genes have more chance of fitting into a direct repeat when being amplified. It has been suggested that the transition from a clustered gene to a tandem repeat can arise rapidly under appropriate selection (Graham, 1995). During amplification, preferred breaking points are used, but this can change as the copy number changes and in different cell types (Schimke, 1984).

The AFP genes have been suggested to have arisen during the onset of the last glaciation age which started in the beginning of Cenozoic age, approximately 25

million years ago. This period had several remissions during the last 5 million years (Scott *et al.*, 1986), with cold and warm periods interchanging. This should have resulted in relaxed and increased selection for these genes. The AFP Type III genes have evolved, responding to two different environmental surroundings. The North Atlantic and Antarctic ocean provide a different environment, based on a seasonal variation in water temperature. The temperature in the Arctic waters can fluctuate from around zero degrees to over 15°C whereas the temperature in the Antarctic sea waters remains around and below 0°C. This results in a seasonal expression of AFP in fishes off the coast of Newfoundland, whereas fishes in the Antarctic ocean maintain high AFP throughout the year (DeVries, 1980; Davies *et al.*, 1988).

This DNA alignment supports earlier notions that the Type III AFP genes are closely related. Furthermore, the surrounding regions are highly similar even though the genomic arrangements differ considerably. All these genes have common DNA sequences, or core regions, and unique segments, which are only present in one or few of the genes, but not in all. The common core segments that have several putative elements in common, such as "a, b, c, f, s and Sp1", are likely to provide a basic gene control for the Type III AFP. The "a" element has similarities to the element of previously identified ribosomal protein genes from *S. pombe*. These elements mediate activation for these genes (Witt *et al.*, 1993; Witt *et al.*, 1995). The 12 bp "b" element is repeated twice in WO1 and WO2 sequences, in contrast to only one "b" element in other genes which have been characterized, including a third wolffish sequence. This results in one "b" element in a core sequence, whereas the second element is unique to WO1 and WO2. The region previously identified as harboring repressor activity is located between -220 to -90 in WO1 and WO2 (positions 2300 to 2816, in Figure 2-1) (Gong & Hew, 1993), and has the "b" elements in the center. Thus, this segment is divided between two core regions, intervened with unique sequences.

The AEP40, 42 and 43 sequences were thought contain several exons corresponding to the second exon in other Type III AFP genes. The multiple 2nd exons, spliced together, were believed to form one protein with multiple antifreeze domains. This alignment shows that each exon contains TATA, CAAT and Sp1 elements, prior to the conserved transcriptional start site. Therefore, each exon has the potential to be transcribed from an independent promoter, or from one common promoter, forming one long transcript, which is subsequently spliced. This particular genomic clone lacks the 5' end (Wang *et al.*, 1995b). As a result, no information is available regarding the presence of the first exon. Combined, this raises the question if the AFP from the Antarctic eel pout has internal promoters, that could provide additional control for AFP gene expression. Interestingly, if this is true, then the internal promoters would direct the regulation of an internal AFP that lacks the pro-sequences and hence, would not be secreted. A weak Poly A signal is located 14 bp after the connector sequence, containing the sequence AATATA repeated (Sheets *et al.*, 1990). This could possibly allow the RNA polymerase to utilize a single promoter and by reading through several exons a multi-exon mRNA would be synthesized. which subsequently would be translated into multi-domain AFP as has been suggested (Wang *et al.*, 1995b; Wang *et al.*, 1995a). Alternatively, the promoters prior to each exon could generate nonsecreted AFP. The winter flounder has skin AFP that is not secreted. This variant, lacks both pre- and pro- sequences, which are required for the secretion of liver Type I AFP (Gong *et al.*, 1996).

The high year-round presence of AFP in Antarctic fishes, in contrast with the seasonal expression of AFPs in fishes from the North Atlantic waters, does not seem to support the notion that a common transcriptional control is under the influence of the photoperiod. If the North Atlantic fishes, which express the Type III AFP, are under the control of the photoperiod, this control would be expected to be mediated by DNA

segments which are unique to both ocean pout and wolffish. On the other hand, if sea water temperature is controlling the transcription of the AFP genes, these elements would be expected to be within the core region.

Chapter 3

CHARACTERIZATION OF A REPRESSOR ELEMENT FOR THE GENE FOR THE AFP IN WOLFFISH

Abstract

The region upstream of the gene for Type III AFP in wolffish was analyzed. The segment, spanning from -220 to -90 from the transcriptional start site, is shown to interact with several proteins in the nuclear extract from rainbow trout liver. The strongest signal is with nuclear extract prepared from liver. It is also demonstrated that this promoter region harbors a strong repressor element, which was subsequently located to two 12-base-pair direct-repeated sequences, called the "b" elements. When tested in transfection studies, the strongest repressor signal is seen in transfected fish cell lines. Furthermore, the -220 to -90 region is shown to have potential elements that bind to Sp1 transcription factor and a protein that binds to the proximal promoter of the ribosomal genes in fission yeast, *Schizosaccharomyces pombe*.

Introduction

Most AFP genes in fish are seasonally controlled a with maximal amount of protein produced during the winter season (Davies *et al.*, 1988). Although four different groups of antifreeze proteins have been characterized, the transcriptional control of the corresponding genes is poorly understood. The AFP gene copy ranges widely, from 12 in sea raven (Hayes *et al.*, 1989) to around 150 in ocean pout (Hew *et al.*, 1988). No similar sequences in the flanking DNA sequences between different groups of AFP genes has been reported, and hence there was no suggestion of common elements for gene control. In winter flounder and ocean pout, which synthesize Type I and III AFP, respectively, the tissue expression is ubiquitous, although predominantly in liver. In sea raven, which synthesizes Type II AFP, the expression is limited to liver (Gong *et al.*, 1992). The Type I AFP in the winter flounder is under the control of the photoperiod, with maximal expression during the shortest days of the winter (Fletcher *et al.*, 1978; Fletcher, 1979; Fourney *et al.*, 1984a). This is mediated by growth hormone, which is released from the pituitary gland depending on the day length, and which negatively controls the AFP expression (Fletcher *et al.*, 1978; Fletcher, 1979; Hew & Fletcher, 1979). This is in contrast to the cod, where cold sea water temperature increases the AFGP in the plasma (Kao & Fletcher, 1988).

The Type III AFP has the most limited phylogenetic distribution of all the AFPs examined. It is found only in the suborder of Zoarcidea (Davies *et al.*, 1988). It is interesting to note that the Type III AFP seems to be under unique control, when compared to other groups of AFP. When compared to other species off the shore of Newfoundland, the plasma from ocean pout has the highest thermal hysteresis, which is the difference between freezing and melting point. During the summer, the ocean

pout has even higher AFP activity than found during the winter, in any other species in the same region (Davies *et al.*, 1988). In contrast, the Antarctic eel pout, found off the shore of Antarctica, which also synthesizes Type III AFP, has high levels of AFP all year-round in response to constant cold sea water temperature (DeVries, 1980). The strong expression of the ocean pout AFP Type III promoters has been used to generate "all-fish" gene constructs to express the salmon growth hormone gene in transgenic Atlantic salmon (*Salmon salar)*. This resulted in markedly increased growth of the fish (Du *et al.*, 1992b).

Progressive deletion of the 5' end of the Type I, II and III AFP genes indicated that several regions of each gene mediated both positive and negative transcription. Transfection of these constructs into Japanese medaka embryos (*Oryzias latipes*) resulted in a developmentally dependent increase of transcription between days 5-13 after fertilization. This increased expression due to the AFP promoters was specific for the AFP constructs, as the controls, which contained a promoter from the Rous sarcoma virus or SV40 virus, did not show any marked increase in transcription (Gong *et al.*, 1991). All Type III AFP promoters have three 8 bp repeated elements upstream of the promoter (elements "a") (second chapter). Furthermore, two out of three wolffish promoters sequenced so far, have two 12 bp elements ("b" elements) repeated between the "a" elements and the promoter. All other characterized promoters for the AFP Type III, have only one "b" element. This includes two ocean pout sequences, one wolffish and seven Antarctic eel pout AFP genes (Hew *et al.*, 1988; Scott *et al.*, 1988; Wang *et al.*, 1995b). Deletion of the three "a" elements and surrounding regions in both wolffish and ocean pout decreased the promoter activity, suggesting that this region might contain positive element(s). When the sequence around and including the two "b" elements from wolffish was tested in conjunction with sea raven AFP promoter and thymidine kinase promoter from herpes simplex

virus, it reduced the transcription, indicating that this region contained a strong repressor element. Deletion of the region around and including one "b" element from ocean pout, also suggested that this region might contain a repressor (Gong *et al.*, 1991; Gong & Hew, 1993).

The unique nature of the Type III promoters makes them ideal candidates for studying seasonal and liver enriched gene expression. In this study it is shown that there are several proteins interacting with the -220 to -90 fragment from the AFP gene in wolffish including both of the "b" elements. Furthermore, in accordance with previous results, the -220 to -90 fragment is a strong repressor, with strongest transcriptional repression in the fish cell lines, from rainbow trout (RTH-149) and chum heart salmon cell line (CHH-1). The repressor element is furthermore localized to the two "b" elements. A new putative element, which has similarity to Sp1 DNA binding elements, is characterized. Also a new "a" element is also characterized and which is also shown to have resemblance to promoter elements for ribosomal proteins from *Schizosaccharomyces pombe*.

Material and methods

Preparation of Nuclear Extract from Liver from Rainbow Trout

Preparation of nuclear extract followed previous published protocols, with minor modification (Sierra, 1990). Livers were collected from decapitated rainbow trout (Rainbow Trout Hatchery, Pickering, Ontario). The livers were immediately frozen on dry ice and kept at -80°C, until processed. The liver was thawed and quickly processed, first by grinding with a kitchen meat grinder. The homogenate was mixed with ice-cold homogenization buffer consisting of 10 mM HEPES pH 7.8, 0.15 mM spermine, (Sigma) 0.5 mM spermidine (Sigma), 1 mM EDTA, 2.5 M sucrose, 10% glycerol and 1% low fat dry milk. The ratio between liver and buffer was 1:10 (v/v). Just prior to use, the following protease inhibitors were added to indicated concentrations: 1 mM PMSF (Sigma), 1 mM benzamidine (Sigma), 5 mM β-mercaptoethanol and 0.5 μg/ml leupeptin (Sigma). The extract was homogenized by one stroke in a pre-cooled Potter-Elvehjem homogenizer. The nuclear extract was loaded on top of a 10 ml of ice cold homogenization buffer, without the low fat dry milk, and centrifuged for 60 minutes at 24,000 rpm in a pre-cooled SW27 rotor in a Beckman centrifuge at 0°C. The supernatant was discarded, and the walls of the tube were washed with distilled water, taking care not to disturb the pellet, which contained the nuclei. The pellet was resuspended in 10 mM HEPES pH 7.8, 100 mM KCl, 0.1 mM EDTA, 10% glycerol, 3 mM $MgCl_2$ and just prior to use, the aforementioned protease inhibitors were added to the buffer, to the previously indicated concentration. The suspended nuclei were diluted to 0.5 mg/ml DNA according to the following formula: 1 O.D. = 50 μg/ml of DNA, as measured by a spectrophotometer in a 50 fold dilution in 0.5% SDS at 280 nm. To the suspended nuclei was added one tenth of the volume of 4 M $(NH_4)_2SO_4$; then the nuclei were

mixed gently for 30 minutes on ice. The lysed nuclei were centrifuged at 24,000 rpm in SW30 rotor for 60 minutes in a Beckman centrifuge at 0°C. The supernatant was carefully removed and 0.3 g of $(NH_4)_2SO_4$ was added slowly for each milliliter of the solution. Once all the salt had been dissolved, the solution was stirred for an additional 30-60 minutes on ice, followed by a centrifugation in SW30 at 27 rpm for 30 minutes in Beckman centrifuge at 0°C. The supernatant was discarded. The pellet was stored at -80°C or dissolved in dialysis buffer containing 25 mM HEPES buffer pH 7.8, 0.1 mM EDTA, 40 mM KCl, 10% glycerol, 1 mM DTT and the aforementioned protease inhibitors at the same concentrations. The desalted pellet was and dialyzed overnight against 100 times volume of buffer with two changes of the buffer at 4°C. Then the extract was removed and centrifuged in a Eppendorf microfuge for 5 minutes: the supernatant was aliquoted into Eppendorf tubes and stored at -80°C until used.

Subcloning

The DNA construct, pWON, which contains the -220 to -90 fragment (positions 2300 to 2816, in Figure 2-1), from the wolffish (WO1) promoter was cut with Xba I and BamH I; and the 130 bp fragment was isolated by agarose gel electrophoresis. The fragment was cut out of the gel and electroeluted into a dialysis bag as described (Sambrook *et al.*, 1989). The solution was extracted with phenol/chloroform, and the DNA was precipitated with 2.5 volumes of ethanol and 0.25 volumes of 10 M NH_4 acetate and centrifuged in an Eppendorf microfuge at full speed, washed with 70% ethanol and centrifuged again. The pellet was dissolved in 10 mM Tris base pH 8.0 and 1 mM EDTA and quantitated by spectrophotometer as described (Sambrook *et al.*, 1989). DNA oligomers were purchased from the Biotech Service Center, Toronto. Two complementary DNA oligomers, containing the two 12

bp repeat, (ATTGGCAGATGT) corresponding to the two "b" elements

```
5' AGCTTATTGGCAGATGTATTGGCAGATGTA 3'
      3' ATAACCGTCTACATAACCGTCTACATTCGA 5'
```

(positions 2364 to 2388 in Figure 2-1), were annealed at 1 mg/ml, boiled, and slowly cooled to room temperature.

The 130 bp DNA fragment and the "b" oligomer were made blunt ended by incubating with 2 mM of each dNTPs and one unit of Klenow fragment for 30 minutes at 37°C in 1x OPA buffer (10 mM Tris-acetate buffer pH 7.5, 10 mM magnesium acetate, and 50 mM potassium acetate). The enzyme was separated from the solution by spinning the solution through Probind filter (Millipore). The vector pBLCAT2 (Luckow & Schütz, 1987), was digested with Hind III or Sma I, and the Hind III site was made blunt ended as described above. Both restriction sites were dephosphorylated by adding 5-10 units of calf intestinal phosphatase (Boehringer Mannheim) in 20 µl of the solution. Enzymes were separated by spinning the solution through Probind filters. The inserts were ligated into the vector by using 100 times excess of the insert to the vector using ligase and buffer from GIBCO/BRL and incubated over night at 15°C. Of this mixture, 1-5 µl were used to transform 100 µl of the E. coli cell-strain DH5α™, made competent by $CaCl_2$ as described (Miller, 1987) and seeded on LB agar plates and grown overnight at 37°C. The presence and orientation of inserts was confirmed by DNA sequencing.

Radiolabelling of DNA

Two different methods were used for labeling the DNA. Approximately 25 ng of DNA was incubated in a volume of 50 µl with 1x kinase labeling buffer (70 mM Tris buffer pH 7.6, 10 mM $MgCl_2$ and 5 mM DTT), 1-5 units of polynucleotide

kinase and 15 μl of [γ-^{32}P]ATP (Amersham); the mixture was incubated at 37°C for 30 minutes. The free isotope was separated from the labeled DNA by passing the solution through a ProbeQuant G-50 micro column (Pharmacia). First, the column was pretreated by spinning at 3000 rpm for one minute in an Eppendorf centrifuge, washing once with 10 mM Tris base, pH 8,0 and 1 mM EDTA and re-centrifuging at 3000 rpm as before. Second, the DNA solution was applied to the column and centrifuged for two minutes at 3000 rpm in an Eppendorf centrifuge. For DNA footprinting, where only one end of the fragment was radiolabeled, the Bam HI/Xba I fragment was partially filled in with dGTP and dCTP using Klenow fragment; then the fragment was purified by spinning the solution through a Probind filter and finally labeled either with 15 μl of [α-^{32}P]dATP on the Bam HI site or with 15 μl of [α-^{32}P]dTTP (Amersham) on the Xba I site, using 1x kinase buffer in a total volume of 50 μl. The DNA was separated from free isotopes using ProbeQuant G-50 micro columns, which were first prewashed with water, prior to centrifugation. The prewashing of the ProbeQuant G-50 micro columns was done by centrifuging the columns at 3000 rpm in an Eppendorf centrifuge for one minute. The columns were loaded to the top with water and centrifuged again as before. This was repeated 4-5 times.

Experiments with Gel Mobility Shift

Mobility shift experiments were carried out by mixing together 1 μl of 1 μg/μl poly[dI·dC]·poly[dI·dC] (Pharmacia) and 5-10 μg of nuclear extract with 1 μl of MS buffer (100 mM HEPES pH 7.8, 5 mM $MgCl_2$, 200 mM KCl, 20 mM DTT, 2.5 mg/ml BSA, 80% glycerol); the volume was adjusted to 19 μl with water and incubated on ice for 30 minutes. Then 1 μl of radiolabeled probe was added (0.5 ng) and incubated for an additional 30 minutes on ice. Just prior to loading the samples

on the gel, 1 µl of bromophenol blue solution was added (0.1% bromophenol blue in 1/20 MS buffer). Electrophoresis was done in 5% acrylamide gel (80:1, acrylamide:BIS) in a Hoefer minigel apparatus with 1x TBE buffer that consists of 89 mM Tris base, 89 mM boric acid and 2 mM EDTA. Gels were run for approximately 20 minutes at a starting temperatures of about 0°C with 35V/cm. Each gel was exposed to a Kodachrome film (Kodak) overnight at -70°C, with intensifying screens, and developed by an automated film processor (Toronto General Hospital). For competition experiments, either unlabeled "b" oligomers or unlabeled "a" oligomer were used.

```
5' CTATGTCATTGTGTCATTGTGCTATTG 3'
   3' TACAGTAACACAGTAACACGATAACAG 5'
```

This "a" oligomer contains three direct repeats of eight bp "a" elements as found upstream of the wolffish promoter (2261 to 2283 in Figure 2-1). The three "a" elements contain the sequence TGTCATTG, TGTCATTG and TGCTATTG. The oligomer was preincubated with the nuclear extract in 100 fold excess to radiolabeled probe. The "a" oligomers were hybridized as described in the subcloning section. For radioisotope counting of the shifted bands, each lane was excised from the gel and suspended in a scintillation liquid, aquasol-2 (Du Pont) and counted in a scintillation counter 1219 Rackbeta (LKB) for five minutes.

Measurements of Protein Concentration

The concentration of proteins was measured by the Bradford method (Stoscheck, 1990). To 2-10 µl of sample was mixed 900 µl of Bradford reagent (100 mg of Serva Blue G. 100 ml of 85% phosphoric acid and 50 ml of 95% ethanol), and the volume was brought to 1 ml with water. The optical density was read at 590 nm

in a spectrophotometer and compared to a standard curve created by known concentration of BSA (Sigma), prepared simultaneously, in Bradford reagent.

Cell Cultures

RTH-149 cells, from rainbow trout liver and CHH-1 from a chum salmon (*Oncorhynchus keta*) heart cell line were grown at 19°C in sealed flasks in minimal essential medium with Earl's salts, L-glutamine and non-essential amino acids. The medium, 9.6 g, was dissolved in 920 ml of water with 1.2 g of HEPES and 0.37 g $NaHCO_3$. The medium was adjusted to pH 7.4 with KOH. To the medium was added 10 ml of an antibiotic-actinomycotic (100X) (GIBCO/BRL), followed by filter-sterilization with vacuum filters (Corning). Fetal bovine serum (FBS)(GIBCO/BRL) was heat treated at 56°C for one hour and added to the medium to 15%. HeLa and HepG2 cells were grown in minimal essential medium with high glucose (Hospital for Sick Children) in 10% FBS with 1x Antibiotic-actinomycotic and 100 mM HEPES (GIBCO/BRL). In addition, the medium for HepG2 had 1x sodium pyruvate and 1x nonessential amino acids (GIBCO/BRL). The medium was stored at 4°C until used. HeLa cells, a human carcinoma cell line and HepG2, a human hepatoma cell line, were grown in Petri dishes at 37°C in a humidified chamber, supported by 5% CO_2. For passage of the cells, the medium was aspirated and the cells were washed once with 1x PBS,, then the cells were covered with 1x trypsin solution. The trypsin solution was prepared by diluting 100x trypsin (GIBCO/BRL) to 1x concentration by diluting it with 1x PBS. The cell culture was left at room temperature until all the cells had been dispatched. Immediately, medium with FBS was added to stop the enzyme, and cells were placed into new bottles or Petri dishes.

Transfection of Cells

Cells were seeded at $0.5x10^5$ cells per Petri dish (60x10, Corning) with the appropriate medium. The following day the medium was aspirated and fresh medium was added to the culture dishes. DNA constructs were transfected into the cells with the calcium phosphate method with minor modifications (Okayama & Chen, 1991). To 10 μg of DNA construct solution was mixed 1 μg reporter gene, 31 μl of filter-sterilized 2.5 M $CaCl_2$ and the volume was brought to 250 μl with water. Reporter plasmids were either pCMVIEP (MacGregor & Caskey, 1989) with the β-galactosidase gene, or pRSVLUC with the luciferase gene. To the mixture with the DNA constructs was added 250 μl of filter-sterilized 2x HBS composed of 280 mM NaCl, 1.5 mM Na_2HPO_4, 50 mM HEPES, pH exactly to 7.10 with NaOH at 25°C. The resulting mixture was kept at room temperature for no longer than one minute and then added dropwise to the cell medium with constant shaking of the Petri dish. After incubating the cells for 24 hours, the medium was aspirated and the cells were washed once with 1x PBS. The RTH-149 and CHH-1 cells were not treated any further and fresh medium was added. HeLa and HepG2 cells were treated with 12% glycerol in 1x HBS for 3 minutes at room temperature. The solution was quickly aspirated and washed twice with 1x PBS and finally fresh medium was added. Cells were grown for an additional 24 hours, the medium was aspirated and 200 μl of 1x PBS was added to each plate. The cells were scraped off with a rubber policeman and collected into an Eppendorf tube. The cells were washed with 1x PBS and centrifuged; the liquid was aspirated off and the washing was repeated. The HeLa and HepG2 cells were disrupted by adding 250 μl of lysis buffer containing 250 mM Tris base pH 7.8 and 0.1% Triton X-100. The cells were suspended by gentle pipetting and then freeze-thawed by transferring them between a dry ice and a 37°C water bath. The RTH-149 and CHH-1 cells were similarly treated except that the lysis

buffer was 200 µl of 250 mM Tris buffer pH 7.8 and 0.5% Triton X-100. The extract was centrifuged in an Eppendorf microfuge at full speed for 5 minutes;; the precipitate was discarded and the supernatant was stored at -80°C or processed immediately.

Estimation of Promoter Strength

Two measurements were done for each transfected cell culture. First. thechloramphenicol acetyl transferase (CAT) assay indicated the relative promoter strength, and second either the β-galactosidase or luciferase assay was performed and used to correct for variations in transfection. The values from the CAT assays were divided by the values for either the β-galactosidase or luciferase assay to give corrected values. Calculations, graphs and error bars were generated by CA-Cricketgraph III (Computer Associates International, Inc.)

Chloramphenicol Acetyltransferase Assay

The CAT assay was performed essentially as described for all cell lines with minor modifications (Sambrook *et al.*, 1989). To 100 µl of the cell extract from the transfection experiments was added 40 µl of reaction mixture, which was made by adding together: 19 µl of 75 µM HCl, 0.5 µl acetyl CoA (Sigma), 20 µl of 5 mM chloramphenicol and 0.5 µl of ^{3}H acetyl CoA (200 mCi/mmol). The reaction was incubated overnight at 37°C and then transferred to a scintillation counting vial. Into each vial was then added 100 µl of 5 M urea to stop the reaction and 5 ml of Econofluor-2 (Du Pont). Radioactivity in each vial was counted in a Rackbeta scintillation counter.

Measurement of β-Galactosidase Activity

The HeLa and HepG2 cell lines were co-transfected with a pRSVgal construct (MacGregor & Caskey, 1989) and measured as previously described with minor modifications (Sambrook *et al.*, 1989). To 30 μl of the transfected cell extract was added 3 μl of 100x Mg solution, (0.1 M $MgCl_2$, 4.5 M β-mercaptoethanol), 66 μl of ONPG (*o*-nitrophenyl-β-galactopyranoside) and 201 μl of 0.1 M sodium phosphate buffer pH 7.5. This mixture was incubated for 1 hour at 37°C and the reaction was stopped by the addition of 500 μl of 1 M Na_2CO_3. The optical density was read at 420 nm.

Measurement of Luciferase Activity

To 100 μl of the transfected cell extract was added 100 μl of 50 mM Tris base pH 7.8, 1 mM DTT and 0.1 % Triton. Just prior to measurement, 15 μl of cocktail mixture, consisting of 750 mM Tris buffer pH 7.8 with MES, 150 mM Magnesium acetate and 40 mM ATP was added. The luminometer (Berthold) added 60 μl of 10 mg/ml luciferin (in 10 mM Na_2HP0_4 buffer pH 7.5) and measured light emission for 10 seconds.

DNA Footprinting

DNA footprinting was performed as described (Henninghausen & Lubon, 1987) with several modifications. Various amounts of nuclear extracts (from 0 to 200 μg) were mixed with 5 μg of poly[dI·dC]·poly[dI·dC] and 5 μl of MS buffer in a total volume of 95 μl and incubated on ice for 30 minutes, followed by the addition of 2.5 ng of probe in a volume of 5 μl. The probe was the 130 bp fragment from the AFP promoter in wolffish, containing the two "b" elements. The DNA was end labeled on only one strand. The tubes were incubated for an additional 30 minutes on

ice and then warmed to room temperature for 5-7 minutes. To each tube was added 4 μl of DNase I solution consisting of: 62.25 mM $MgCl_2$, 25 mM $CaCl_2$, 12.5 mM NaCl, 5 mM Tris base pH 7.5 and 0.25 mM DTT with 3.3 mU/ml of DNase I (Promega). The tubes were incubated at room temperature for 20 seconds. Immediately, 200 μl of stop solution was added, consisting of: 0.375% SDS, 15 mM EDTA, 100 mM NaCl, 100 mM Tris base pH 7.8, 50 μg/ml sheared salmon sperm DNA and 100 μg/ml proteinase K (Boehringer Mannheim). The salmon sperm DNA was sheared by passing the DNA solution through an 18 gauge needle and boiled for 10 minutes. After the addition of the stop solution the tubes were kept on ice, until they all had been processed, and then incubated simultaneously at 37°C for 15 minutes. The solution was centrifuged three times through Probind filters, followed by 2 minutes incubation at 90°C. This was followed by extensive phenol/chloroform extraction until no precipitated protein was observed. The solution was brought down to a volume of about 25 μl with a Speed Vac Concentrator. The DNA was then precipitated by the addition of 1 μl of 20 mg/ml glycogen (Boerhinger Mannheim) and 0.25 times volume of 10 M ammonium acetate and 2.5 volume of ethanol, washed once with 70% (v/v) ethanol and dried. The pellet was dissolved in 3 μl of Sequencing Stop solution (Pharmacia) and loaded on a denaturing DNA-sequencing gel.

Maxam and Gilbert Reaction

Size markers were created with the Maxam and Gilbert method as described (Sambrook *et al.*, 1989) with minor modification. In a glass tube, 4 μl of 1 mg/ml sheared salmon sperm DNA was mixed with 10 μl of 2.5 ng/μl of single stranded radiolabeled probe and 3 μl of 8.8% (v/v) formic acid and then the mixture was incubated for 7 minutes at 37°C, followed by the addition of 150 μl of 1 M

piperidine, and incubated at 90°C for an additional 30 minutes. The solution was cooled on ice for 2-3 minutes and the DNA was precipitated by the addition of 1.2 ml of 1-butanol, then vortexed for 1 minute and centrifuged in a microcentrifuge for 5 minutes. The supernatant was completely removed and the pellet was dissolved in 150 μl of 1% SDS and re-precipitated again by the addition of 1.5 ml of 1-butanol, vortexed and centrifuged as before. The supernatant was completely removed and the pellet was dried in a Speed Vac Concentrator. The pellet was then repeatedly dissolved in 20 μl of water and dried over a period of 6-8 hours. The DNA was then dissolved in Sequencing Stop Solution (Pharmacia), boiled for two minutes, cooled on ice, loaded on to the sequencing gel, and then run in parallel with the DNA-footprinting samples.

Large Scale Preparation of DNA Plasmids and Sequencing of DNA

Large scale preparation of plasmid DNA for transfection experiments was done with a Maxiprep kit from Qiagen, as described by the manufacturer. The pure plasmids were dissolved in water. DNA sequencing was done with [α-^{32}P]dATP or [α-^{32}P]dCTP to enhance reading close to the primers. A sequencing kit from Pharmacia was used according to the manufacturer's instructions. Sequencing reactions were run on 7% denaturing acrylamide gel (28:2, acrylamide:BIS), in 1x TBE buffer as described (Ausubel *et al.*, 1992). Gels from DNA-footprinting were fixed for 10 minutes in 10% methanol and 10% acetic acid, then dried before exposed to X-ray film.

Results

Experiments with Mobility Shift Assay

Nuclear extract from various sources was tested with the DNA fragment from -220 to -90 from the wolffish AFP promoter. The strongest signal was observed in liver extracts from winter flounder and rainbow trout (Figure 3-1, lanes 3 and 4), but much weaker signals were observed in extracts from HeLa and GH_4 cell lines (Figure 3-1, lanes 1 and 2), suggesting that an ubiquitous but liver enriched, factor(s) is interacting with this DNA region. The same DNA fragment, in combination with nuclear extract from rainbow trout liver, showed several bands when increasing amounts of nuclear extract were added (Figure 3-2A). To test if this is due to synergistic binding of several proteins, the bands were cut out and radioactivity was measured. When the radioactivity is plotted against the amount of nuclear extract. a possible sigmoidal curve was observed. This curve was reproducible (Figure 3- 2B). Combined, this suggests that several proteins might be involved in interaction with the -220 to -90 bp fragment.

An oligomer corresponding to the two "b" elements was synthesized and used in the mobility shift assay. Tested with nuclear extract from rainbow trout liver. a shift indicated a protein-DNA interaction (Figure 3-3B, lane 4). To discern if the protein is interacting with the "b" element or a possible overlapping "a" element sequence (Figure 3-3A), a competition was performed. As expected, no band was observed when 50 fold excess of non-labeled "b" fragment was included (Figure 3-3B, lane 2). In contrast, competition with a 50 fold excess of non-labeled "a" fragment, did not affect the signal of the shifted DNA-protein complex (Figure 3-3B lane 3). This experiment also indicates that the protein-DNA interaction is sequence specific for the "b" element.

Studies with DNA-Transfection

Previous studies had suggested that the -220 to -90 fragment in the AFP from wolffish contained a repressor element (Gong & Hew, 1993). In order to localize the repressor element, several DNA-constructs were made. The -220 to -90 fragment was ligated in front of a thymidine kinase (TK) promoter from herpes simplex virus to form the construct: pBLCAT-220/-90 (Luckow & Schütz, 1987). The oligomer for the two "b" elements were ligated at different distances and orientation with respect to the TK promoter (Figure 3-4). The construct pBLCATb5'+ has the "b" element located upstream of the TK promoter (5') whereas the pBLCATb3'+ and pBLCATb3'- have the element downstream of the promoter (3'), either in the same orientation as they occur in the wolffish promoter (+) or in the opposite orientation (-). These constructs were tested for promoter activity in several cell lines using the chloramphenicol acetyl transferase (CAT) assay. HeLa and HepG2 cell lines were co-transfected with the β-galactosidase construct and RTH -149 and CHH-1 cells were co-transfected with the luciferase construct, to correct for variations in transfections. The pBLCAT-220/-90 construct, with the -220 to -90 fragment, showed little effect in HeLa cells. In contrast, the construct in HepG2, cells it has a possible enhancer effects. The pBLCATb5'+ construct has repressor activity in transfected HeLa cells, whereas both the pBLCATb3'+ and pBLCATb3'- show much less repressor activity when transfected into the same cell line. The HepG2 cell line shows little if any repressor activity, with the exception of construct pBLCATb5'+ (Figure 3- 5, A and B). In contrast, the fish cell lines, RTH-149 and CHH-1 have lower transcriptional activity for all the constructs tested, with lowest activity in construct pBLCATb3'– in both fish cell lines. A stronger repressor signal is in the RTH-149 cell line than in the CHH-1 cell line. Therefor a repressor element in the fragment -220 to -90 in the wolffish promoter is localized to the two "b" elements (Figure 3-6, A and B).

Studies with DNase-Footprinting

To further delineate the protein-DNA interaction in the -220 to -90 fragment, it was tested with nuclear extract from rainbow trout liver in a DNase-footprinting experiment. Several regions were protected in this assay (Figure 3-7). Both the "b" repeats show protection from -132 to -158 indicating that both elements are occupied under these conditions. Additional regions with sequence similarities to the Sp1 element from -200 to -182 also show protection in this assay. The consensus sequence for the Sp1 element is A/CT/CC/TAGCCT/CA/T (Faisst & Meyer, 1992) and the partial sequence protected in this assay is **AACAGCCCC**, (bold-faced letters indicate identical bases). There are several proteins capable of affecting transcription through the so-called G/C boxes, which includes the Sp1 binding element (Kingsley & Winoto, 1992). The consensus sequence for the "a" element, TGTCATTG, has similarities to a third weakly protected region which contains the sequence **TGTGCATG**, (bold-faced letter indicate identical sequences) (Figure 3- 7). This further supports that several proteins are interacting with the -220 to -90 fragment, as was suggested with the mobility shift experiment.

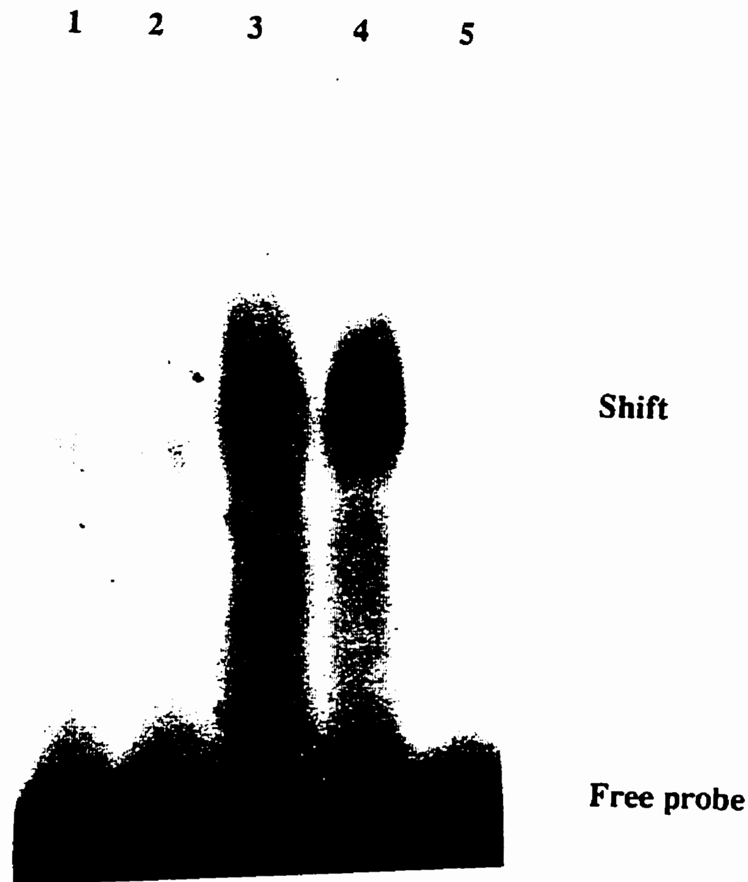


Figure 3-1. Binding of nuclear proteins with the -220 to -90 DNA-fragment from the AFP promoter wolffish. A DNA fragment from the -220 to -90 region from the AFP gene in wolffish was tested with nuclear extract from various tissues and organisms. Nuclear extract (10 μg) from HeLa cell line (lane 1), GH_4 cell line (lane 2), rainbow trout liver (lane 3) and winter flounder liver (lane 4) was tested with the -220 to -90 fragment in a mobility shift assay (total 20 μl). Lane 5 has no nuclear extract.

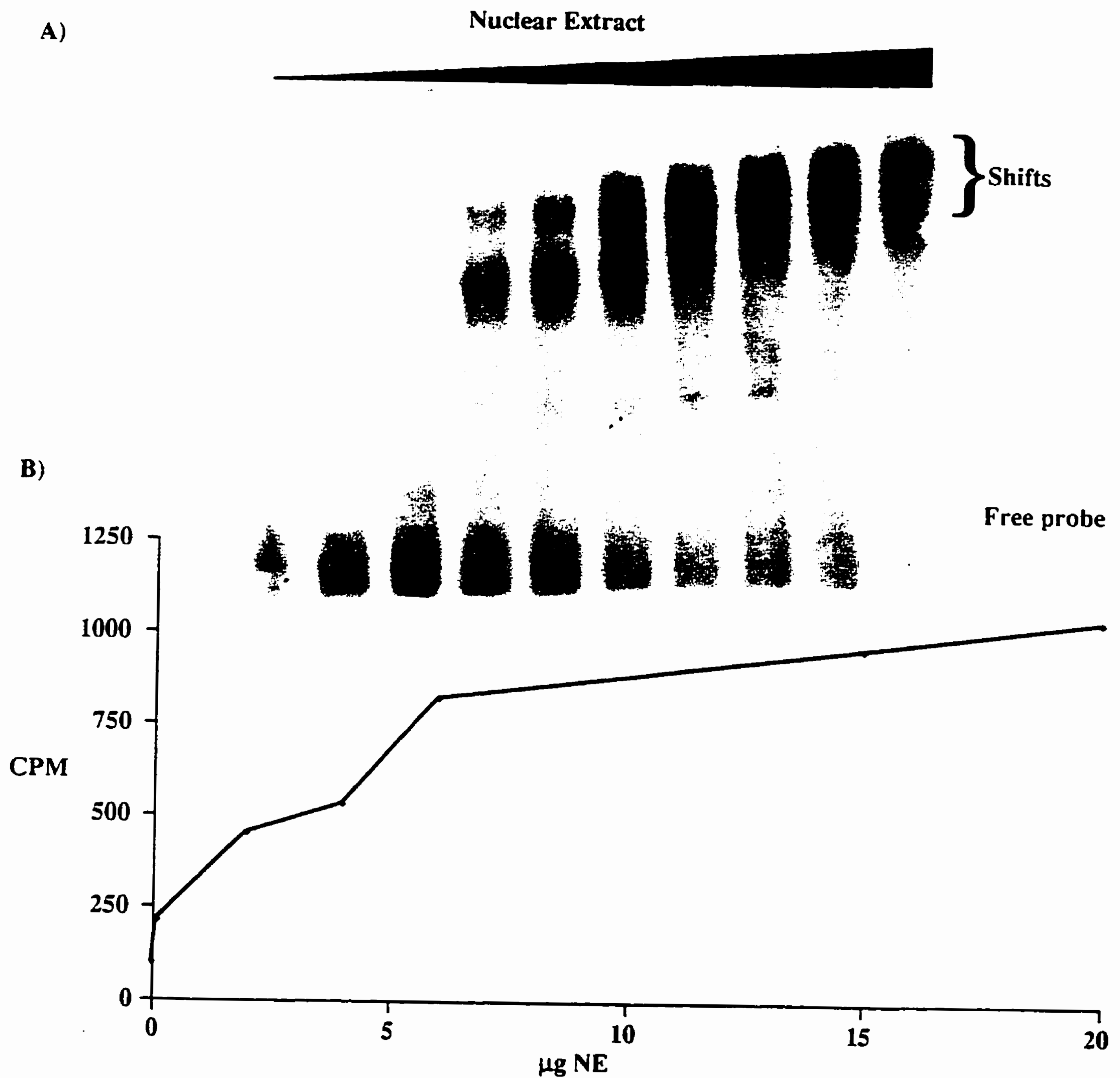
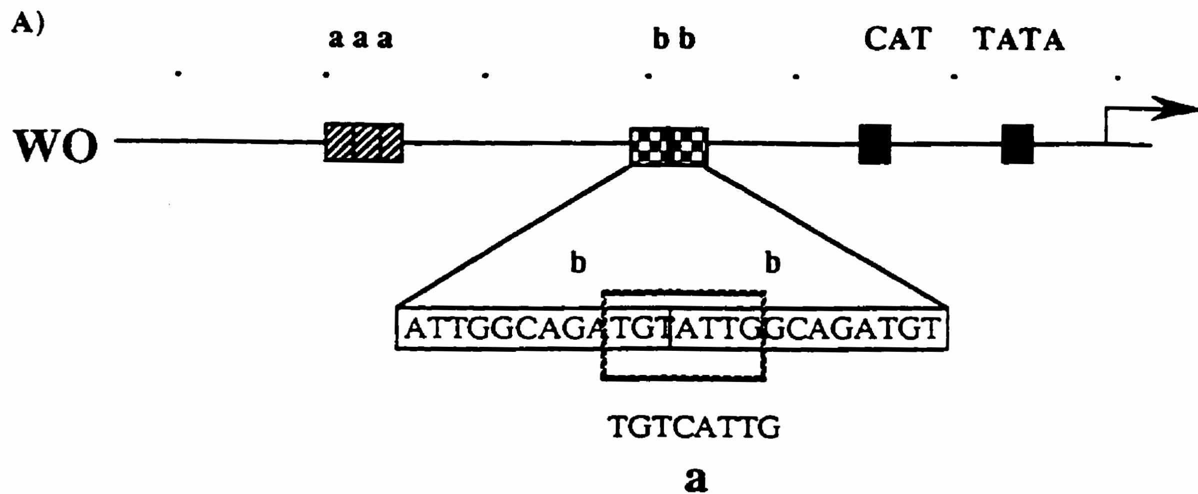


Figure 3-2. Multiple bands appear in a mobility shift assay with a fragment from wolffish promoter. (A) Increasing the concentration of nuclear extract from 0 to 20 µg shows several shifted bands when incubated with the -220 to -90 fragment from wolffish. (B) Graphic display of the radioactivity in shifted bands against amounts of rainbow trout nuclear extract shows a possible sigmoidal binding curve.



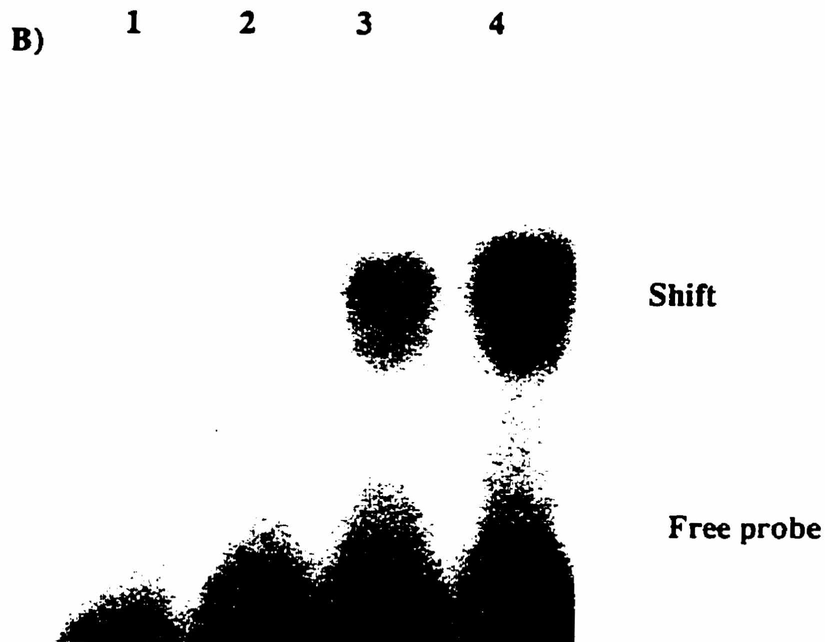


Figure 3-3. Oligomer with two "b" elements is specific and does not bind to putative "a" binding proteins. (A) Three direct repeats of "a" element are located upstream of the "b" element. Furthermore, the two "b" elements have a putative "a" element overlapping in the center. Dots mark 50 bp intervals. (B) Competition between a radiolabeled "b" probe and an oligomer with the three "a" elements. Lane 1: free probe, lane 2: competition with nonlabelled "b" oligomer, lane 3: competition with nonlabelled "a" oligomer, and lane 4; no competition.

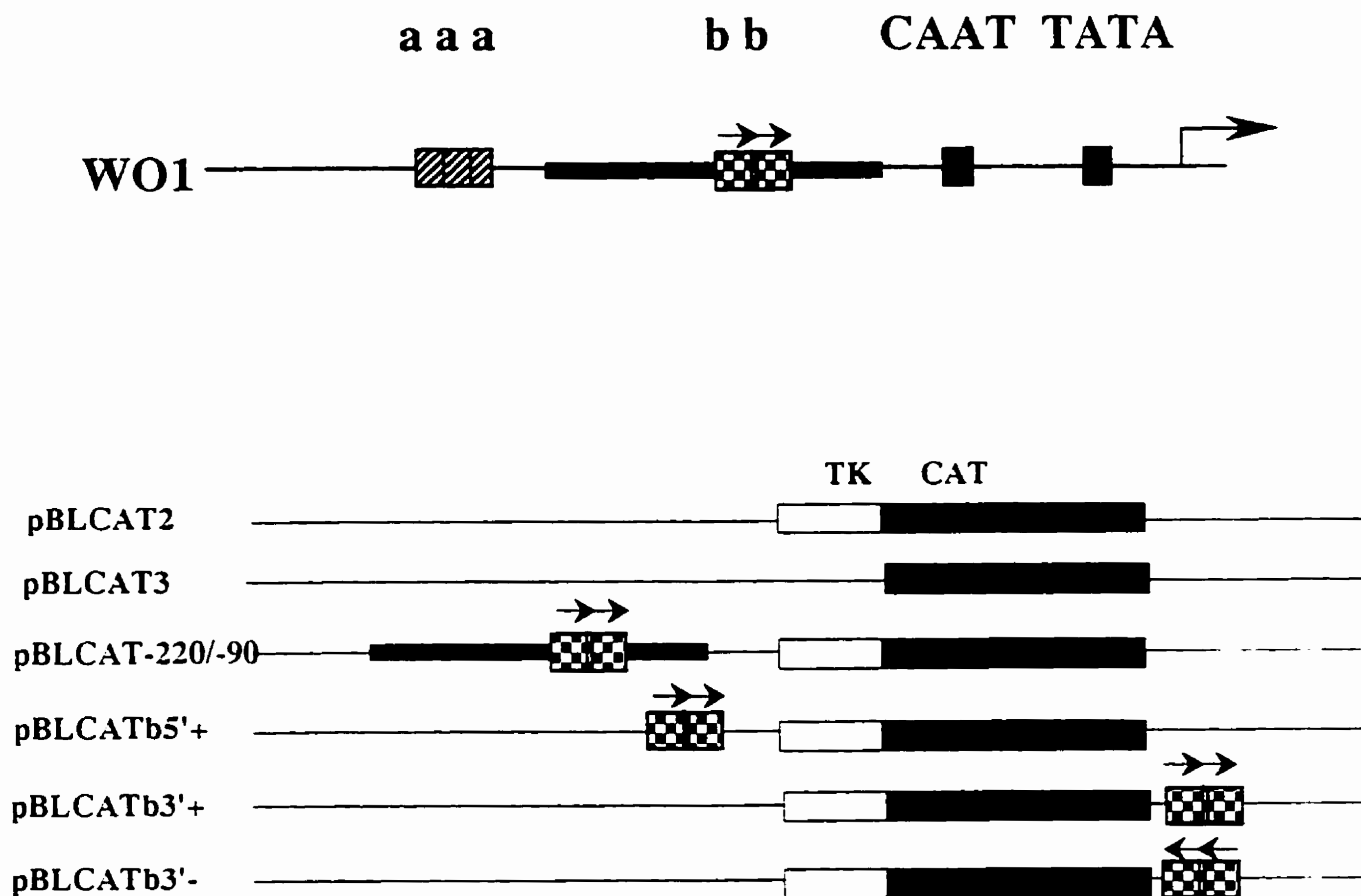
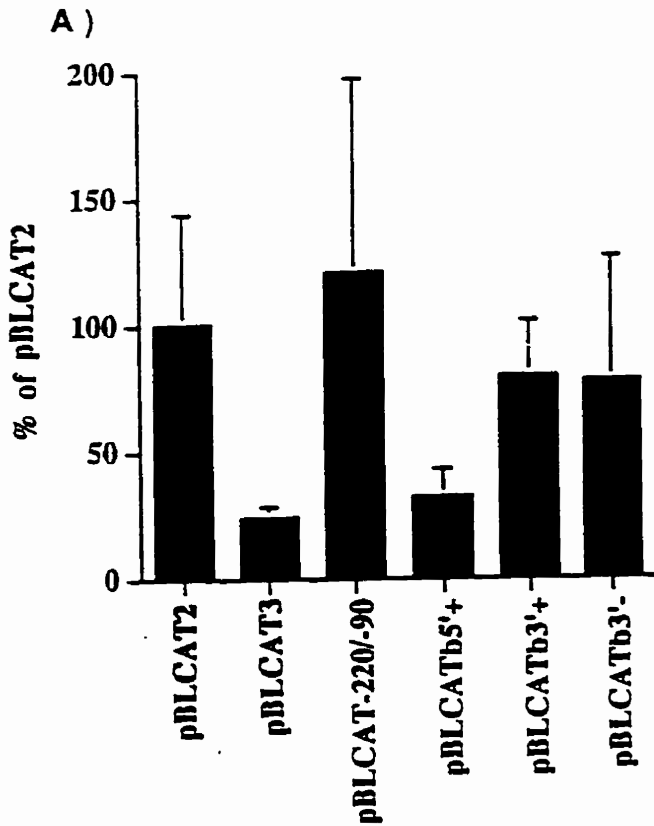


Figure 3-4. DNA constructs for transfections. The construct pBLCAT2 contains a thymidine kinase promoter (TK) from herpes simplex virus, which drives the reporter gene chloramphenicol acetyl transferase (CAT), and is used as positive control. The pBLCAT3 is promoterless CAT plasmid, and is a negative control. Thin lines in the DNA-constructs correspond to plasmid sequences (Luckow & Schütz, 1987). The top figure is a schematic diagram for wolffish 1 (WO1), of the upstream region from the AFP gene. The -220 to -90 fragment from (WO1) (solid line, top), had been subcloned into pWON (Gong & Hew, 1993). This fragment was released with restriction enzymes, and subsequently inserted into a Hind III site of pBLCAT2 to create pBLCAT-220/-90. The "b" oligomer, containing both the "b" elements, was inserted into the Hind III (pBLCATb5'+) or Sma I site in indicated directions (pBLCATb3' + and -).



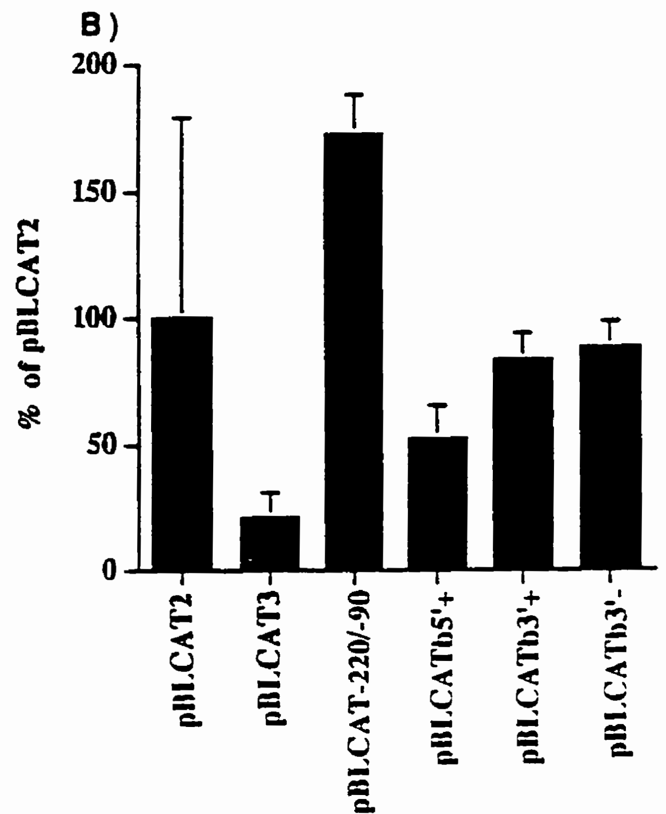


Figure 3-5. A repressor element is located to the two "b" elements. The constructs in Figure 3-4 were transfected into HeLa (A), HepG2 (B). The vertical axis denotes corrected activity of the promoters. The constructs on the horizontal axis correspond to those shown in Figure 3-4.

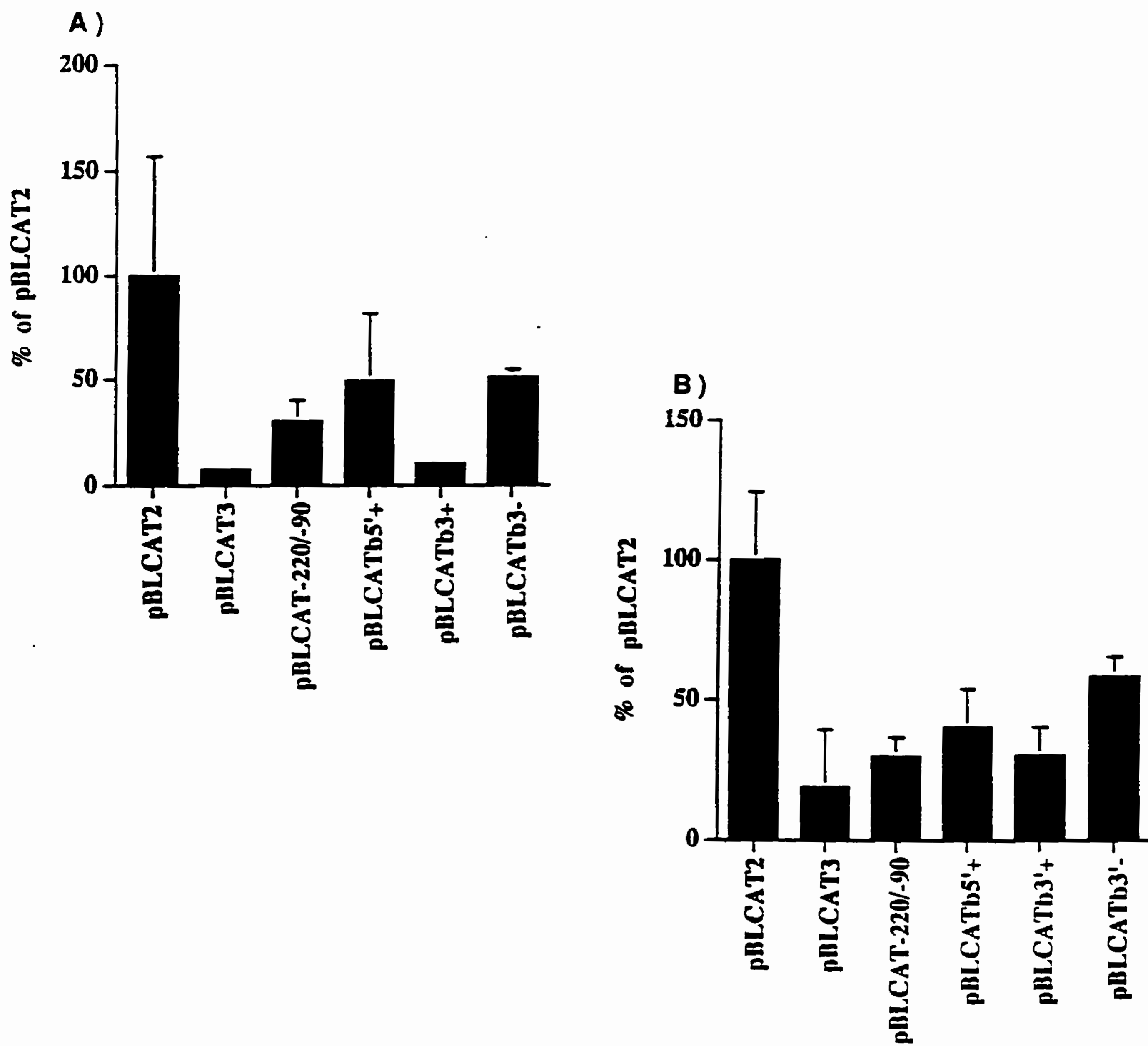


Figure 3-6. The "b" element has strong repressor activity in fish cell lines. RTH-148 (A) and CHH-1 (B) cell lines. The vertical axis denotes corrected activity of the promoters. Constructs shown on the horizontal axis correspond to those shown in Figure 3-4.

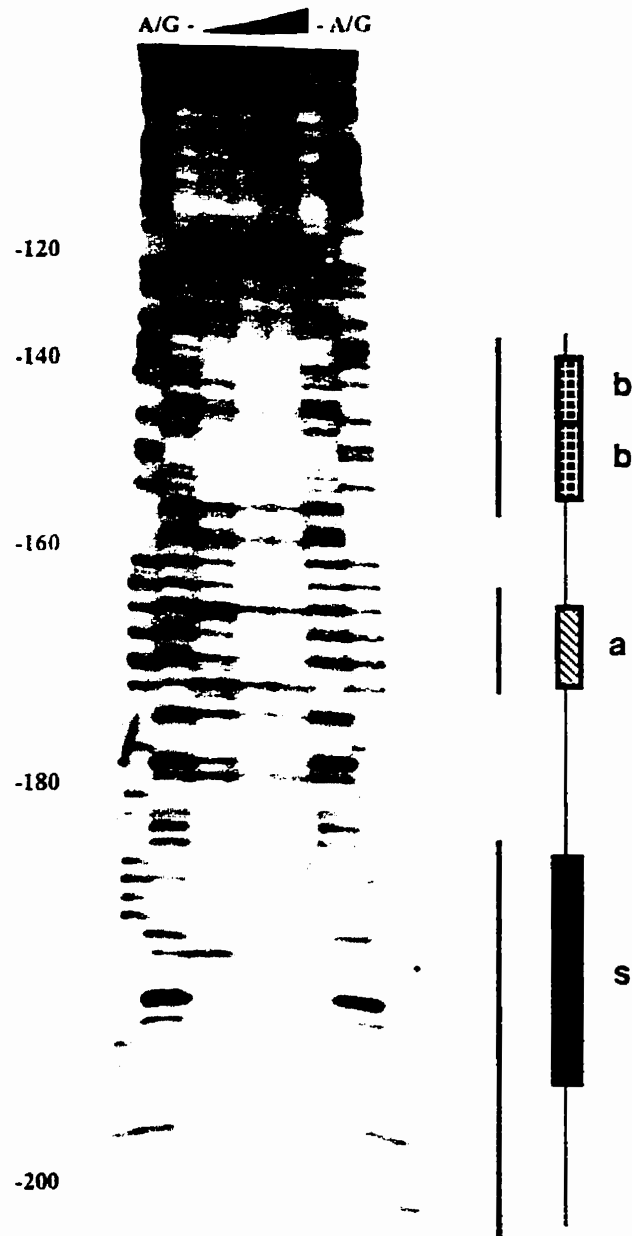


Figure 3-7. The "b" element is protected from DNase I digestion, as are sequences with similarities to Sp1 and the "a" elements. DNA-footprinting of the -220 to -90 fragment, labeled on the Xba I site. A/G, Maxam and Gilbert G+A ladder. –, no nuclear extract, the wedge indicates increasing amounts of nuclear extract from rainbow trout liver. On the right side of the picture, bars high-light protected nucleotides, and boxes indicate where the "b, a and S" elements are located. The utilization of the Xba I restriction site, results in labeling of the complementary strand, from what is shown in Figure 2-1.

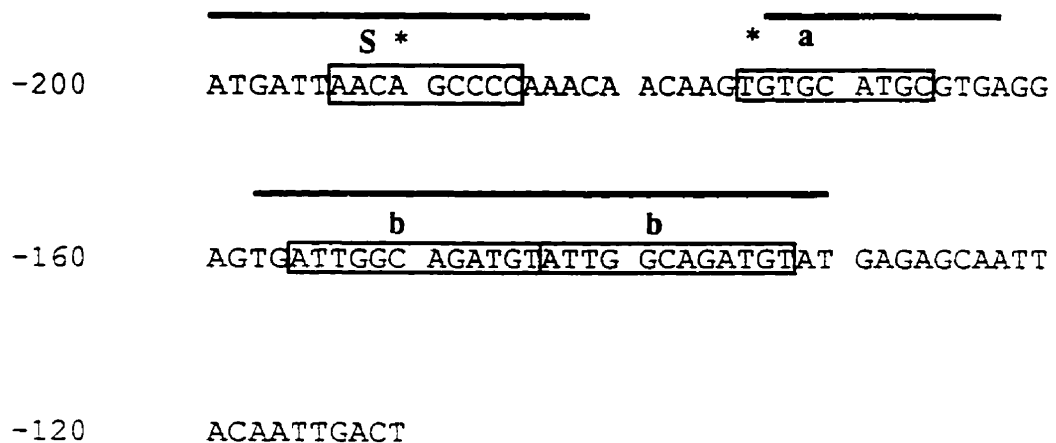


Figure 3-8. Sequence protected from DNase I digestion. Sequences from -200 to -110 of the wolffish AFP promoter, bars indicate protected bases by DNase I digestion, asterisk indicates hypersensitive bases, the elements, "a", "b" and "S" are boxed.

Discussion

In this study several elements are delineated in the -220 to -90 fragment from the AFP gene in wolffish. This fragment binds to proteins from several different nuclear extracts, both of human and fish origin. An element termed "S" might be an Sp1 binding site based on sequence similarity to published consensus sequence. These elements, also termed GC boxes, interact with several proteins of the Sp1 family, which both mediate activation and repression for transcription (Imataka *et al.*, 1992; Kingsley & Winoto, 1992; Hagen *et al.*, 1994).

A second element is observed, which has similar sequences to the three repeated "a" elements, located upstream of the fragment being tested. The salient feature of these experiments is the characterization of the "b" element, which is shown to mediate a transcriptional repression. The variable number and location of the "b" elements from various AFP genes is striking. All Antarctic eel pouts and both ocean pout AFP sequences have only one "b" element (Hew *et al.*, 1988; Du *et al.*, 1992a; Wang *et al.*, 1995b) (see also Chapter 2). Two wolffish promoters sequenced have the "b" element in direct repeat (WO1 and WO2)(Scott *et al.*, 1988) and the third wolffish sequence (WO3) has only one "b" element (Wu and Davies, unpublished results). This results in one "b" element in the common core sequences found in all AFP Type III genes; the extra "b" element is unique to the two wolffish sequences (2nd Chapter). Furthermore these "b" elements are located at various distances from the transcriptional start site, depending on the AFP gene. In OP5 the "b" element is located at -505 bp upstream from the start site, whereas in OP3 it is at -139. The "b" element in WO1 and WO2 is located around -163 whereas the Antarctic eel pout (AEP37 to 39) is at -552. The distance from the "b" element to the upstream "a" elements on the other hand is relatively fixed in all the AFP genes

available, with the exception of AEP37 to 39, which has a 7 bp deletion dispersed between these elements. By increasing the distance of the Sp1 element from the transcriptional start site, the effects of the element are reduced (Sassone-Corsi & Borrelli, 1986). Likewise, increasing the distance between two cooperatively interacting glucocorticoid elements results in a reduced interaction between the transacting factors (Schmid *et al.*, 1989). This raises the possibility that the more distant elements have less effect on transcription than the proximal ones. Increasing the number of elements can result in a synergistic interaction between the corresponding DNA-binding proteins (Carey *et al.*, 1990). This might argue for synchronous activity of the elements in the region beyond the "b" element as they are always grouped together.

It is interesting to note that WO1 and ocean pout promoter constructs show different activity when sequences upstream of the "b" element are deleted. The wolffish construct has lower activity, probably due to the two "b" elements, versus only one in the ocean pout (Gong & Hew, 1993). It is worth noticing that the region surrounding and including the three "a" elements is not perfectly conserved between these species. In contrast to the "b" element which has no published function. the consensus of the "a" element, TGTCATTG, shows similarities to a regulatory element in a gene for ribosomal proteins in *Schizosaccharomyces pombe*. TGTGACTG and differs only in two nucleotides. This element mediates activation for the ribosomal genes in *S. pombe* in conjunction with a nearby direct repeat (Witt *et al.*, 1993; Witt *et al.*, 1995).

The fact that the repressor also acts through a heterologous promoter and independent of distance or orientation, suggests that it either interacts through the general transcription factors or possibly through TATA or CAAT elements, the two latter elements being in common with the tested promoters. To my knowledge no

known DNA-bound repressor has been shown to mediate its action through the CAAT binding proteins. Several repressors (Mot 1, HMG1, HMG2, Dr1 and Topo I) are known to interact directly with TATA-binding-protein (TBP) and TBP associated factors (TAFs), which together make TFIID complex. None of these factors has a distinct DNA binding element. The interaction between these repressors and the TBP or the TAFs is direct and thereby these proteins negate the transcription (Inostroza *et al.*, 1992; Auble & Hahn, 1993; Merino *et al.*, 1993; Ge & Roeder, 1994; Stelzer *et al.*, 1994). The Krüppel (Kr) protein from *Drosophila* provides an example of a DNA-bound *trans*-acting factor with known interaction with the initiation complex. The Kr protein can act either as an activator or a repressor, depending on its cellular concentration (Sauer & Jäckle, 1991). At low concentrations the Kr protein interacts with TFIIB as a monomer, and activates transcription, whereas at higher concentrations the protein forms homodimers and then interacts with TFIIEβ, but not TFIIB, and represses transcription (Sauer *et al.*, 1995).

Chapter 4

PURIFICATION AND CHARACTERIZATION OF A REPRESSOR PROTEIN FOR THE TRANSCRIPTION OF TYPE III AFP GENE

Abstract

To prevent freezing, fish living in cold sea water produce an antifreeze protein that is secreted into their blood. These proteins are usually under circannual control, with maximal levels during the winter months. A repressor element, called "b" element, is present in all Type III AFP promoters; two copies are present in two out of three wolffish sequences. To elucidate the transcriptional control of the antifreeze protein genes in wolffish, a purification of the repressor binding protein was attempted. Through a series of chromatographic steps, three proteins of molecular mass between 34 and 38 kDa were characterized. All three proteins have an identical N-terminal sequence and a similar amino acid composition. These proteins appear in a 60 kDa or larger complex within the nucleus, presumably as a multi-protein complex. This putative repressor, designated as pBB for protein binding to "b" element, appears to be a unique transcription factor.

Introduction

Adaptation to seasonal cycles involves intricate signaling mechanisms. Two important environmental cues for seasonal changes are day length and temperature. The pineal gland, an endocrine organ located in the posterior of the brain, signals the body both the time of the day and the year. The main hormone in the gland is melatonin, which is released with the onset of night. With longer nights during the winter months, more melatonin is released, thus signaling the body both the time of the day and season. Optic nerves direct light signals either directly to the pineal gland, or in mammalian species, via the suprachiasmatic nuclei (Korf, 1994). Furthermore, by *trans*-planting a pineal gland from an old mouse to a younger one, and vice versa, the life expectancy of these animals changes. The old mouse, that has the young pineal gland, delays aging, whereas the young mouse dies prematurely. Therefore the pineal gland synchronizes not only the day and season, but also controls the pace of aging (Lesnikov & Pierpaoli, 1994).

In many animals the photoperiod signals reproduction which is mediated by the pineal gland. By misleading the body with controlled injections of melatonin, simulating the time of the year that reproduction takes place, the oestrous cycle can be controlled. Thus, timed melatonin injections in both sheep (*Ovis spp*) and hamsters (*Mesocricecetus spp*) can be used to induce the reproduction cycle, which usually occurs during mid winter and spring, respectively, independent of seasons (Reiter, 1993). Melatonin appears to bind to two different receptors;; a membrane receptor , that results in elevated amounts of cAMP and an orphan receptor from the steroid receptor family: RZR (Calberg & Wiesenberg, 1995).

Another seasonal cycle is hibernation of during the coldest period of the year. This process is likely mediated by many different hormones, some of them released by

the pituitary gland (Nürnberger, 1995). The suprachiasmatic nuclei become more metabolically active as hibernation progresses, concomitantly to lowering body temperatures and reducing metabolic activity of the animal (Kilduff *et al.*, 1993). These animals can be grouped into two different categories, each controlled by a different set of cues for entering hibernation. The first group, *facultative hibernators*, depends on environmental signals such as cold temperatures, short photoperiods and restricted food supply to induce hibernation. The hamster is a representative of this group. The second group, *obligate hibernators*, depend more on intrinsic circannual changes in their physiology rather than on environmental cues for hibernation. The ground squirrels (*Citellus or Spermophilus spp*) represent this group (Kilduff *et al.*, 1993).

Many additional biochemical changes are observed in animals during the course of the year. Several neuropeptides in the pituitary gland in rats (*Rattus norvegicus*) have circannual cycles, independent of photoperiodic changes or any other external signals (Bissette *et al.*, 1995). Somatolactin, a pituitary hormone in fish, with an unknown function, has strong circannual expression. In rainbow trout its expression correlates with ambient water temperature, high in warm and low in cold water. The hormone was out of phase with the photoperiod, peaking one month after the longest day (Rand-Weaver *et al.*, 1995).

Atypical behavior can be prompted by the seasonal cycle as well. In humans, some mood disorders are known to be correlate with seasonal changes. Mania, unipolar and bipolar depression have been suggested to peak in winter and spring (Szabo & Blanche, 1995). In the Belgian population violent suicides, peak in the spring and fall as opposed to non-violent suicides, which are not affected by the seasonal cycle (Maes *et al.*, 1995). Sudden infant death syndrome in newborn children, is more prominent during the winter than summer months (Haglund *et al.*, 1995).

Fish living in the polar oceans are protected from freezing by antifreeze proteins, which are secreted into the plasma. These proteins are typically synthesized during the winter months, concomitantly when the temperature is coldest in the sea water (Davies *et al.*, 1988). The seasonal control of the winter flounder AFP genes is under the negative control of growth hormone, which is secreted from the pituitary gland. During summer months the hormone suppresses the synthesis of AFP, whereas its absence in the winter months allows the synthesis of the protein to resume (Fletcher, 1979; Hew & Fletcher, 1979). In cod, ambient sea water temperature affects the synthezes of the Type II AFP (Fletcher *et al.*, 1987). The ocean pout, which synthesis the Type III AFP, has the largest difference between summer and winter concentrations of any AFP measured (Davies *et al.*, 1988), but the gene control is poorly understood.

The genes encoding Type III antifreeze proteins have a strong repressor element, referred to as the "b" element, which is located upstream of the promoter (Chapters 2 annd 3). To elucidate the transcriptional control of the Type III antifreeze proteins with the aim of explaining the circannual cycle, the putative repressor protein was purified and characterized. Three individual proteins were purified, ranging in size from 34 kDa to 38.5 kDa. All proteins have an identical N-terminal end and similar amino acid compositions. The protein, named protein binding to "b" element (pBB), appears to be a novel transcription factor.

Materials and methods

Preparation of pBLXb Construct

The pWON construct containing the-220 to -90 fragment (positions 2300 to 2816, in Figure 2-1) was digested with Xba I restriction enzyme, and the 130 bp fragment was isolated by electrophoresis on a 1% agarose gel, electroeluted, precipitated and dissolved in water. Bluescript (KS) vector (Promega) was digested with Xba I enzyme, phosphatased and the reaction mixture was centrifuged through Probind filters. The 130 bp fragment was ligated into the vector and transformed into the DH5α strain of E. coli and grown up on LB agarose plate medium with ampicillin. The presence and orientation of the insert was verified by DNA sequencing.

Preparation of (Xba I)$_n$ DNA

Bacteria with the pBLXba I construct were grown up in 11 liters of LB medium in the presence of ampicillin by overnight incubation at 37°C. The plasmid was purified as previously described with minor modifications (Sambrook *et al.*, 1989). The cells were centrifuged at 6000 rpm for 5 minutes using a GSA rotor in a Sorvall centrifuge and resupsended in 1x STE buffer made of 10 mM Tris buffer pH 8.0, 1 mM EDTA, 100 mM NaCl, and centrifuged as before. Bacteria were resuspended in lysozyme solution (50 mM glucose, 25 mM Tris buffer, 10 mM EDTA and 2 mg/ml lysozyme) and incubated at room temperature for 30 minutes. The bacteria were then lysed with 0.2 M of NaOH and 1% (w/v) SDS followed by a 20 minute incubation at 0°C. The solution was neutralized by the addition of 3 M potassium and 5 M acetic acid and incubated for 30 minutes on ice, followed by centrifugation as described above. The supernatant was transferred to a clean tube and the DNA was precipitated by adding 0.6 volumes of isopropanol, followed by 20 minutes incubation at room temperature. The DNA was

pelleted by centrifuging the bottles for 30 minutes at 11000 rpm in a GSA rotor in a Sorvall centrifuge and washed once with 70% (v/v) ethanol and centrifuged again. The DNA was separated by CsCl gradient as described (Sambrook *et al.*, 1989) with minor modifications. The plasmid was dissolved in 120 ml of a solution containing 1 g/ml CsCl and 1.3 mg/ml of ethidium bromide. The mixture was poured into quick seal centrifuge tubes and centrifuged for 36 hours at 35,000 rpm in Ti50 rotor for in Beckman centrifuge at room temperature. The plasmid was isolated using a syringe and the solution was extracted several times with 1-butanol until all the ethidium bromide had been removed. The DNA was precipitated with ethanol and washed with 70% (v/v) ethanol and dried in a Speed Vac Centrifuge. This resulted in 90 mg of plasmid. The insert was then excised out of the plasmid with BamH I restriction enzyme and the fragment was separated from the plasmid with electrophoresis on 1% agarose gel. The area containing the fragment was cut out of the gel, electroeluted and purified as described in the Chapter 3. The fragment was ligated into 4mer on the average, as observed on a ethidium bromide stained agarose gel. Around 1 ng of the isolated DNA was radiolabeled and mixed with the remaining DNA solution. This radioactivity was used to monitor coupling efficiency of the DNA to the Sepharose (see preparations of (Xba I)$_n$ DNA).

Cross-linking Proteins to DNA

A bromo-labeled oligomer and an oligomer complementary in sequence were synthesized (Biotech Service Center) and hybridized as described (Chapter 3).

```
AGTGAGGAGZGATZGGCAGAZGTATGAGA
  ACTCCTCACTAACCGTCTACATACTCTCT
```

The "Z" denotes equal ratio of the 5-bromo-2'-deoxyuridine nucleotide and thymidine nucleotide. The oligomer was end labeled by kinase treatment and purified from non-incorporated nucleotides as described (Chapter 3). A total of 15 µg nuclear extract, was preincubated on ice with 1 µg of poly[dI·dC]·poly[dI·dC] in a total volume of 19 µl in MS buffer (Chapter 3) for 30 minutes. Into this mixture was added 1 µl of radiolabelled probe and the mixture was incubated an for additional 30 minutes on ice. The Eppendorf tubes were opened and irradiated with UV light for 2 minutes on an ultraviolet transilluminator (La Jolla Scientific Co. Inc). To each tube was then added 20 µl of 2x loading buffer. (20% glycerol, 20% SDS (w/v), 25 µg/ml Bromophenol blue and 0.72 M mercaptoethanol). Samples were boiled for 2 minutes, cooled on ice and run on a 12.5% acrylamide gel. During the handling of the oligomers, they were shielded from light by wrapping the vials in aluminium foil.

Mobility Shift Experiments

Mobility shift experiments were done as described in the Chapter 3. with minor modifications. When testing fractions from the affinity columns. poly[dI·dC]·poly[dI·dC] was omitted and samples were adjusted to 10 mM KCl. by measuring the conductivity of the fractions with a conductivity meter (VWR Scientific). An oligomer encompassing the "S" element was synthesized (Biotech Service Center), and annealed to form the duplex::

```
   ATTAACAGCCCAAACAAGA
CTTAATTGTCGGGTTTGTT
```

This oligomer was radiolabelled and purified as described in the Chapter 3. Pre-annealed, double stranded oligomer for the Ap3 element was purchased from Promega.

Construction of (Xba I)$_n$ Affinity Column

The coupling of the DNA to the matrix was done as previously described with minor modifications (Ausubel *et al.*, 1992). Around 20 ml of Sepharose 4B slurry (Pharmacia) was placed in a Buchner funnel attached to an Erlenmeyer flask that was connected to a water aspirator. The slurry was washed with 200 ml of water, all excess water was drained off and the Sepharose was resuspended in an equal volume of 0.2 M Na_2CO_3 and 640 µl of 62.5% (w/v) cyanogen bromide was added drop wise and stirred for 5 minutes. The Sepharose was then filtered as before and washed with 100 ml of 1 mM HCl. The slurry was resuspended in 1 mM HCl and around 2 ml of the Sepharose was used for DNA-coupling. The DNA was dissolved in 100 µl of 0.1 M $NaHCO_3$, pH 8.3 and mixed with the activated Sepharose. The slurry was put on a rotator and slowly inverted for 16 hours at 4°C. Uncoupled ligand was washed out with 100 ml of 0.1 M $NaHCO_3$, pH 8.3 and the remaining active groups were blocked by soaking the slurry in 0.1 Tris buffer, pH 8.0 for two hours at 4°C. The matrix was washed with 0.1 M acetic acid, pH 4.0 and 0.5 M NaCl, and then with 0.1 M Tris buffer, pH 8.0 and 0.5 M NaCl. This cycle was repeated three times. The (Xba I)$_n$ affinity column was packed into a column and stored in 0.1 M Tris buffer, pH 8.0, 5 M NaCl and 0.01 % NaN_3.

Oligomer "b" Affinity Column

An oligomer with a biotin group attached was synthesized (Biotech Service Center) and hybridized to a complementary strand

GAGTGATTGGCAGATGTATGAGAGAT-Biotin
ACTCTCACTAACCGTCTACATACTCT

in water as described in the Chapter 3. The complementary strand was in 10% excess to the biotin-labeled strand during hybridization. A total of 23.4 nmols of double stranded

(ds) biotin-labeled DNA was used for coupling to the gel. The dsDNA was coupled to 0.5 ml of streptavidin-labeled agarose (Sigma). The streptavidin was attached to the agarose via a 7 atom spacer. The coupling was done by mixing 430 µl of dsDNA to 500 µl of the streptavidin-agarose with constant rotation at 4°C for 16 hours. The final composition of the coupling mixture was 5 mM phosphate buffer, pH 7.2, 75 mM NaCl and 0.1 % NaN_3. Non-coupled DNA was eluted with 10 mM Tris buffer pH 7.8 and 1 mM EDTA. An optical density at 280 nm was used to determine the coupling efficiency. This resulted in 45 nmols of dsDNA/ml gel. Columns were stored at 4°C with NaN_3.

Electrophoresis of Proteins

Denaturing electrophoresis for proteins was performed as described with minor changes (Ausubel *et al.*, 1992). The separating gel was 12.5% acrylamide. The gel was prepared by mixing together 6.25 ml of 30% acrylamide:bis (30:0.8), 3.75 ml of 4x running buffer, consisting of 1.5 M Tris buffer, 0.4% SDS pH 8.8 and polymerization was started by adding 50 µl of 10% ammonium persulfate (ICN) and 10 µl of TEMED (BioRad). Both the acrylamide and running buffer were filtered through a 0.45 µm nitrocellulose membrane prior to gel casting. The separating gel mixture was poured into a gel sandwich separated by 0:75 mm spacers. Isobutyl alcohol was loaded on the top of the solution to form a smooth surface. After polymerization, the isobutyl alcohol was washed off, and the stacking gel was loaded on top. It was composed of 0.65 ml of 30% acrylamide:bis solution and 1.25 ml of 4x stacking gel buffer, consisting of 0.5 M Tris buffer and 0.4% SDS, pH 6.8. Polymerization was initiated by adding 25 µl of 10% ammonium persulfate and 5 µl of TEMED. As before, the acrylamide and the stacking gel buffer were filtered through 0.45 µm nitrocellulose membrane. Samples were mixed with 2x loading buffer, boiled for 2 minutes, cooled on ice and loaded onto the gel. Gels were run with 10V/cm until the dye ran off the gel.

Gel Staining

Staining of acrylamide gels was done as described with minor modifications (Ausubel *et al.*, 1992). After electrophoresis, the gel was fixed in 50% methanol (v/v) and 10% (v/v) acetic acid for 2 hours with gentle shaking. The gel was then stained with 0.05% Coomassie brilliant blue R-250 (BioRad) in the fixer solution by gentle shaking overnight at room temperature. The staining solution was poured off and the gel was rinsed several times with destaining solution, consisting of 16.5% methanol and 5% acetic acid. Fresh destaining solution was added and the gel was shaken until a clear background was obtained. The gel was stored in 7% (v/v) acetic acid.

N-terminal Sequencing and Amino Acid Composition Analysis

PVDF membranes with blotted proteins, were submitted to Biotech Service Center, which performed the N-terminal sequencing and amino acid composition analysis. The amino acid composition was done with Waters PICO-TAG system after the protein had been hydrolyzed in 6M HCl and 1% phenol for 24 hours at 110°C. N-terminal sequencing was done by using a Porton gas phase microsequencer, model 2090 with online phenylthiohydantoin analysis.

Bio-Gel P-60 Chromatography

The Bio-Gel P-60 (BioRad) matrix was hydrated in elution buffer (20 mM HEPES buffer pH 8.0, 10 mM $MgCl_2$, 16% glycerol and 100 mM KCl and 0.1% Nonidet P-40) overnight at room temperature. The slurry was degassed for 30 minutes and 2 volumes of degassed elution buffer was added. The gel was gently stirred and allowed to settle, then the supernatant was removed to eliminate fine particles. The removal of fine particles was repeated four times. The remaining slurry was poured into a column of 1.5 cm width and 50 cm length. After the gel had settled, the column was

washed with 200 ml of elution buffer with a flow rate of around 15 cm/hr at 4°C. Around 50 mg of nuclear extract in a 2 ml volume was loaded onto the column and eluted with the elution buffer. The eluate from the column was collected into 1.5 ml fractions, and tested in a mobility shift assay using the 130 bp promoter fragment from the AFP in wolffish as a probe. Positive samples were pooled and concentrated.

Concentration of Samples from the Chromatography Columns

Positive fractions from the columns were pooled and BSA was added to a final concentration of 0.5 mg/ml. The samples were then concentrated by centrifugation with Ultra filtration Devices-series (PGC Scientific) using a membrane with a 20 kDa molecular weight cut off. The samples were centrifuged at 4°C in a GSA rotor in a Sorvall centrifuge at 2,500xg.

Chromatography with $(Xba\ I)_n$ Affinity Column

The $(Xba\ I)_n$ chromatography affinity column was equilibrated with binding buffer consisting of 20 mM HEPES buffer pH 7.8, 1.25 mM MgCl2, 16% glycerol, 100 mM KCl, 0.5 mg/ml BSA , 0.1 mM DTT and 0.01% Nonidet P-40 with a flow rate of 25 cm/hr. To a 1.5 ml concentrated fraction from the Bio-Gel P-60 column was added 75 µg of poly[dI·dC]·poly[dI·dC]. This mixture was circulated through the $(Xba\ I)_n$ column for three hours at 4°C. The column was then washed with 20 ml of binding buffer excluding the BSA and poly[dI·dC]·poly[dI·dC]. A linear salt gradient, from 100 mM to 800 mM concentration, was established with KCl, excluding the BSA and poly[dI·dC]·poly[dI·dC] to elute bound proteins from the column. Samples were collected in 500 µl fractions and conductance was measured with a conductivity meter (VWR Scientific). DNA-binding activity was measured in the mobility shift assay, using the 130 bp fragment as a probe. Positive fractions were pooled and concentrated.

Chromatography with Oligomer "b" Affinity Column

The column was pre washed with 10 ml of binding buffer, consisting of 10 mM HEPES buffer pH 7.8, 1.25 mM $MgCl_2$, 16% glycerol, 0.1 mM DTT, 0.01 % Nonidet P-40, 50 mM KCl and 0.5 mg/ml BSA. The buffer of the pooled and concentrated fractions from the $(Xba\ I)_n$ column was changed to the above binding buffer with Ultra filtration Devices. The final volume was 2 ml, which was mixed with the resin and gently shaken for 3 hours at 4°C. The matrix was washed with 5 ml of the binding buffer, excluding the BSA. Bound protein was washed off the column with the binding buffer including 800 mM KCl, but excluding the BSA.

Electroblotting and Staining of Membrane

After electrophoresis the gel was placed on a PVDV membrane (BioRad) supported by filter papers and sponges, and sealed inside a sandwich for the electroblotting minigel apparatus (BioRad). All components were immersed into blotting buffer (39 mM glycine, 48 mM Tris base, 0.037% SDS and 20% (v/v) methanol, pH 8.3). The membrane was pre wetted in methanol for 3 seconds prior to use. The blotting was done for 30 minutes at 4°C at 85 volts. The membrane was stained by immersing the membrane into 40% (v/v) methanol and 0.025% (w/v) Coomassie blue R-250 for 30 minutes at room temperature. Destaining was done by placing the membrane in a Vacuum-Blot apparatus (LKB) and 40% (v/v) methanol was slowly filtered through the membrane until the bands were clearly visible.

Large Scale Preparation of Nuclei

The preparation of nuclei was done essentially as described in the Chapter 3, with minor modifications as previously published (Sierra, 1990). The minced liver was mixed

with the resuspension buffer in the ratio of 3 g:10 ml (instead of 1 g:10 ml) and the buffer contained 2.4 M sucrose instead of 2.0 M.

Results

Cross-Linking the "b" Element to the Putative Repressor

An estimation of the size of the putative repressor protein is a prerequisite for its purification. An oligomer was synthesized which had equal ratios of bromo-labeled uridine and thymidine nucleotides at three positions within and around the repeated element and annealed to a complementary strand (Figure 4-1A). The double stranded oligomers were therefore a mixture of eight different sequences, ranging from an oligomer with no substitutions to oligomers with three incorporated bromo-labeled nucleotides. This design of heterogeneous oligomers was to minimize the effects of possible steric hindrance due to the substituted nucleotide. Incorporation of a modified nucleotide which would prevent protein binding, would be compensated with a population of four other different oligomers, which would not utilize that particular site. The bromo-labeled oligonucleotide was radiolabelled and mixed with crude nuclear extract from rainbow trout livers under conditions identical for mobility shift binding. except that all steps were carried out in a darkroom under a red safety light. After UV exposure, the mixture was run on a SDS-PAGE gel and exposed to X-ray film. One band was observed in the nuclear extract lane, with a molecular weight around 36 kDa (Figure 4-1B).

Purification of the Repressor Protein

The rationale of the purification process for the repressor protein was based on the assumption that several proteins might be in a complex or an individual protein might be interacting with the "b" element (Figure 4-2). Nuclear extract was isolated from rainbow trout liver and used as a source to isolate the protein. The extract was then precipitated with ammonium sulfate and dialyzed against the buffer used for the mobility shift

experiment. The same buffer was also used for elution. The extract was first loaded on to a Bio-Gel P-60 column. Fractions collected from the column were tested in the mobility shift assay using the 130 bp fragment as a probe. Positive fractions eluted in the void volume (Figure 4- 3), which would suggest that the size of the protein or the complex is 60 kDa or larger. Positive fractions were pooled and concentrated.

The collected fractions were mixed with competitor DNA to eliminate non-specific binding and loaded on to a $(Xba\ I)_n$ affinity column in the buffer used for the mobility shift assay. The sample was repeatedly circulated through the column to facilitate binding to the DNA. After washing, the column was eluted with a linear salt gradient. Each fraction, 0.5 ml in size, was tested in the mobility shift assay using the same probe as before after adjusting the salt concentration to 50 to 100 mM KCl. Positive fractions eluted at the beginning of the salt gradient (Figure 4-4). The 130 bp fragment, which was used in the $(Xba\ I)_n$ column has several sites protected in the DNase footprinting assay (Chapter 3). Oligomers corresponding to these sites were synthesized and tested in the mobility shift experiment using positive pooled fractions from the Bio-Gel P-60 column. The oligomers for individual "S, a and b" elements were used in the mobility shift assay, testing the pooled fractions from the $(Xba\ I)_n$ column. The oligomer for the "a" element corresponds to all three "a" elements, upstream of the 130 bp fragment, the same fragment used in the $(Xba\ I)_n$ affinity column. All the oligomers gave a shifted band when tested with the pooled fraction from the $(Xba\ I)_n$ column, in sharp contrast with the oligomer for the Ap3 element, which was only positive in the crude nuclear extract (Figure 4-5). This further confirms multiple protein interaction with the 130 bp fragment and an "a"-like element(s) in this region.

The positive fraction from the $(Xba\ I)_n$ column was mixed with the resin containing the attached "b" element and gently shaken at zero temperature. The matrix was washed and bound protein eluted with high salt solution. Fractions were tested with

the same DNA fragment as before and a shift was observed in the high salt eluate (Figure 4-6). Analysis on a SDS-PAGE gel, with samples from the positive fraction showed three distinct bands when stained with Coomassie Blue (Figure 4-7). These bands were of 34, 36 and 38.5 kDa in size, which is in a good agreement with results obtained with the "b" element cross-linked to its putative repressor.

Chemical Analysis of Purified Proteins

The three proteins were separated on a SDS-PAGE gel and transferred to a PVDF membrane and each band was excised separately. Amino acid analysis and N-terminal sequencing was performed on each band. All three bands show similar amino acid compositions with the highest amount being for the glutamine and glutamic acid combined (Figure 4-8). The presence of glycine in the transfer buffer, leads to artificially high levels in the amino acid composition data. The amino acid from the available N-terminus are identical for all three proteins. The first site contains one of three amino acids: glutamic acid, alanine or glycine, followed by glutamine-proline-glutamic acid-alanine residues. The glycine residue in the first position could be due to its presence in the transfer buffer. Further sequencing was not feasible due to the small quantity of sample.

A)

Oligomer "b" :   AGTGAGGAGZGATZGGCAGAZGTATGAGA
    ACTCCTCACTAACCGTCTACATACTCTCT

B)   **1 2**

Figure 4-1. The "b" element cross-links to a protein of around 36 kDa in size. A) An oligomer spanning one "b" element was synthesized with bromo-labeled uridine or thymidine in equal ratios (Z) and annealed to a complementary strand. B) Radiolabelled bromo-substituted oligomer, either without any nuclear extract (Lane 1) or cross-linked in the presence of nuclear extract, interacts with a protein of around 36 kDa in size (Lane 2).

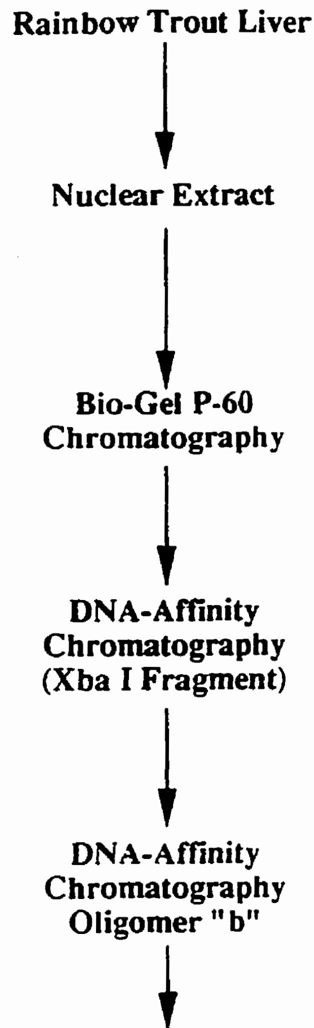


Figure 4-2. Schematic outline for purification of the putative repressor protein for the "b" element. The first step involves the isolation of nuclei from rainbow trout liver, followed by a chromatography through a size exclusion matrix. This was followed with two different DNA-affinity chromatography steps, the first using a 130 bp fragment, encompassing the "b" element and surrounding region from the wolffish AFP promoter. The second affinity step was done with the "b" element alone attached to a supporting matrix.

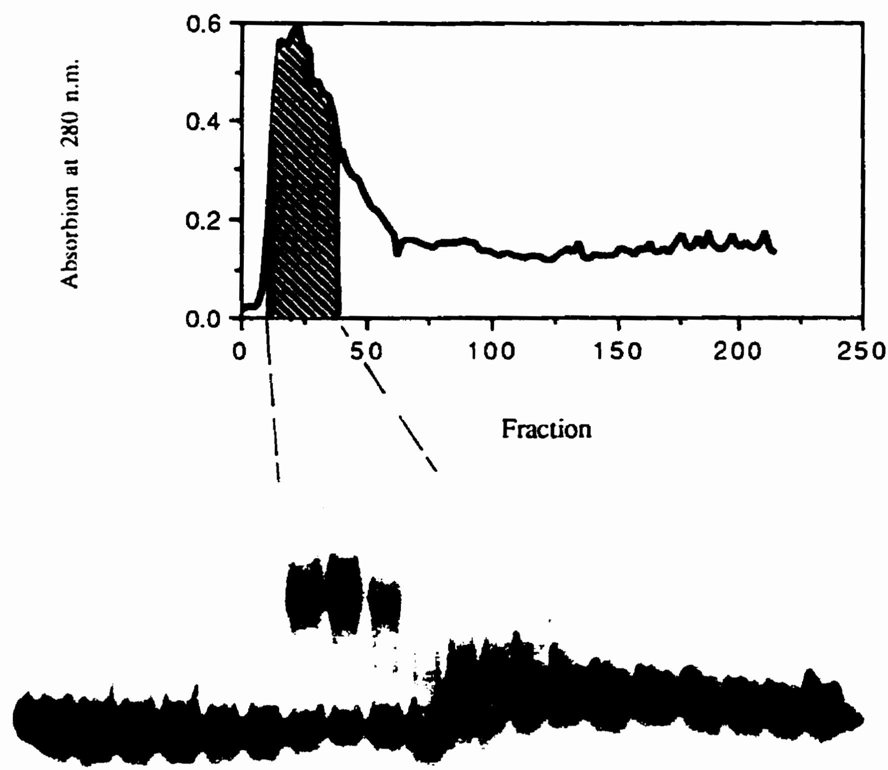


Figure 4-3. Protein components of around 60 kDa or larger bind to the upstream region of the AFP promoter in wolffish. The elution profile from Bio-Gel P-60 column for the extract from rainbow trout liver, were collected into 1.5 ml fractions and measured at 280 nm (top). Fractions positive for DNA-binding were identified with mobility shift using the 130 bp fragment from the wolffish AFP promoter as a probe (lower). Positive fractions (shaded area) elute in the void volume of the column suggesting that the size of the component is 60 kDa or larger.

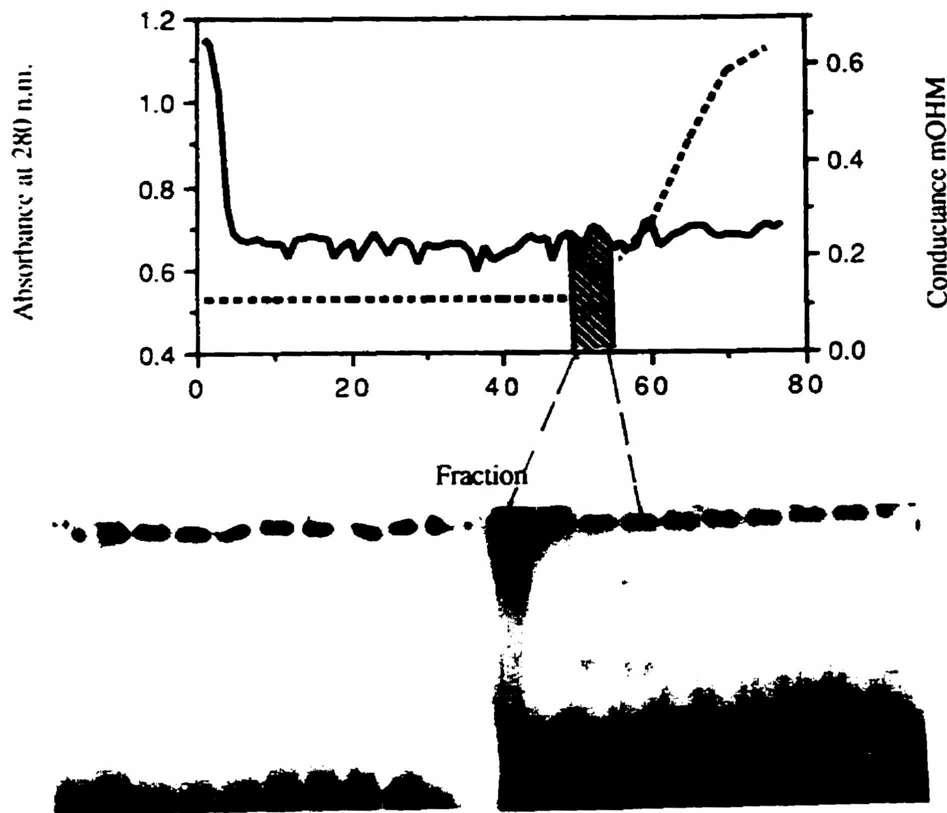


Figure 4-4. Chromatography with the (Xba I)$_n$ DNA-affinity column. The elution profile for the positive fractions from the Bio-gel P-60 column, passed through a DNA-affinity column. The column consisted of concatenated 130 bp fragment from the wolffish AFP promoter, which was attached to supporting matrix. Samples were collected in 0.5 ml fractions and measured at 280 nm (top). Bound protein was eluted with increasing concentration of KCl. Positive fractions were identified with mobility shift assay (shaded area), using the 130 bp fragment as a probe with the salt concentration adjusted 50 to 100 mM. The dashed line corresponds to the conductance, and the solid line denotes the optical density read at 280 nm.

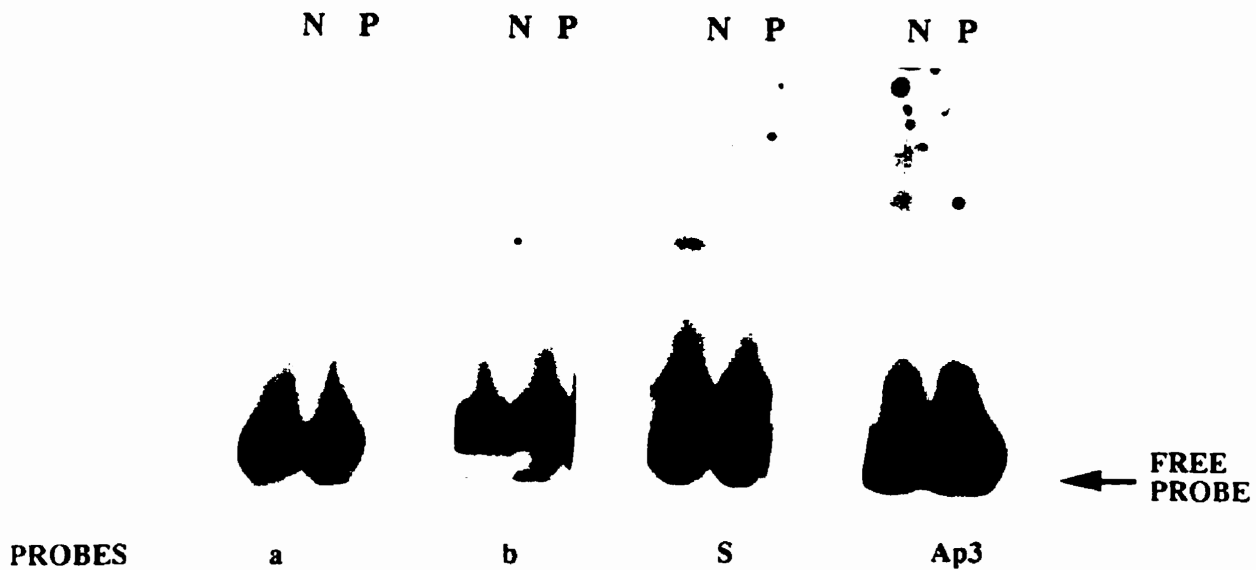


Figure 4-5. Positive fractions from the (Xba I)$_n$ column interact with the "a, b and S" elements. The positive fractions from the (Xba I)$_n$ column were pooled concentrated. The pooled fraction was tested for DNA-binding activity (lanes P) with oligomers containing either the "a, S, b" or Ap3 elements and crude nuclear extract from rainbow trout livers (lanes N).

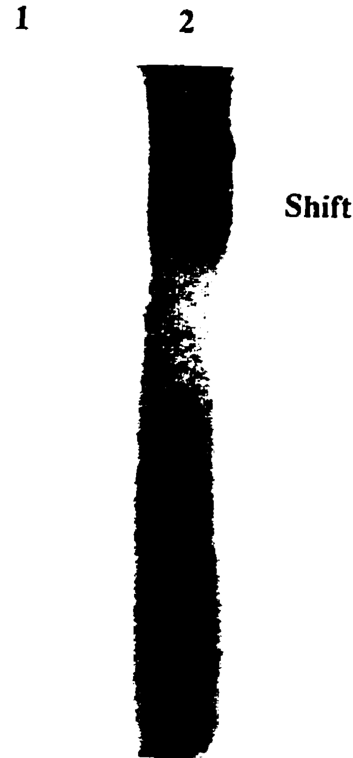


Figure 4-6. Positive fraction from the "b" affinity column interacts with the "b" element in mobility shift assay. DNA-affinity chromatography with the "b" element affixed to agarose matrix was used on positive fractions from the (Xba I)$_n$ column. The fractions from (Xba I)$_n$ column were pooled, concentrated and dialyzed against buffer suitable for DNA binding and loaded on to a DNA-affinity column with the "b" element attached to a supporting matrix. Fractions were tested in the mobility shift assay using the 130 bp fragment as a probe. Lane 1; free probe, lane 2; high salt elute fraction.

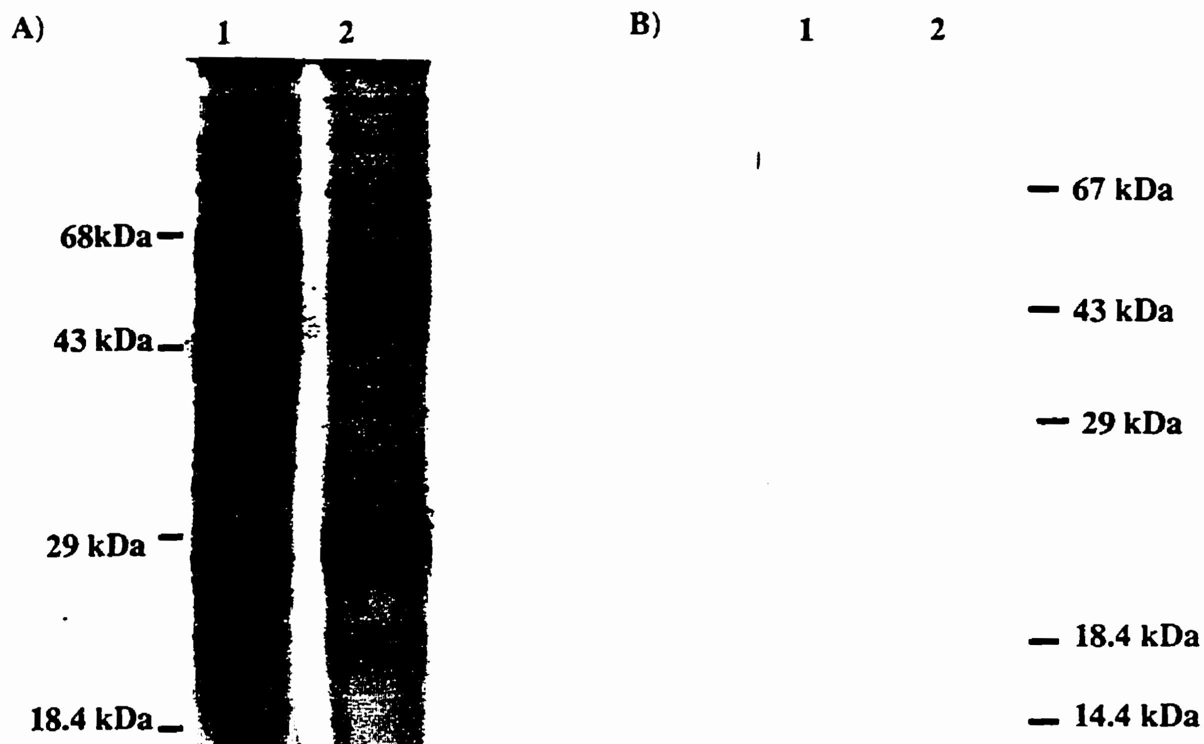


Figure 4-7. Positive fraction from the "b" affinity column shows three distinct bands. Electrophoresis of samples from the chromatography columns. A) Lane 1; Crude nuclear extract, lane 2; pooled positive samples from the Bio-Gel P-60 column. B) pooled positive samples from DNA-affinity chromatography, lane 1; $(XbaI)_n$ column and lane 2; after "b" affinity column. Three distinct bands are apparent after gel staining, 34 , 36 and 38.5 kDa in size.

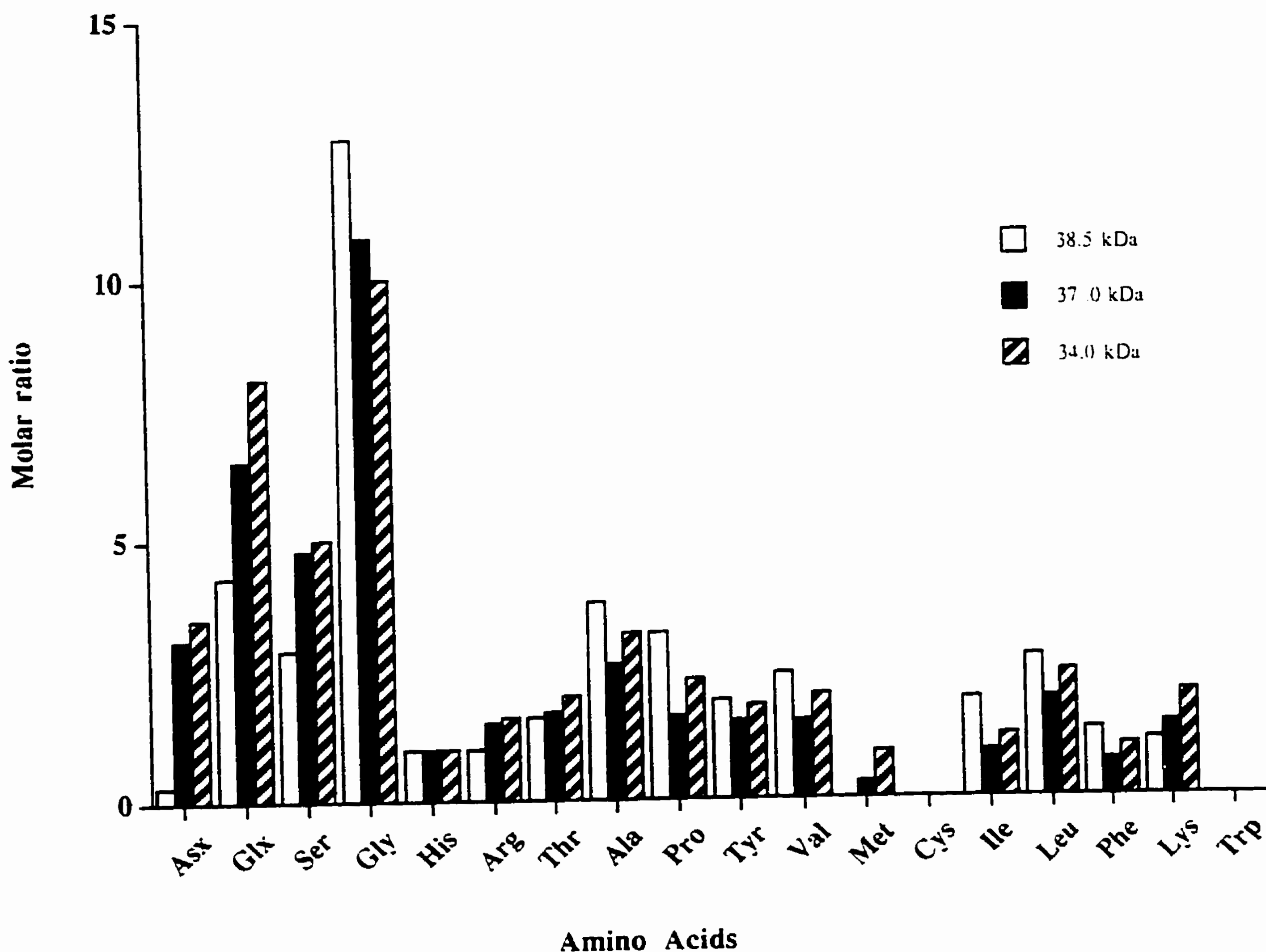


Figure 4-8. Amino acid composition of the three proteins isolated with the "b" affinity column. Positive fractions were pooled, concentrated and separated on denaturing gel with electrophoresis, followed by a transfer to a PVDF membrane for amino acid composition and N-terminal sequencing analysis. The three isolated proteins have similar amino acid composition and identical N-terminal sequence. All N-terminal sequences start with the possibilities of one of three amino acids: Glu, Ala or Gly, followed by the sequence: Gln-Pro-Glu-Ala.

Discussion

Purification of transcription factors from eukaryotic sources has been hampered by their scarcity within the cells and the yield for a purified protein is commonly in the range of a few micrograms or less (Briggs *et al.*, 1986; Jackson & Tjian, 1989). The development of new methods has circumvented these problems, such as the one hybrid system and the *in vitro* screening-method (Singh *et al.*, 1988; Silver & Humt, 1993). The caveat with these latter methods is the fact that several different proteins can interact through the same element. sometimes with opposite effects on transcription (Imataka *et al.*, 1992), rendering it difficult to pinpoint the protein involved in transcription. In addition, many transcription factors bind as heterodimers, which are not detected with these means.

The purification and characterization of transcription factors has led to the identification of several activation domains, which can be categorically divided into acidic-, glutamine-, proline- and serine-threonine residue rich domains. based on the prevalent amino acid. (Mitchell & Tjian, 1989). Repressor mediating regions are often rich in alanine, charged amino acids and surprisingly, proline or acidic rich residues (Hanna-Rose & Hansen, 1996). Other regions have been characterized as mediating either repression or activation without any predominant amino acid. These regions are believed to interact with the general transcription complex and modulate the effects of transcription (Mitchell & Tjian. 1989; Hanna-Rose & Hansen. 1996). No pattern has emerged with respect to which motif interacts with which general transcription factor. often with clear discrepancies in the literature. For example, the acidic activation domain from the activator Gal4 interacts with TFIIB, and the same activation domain has been shown to interact with TBP-TFIIA DNA-bound complex (Roberts *et al.*. 1993: Stargell & Struhl, 1995). As recently pointed out, substitution mutagenesis has not been

extensively carried out, rendering it difficult to rely only on the amino acid enrichment, when defining functional domains (Hanna-Rose & Hansen, 1996).

In this work the purification and a characterization of a putative repressor binding protein for the Type III AFP genes binds to the "b" element in wolffish is described. The protein was isolated from rainbow trout livers, using a series of chromatographic steps, including two different DNA-affinity columns. Three proteins were isolated with the apparent molecular weight between 34 and 38. 5 kDa, and are called pBB for "b" element binding protein. All three proteins have identical N-terminal sequence and similar amino acid sequence composition. Interestingly, all three proteins also appear to be rich in glutamic acid and glutamine residues, which could possibly suggest transcriptional repression domain.

After the crude nuclear extract is passed through a size exclusion chromatography column, DNA-binding activity is recovered in a fraction which that suggests the component is 60 kDa or larger. On the other hand, the cross-linking experiments are in good agreement with the molecular weight of the purified proteins, which suggested a DNA-binding protein around 36 kDa in size. This discrepancy could possibly reflect a homodimer, or a heterodimer complex, with one or all components being from the isolated proteins. The TATA box binding protein is present within the cell in several multi-protein complexes, that dictate the DNA-binding specificity of the complex (Hernandez, 1993). The presence of the three proteins of different molecular weight with identical N-termini could be due to microheterogenity. This could arise from phosphorylation or glycosylation. Furthermore, the difference of the size could also be explained with alternative splicing (Santoro *et al.*, 1988; Jackson & Tjian, 1989; Mylin *et al.*, 1989; Descombes & Schibler, 1991).

When the repeated arrays of the 130 bp fragment from the upstream region of the AFP gene in wolffish were used for the affinity column, three different DNA binding

activities are observed in the positive fraction. In addition to the expected "b" element, the "a and S" elements were similarly reactive in mobility shift assay. This confirms previous observations that several additional DNA-binding sites are present in this fragment (Chapter 3). Furthermore, an oligomer for the "a" element, also strengthens the previous finding that an "a"-like element is located in the -220 to -90 bp fragment in wolffish. The "S" element, has sequence similarities to the G/C box. This element binds the Sp1 transcriptional factor and several other DNA-binding proteins, which can mediate either activation or repression of transcription, depending on the factor bound (Imataka *et al.*, 1992; Kingsley & Winoto, 1992).

In conclusion, the putative repressor proteins, pBB, purified to homogeneity, and shown to bind to the "b" element, appear to be novel proteins. No identical sequences in protein sequencing data banks are detected with other N-terminal sequences using the Prosis software (1992).

Chapter 5

CONCLUSION

Conclusion and General Discussion

This work was initiated with two main objectives: first, to identify a repressor element in the promoter for the Type III AFP genes from wolffish. and second, to purify and characterize the protein interacting with this element. Both these objectives have been accomplished.

Characterization of a Repressor Element in the AFP Promoter in Wolffish

Before embarking on the actual characterization of the repressor element, a sequence alignment of all Type III AFP genes and their flanking sequences was carried out. The aim was to look for possible repeated DNA sequences from a segment which had been previously shown to harbor repressor activity. The repressor activity had been localized to -220 to -90 in the wolffish promoter (Gong & Hew, 1993). If sequences from this region were repeated elsewhere in and around the Type III AFP genes, either as direct or inverted repeats, then it might reflect functional importance, which would be focused on during the course of this work. The aforementioned region, which was shown to have repressor activity, is between -220 and -90 from the transcriptional start site in wolffish sequences number 1 and 2 (WO1 and WO2). This work did not discover any new repeated sequences within this segment. The -220 to -90 fragment in WO1 and WO2 has two 12 bp tandem directed repeats, as observed previously (Scott *et al.*, 1988). These elements are referred to as the "b" elements, based on alphabetical nomenclature for repeated elements identified in the flanking sequences in the Type III AFP genes. Based on this alignment, the Type III AFP genes can be divided into core and unique sequences. Core sequences occur in all AFP genes and their flanking regions, whereas the unique sequences are missing in one or more segments. The two direct repeats of "b" elements in wolffish, which are located in the -220 to -90 fragment, are located on

boundaries of core and unique regions. One "b" element is present in all Type III sequences, but the second "b" element is unique for WO1 and WO2 sequences.

When the -220 to -90 fragment was tested in the mobility shift assay, several bands were observed. This suggests that several proteins might be interacting with this fragment. Mobility shift experiments with the same fragment gave the strongest signal in liver tissues, when tested with nuclear extract from various sources. This agrees with previous results, which had suggested the liver as the predominant source for Type III AFP synthesis (Gong *et al.*, 1992). Furthermore, the "b" element alone binds strongly to proteins from the nucleus in the same experiments.

To localize the region involved in repression, a series of DNA constructs were made. The -220 to -90 fragment was ligated in front of a reporter gene, which was driven by thymidine kinase (TK) promoter. The "b" elements were ligated at different locations and orientations, into the same vector. In transfection studies with fish cell lines, the -220 to -90 fragment is a strong repressor as predicted. Furthermore, the "b" elements alone are shown to mediate repression, independent of their orientation and distance with respect to the transcriptional start site. The strongest repression is observed in fish cell lines as compared to mammalian cell lines.

The repressor is likely to act on the general transcription factors (Figure 5-1). The other alternative would be that the repressor acts through DNA bound activators, which would have to be common to both Type III AFP and TK promoter. These two promoters have in common, TATA, CAAT and G/C boxes. Both the CAAT and the G/C boxes can bind to number of different proteins. This would render it unlikely that the same factor would be binding to both promoters. Even though the repressor protein acts by interference with these aforementioned elements it should only affect activated transcription. In contrast, complete inhibition is sometimes observed which suggest that basal transcription is affected. Basal transcription is not supported by any activators, but

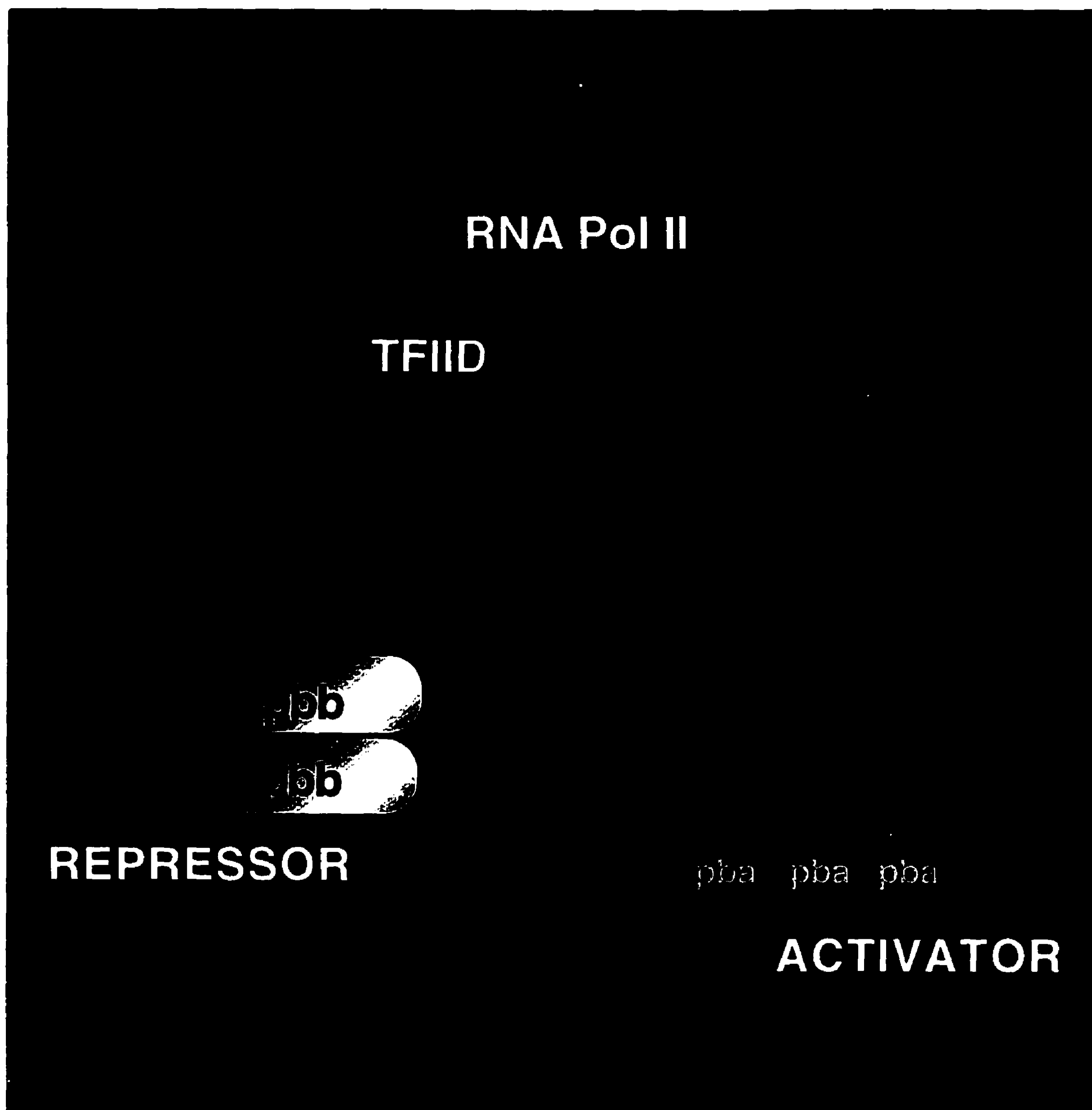


Figure 5-1. Model for the action of the pBB on the transcription complex. The repressor protein, pBB, is presumed to act directly on the general transcription machinery and repress transcription, where as the region surrounding and including the "a" elements, act as activators. This could either occur through the pBB or by direct interactions with the transcription complex.

utilizes only general transcription factors (Merino *et al.*, 1993). There are many putative interaction sites within the initiation complex, which have been previously shown to interact with repressors. TFIID interacts with several repressor protein, which prevents further assembly of the initiation complex (Inostroza *et al.*, 1992; Auble & Hahn, 1993; Merino *et al.*, 1993; Ge & Roeder, 1994; Stelzer *et al.*, 1994). The best characterized DNA-bound repressor protein to date is Krüppel protein, which can act either as an activator or a repressor, depending on its cellular concentration. At low concentrations the Kr protein acts as a activator, whereas at higher concentration it forms homodimers and represses transcription. As activator, Kr interacts with TFIIB, while as a repressor it interacts with TFIIEβ. Interestingly, TFIIEβ affects TFIIH, which in turn phosphorylates the CTD on RNA II pol concomitant to promoter clearance (Sauer & Jackel, 1991; Sauer *et al.*, 1995). If the repressor protein described here acts in similar way, can only remain speculative at this point.

The observation that several bands appear in the mobility shift experiment, suggests that several proteins might be interacting with the -220 to -90 fragment. To further explore this possibility, DNAse footprinting was performed on this fragment. Both the "b" elements were clearly protected in addition to at least two other sites. One of the sites has sequence similarity to G/C boxes, which interacts with the Sp1 transcription factor, among others. The element is therefore named "S" element. Also protected are sequences with resemblance to the "a" elements. The "a" elements, which were first noticed as being three repeated elements upstream of the "b" element in ocean pout and wolffish, are 8 bp long and show sequence similarities to the activator element for the ribosomal genes in *S. pombe* (Witt *et al.*, 1993; Witt *et al.*, 1995).

Purification of the Repressor Binding Protein

To further understand the transcriptional control of the Type III AFP genes, a project was undertaken to purify the repressor binding protein which interacts with the "b" element. Having established that the repressor binding protein is present and active in livers from rainbow trout, it was decided to use the liver as a source for the purification of the protein. Initially, the size of the putative repressor protein was estimated to be around 36 kDa. This was done by cross-linking a bromo-labeled oligomer for the "b" element, to its putative DNA-binding protein in crude nuclear extract. After large scale preparation of nuclear extract, the mixture was loaded onto a size exclusion chromatography column. Active fractions were eluted in the void volume, suggesting that the component was 60 kDa or larger in size. This might be due to complex formation, either by homo-polymerization or hetero-polymerization with unknown components in the nuclei. The positive fractions were sequentially loaded onto two different DNA-affinity columns. The former column consisted of the -220 to -90 DNA fragment, attached to a supporting matrix. Positive fractions were eluted with an increasing linear salt gradient. The DNA fragment on the column was identical to the one used for DNAse footprinting. Therefore, the positive fractions were tested in a mobility shift assay with oligomers corresponding to the "S, a and b" elements. All these oligomers were reactive, as opposed to an oligomer with the Ap3 element, which only binds to proteins in the crude nuclear extract. The second DNA-affinity column was made with the "b" element alone, attached to a supporting matrix. Positive fractions from the previous column were loaded and a high salt wash was used to elute bound protein. The positive fraction from the mobility shift assay showed three distinct bands on a SDS-PAGE gel. The size of the three proteins ranges from 34 to 38.5 kDa in size. This is in good agreement with the cross-linking experiment, which suggested a mass of around 36 kDa. Amino acid composition analysis showed that all three proteins have

high amounts of glutamine and/or glutamic acid residues, which are also present in activation and repression domains of *trans*-acting factors (Stargell & Struhl, 1995: Hanna-Rose & Hansen, 1996).

All three proteins have identical N-terminal sequences, as revealed by Edman degradation. Furthermore, a search in data bases with protein sequences found no similarities to other known N-terminal sequences. Many proteins have these sequences inside their proteins. The fact that these proteins have similar amino acid content, suggest that all three might have arisen due to micro-heterogeneity. This could be caused by postranslational modification that stems from phosphorylation or glycosylation (Jackson & Tjian, 1988; Mylin *et al.*, 1989). Alternative splicing is also possible (Santoro *et al.*, 1988; Descombes & Schibler, 1991).

During the course of this work, several new aspects of transcriptional control for the Type III genes were elucidated. All the genes can be partitioned into core and unique sequences. The two "b" elements, found in WO1 and WO2 are in each separate region. This element, whose function had not been characterized before, was shown to mediate a strong repression for transcription. Furthermore, the purification of its cognate binding protein, revealed three distinct proteins, each with an identical N-terminus. By all account, these appear to be novel proteins.

Future Experiments and Prospects

These experiments will open up several new aspects for the research of the transcriptional control of the Type III AFP genes. The salient feature of these experiments is the isolation of a putative repressor protein, which interacts with the "b" element. The same element is present in all flanking sequences of the Type III AFP genes. Further characterization of the protein, by purification and more N-terminal

sequencing, would allow the cloning of pBB, which then would be accessible to delineate repression domains involved in the transcriptional control.

These experiments did not delineate the seasonal control of the Type III AFP in fish caught off the coast of Newfoundland. Livers collected during the summer and winter months, from ocean pout, can be used for DNase footprinting, to identify different DNA-binding patterns of transcriptional factors during the different seasons. Antibodies raised against the repressor protein, could also be used to probe if the concentration of the protein fluctuates with the seasons.

REFERENCES

(1992). DNASIS™ (Brisbane, California, USA: Hitachi Softwear Engineering LTD).

Abate, C., Luk, D., Gagne, E., Roeder, R. & Curran, T. (1990). Fos and jun cooperate in transcriptional regulation via heterologous activation domains. *Molecular & Cellular Biology* **10,** 5532-5.

Akerblom, I. J., Slater, E. P., Beato, M., Baxter, J. D. & Mellon, P. L. (1988). Negative regulation by glucocorticoids through interference with a cAMP responsive enhancer. *Science* **241,** 350-53.

Alberts, B., BBay, D., Lewis, J., Raff, M., Roberts, K. & Watson, J. D. (1983). Molecular biology of the Cell (New York: Garland).

Arnone, M., Zannini, M. & Di, L. R. (1995). The DNA binding activity and the dimerization ability of the thyroid transcription factor I are redox regulated. *Journal of Biological Chemistry* **270,** 12048-55.

Attardi, L. D. & Tjian, R. (1993). *Drosophila* tissue-specific transcription factor NTF-1 contains a novel isoleucine-rich activation motif. *Genes & Development* **7,** 1341-53.

Auble, D. T. & Hahn, S. (1993). An ATP-dependent inhibitor of TBP binding to DNA. *Genes & Development* **7,** 844-56.

Auble, D. T., Hansen, K. E., Mueller, C. G., Lane, W. S., Thorner, J. & Hahn, S. (1994). Mot1, a global repressor of RNA polymerase II transcription, inhibits TBP binding to DNA by an ATP-dependent mechanism. *Genes & Development* **8,** 1920-34.

Austin, R. J. & Biggin, M. D. (1995). A domain of the even-skipped protein represses transcription by preventing TFIID binding to a promoter: repression by cooperative blocking. *Molecular & Cellular Biology* **15,** 4683-93.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1992). Short Protocols in Molecular Biology, Second Edition (New York: Green Publishing Associates and John Wiley & Sons).

Barinaza, M. (1991). Dimers direct development. *Science* **251,** 1176-77.

Beg, A. A. & Baldwin, A. S. (1993). The IκB proteins: multifunctional regulators of Rel/NF-κB transcriptional factors. *Genes & Development* **7,** 2064-70.

Benezra, R., Davis, R. L., Lockshorn, D., Turner, D. L. & Wintraub, H. (1990). The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell* **61,** 49-59.

Berberich, S. J. & Cole, M. D. (1992). Casein kinase II inhibits the DNA binding activity of Max homodimers, but not Myc/Max heterodimers. *Genes and Development* **6,** 166-76.

Bhat, P. J. & Hopper, J. E. (1992). Overproduction of the GAL1 or GAL3 protein causes galactose-independent activation of the GAL4 protein: evidence for a new model of induction for the yeast GAL/MEL regulon. *Molecular and Cellular Biology* **12,** 2701-7.

Binétruy, B., Smeal, T. & Karin, M. (1991). Ha-Ras augments c-Jun activity and stimulates phosphorylation of its activation domain. *Nature* **351,** 122-27.

Bissette, G., Griff, D., Carnes, M., Goodman, B., Lavine, M. & Levant, B. (1995). Apparent seasonal rhythms in hypothalamic neuropeptides in rats without photoperiod changes. *Endocrinology* **136,** 622-8.

Blank, V., Kourilsky, P. & Israel, A. (1992). NF-κ B and related proteins: Rel/dorsal homologies meet ankyrin-like repeats. *Trends in Biochemical Sciences* **17,** 135-40.

Boyle, W. J., Smeale, T., Defize, L. H., Angel, P., Woodgett, J. R. & Karin, M. (1991). Activation of protein kinase C decreases phosphorylation of c-Jun at site that negatively regulate its DNA-binding activity. *Cell* **64,** 573-84.

Briggs, M., Kadonaga, J., Bell, S. & Tjian, R. (1986). Purification and biochemical characterization of the promoter-specific transcription factor. Sp1. *Science* **234,** 47-52.

Buchanan, B. B., Schurmann, P., Decottignies, P. & Lozano, R. M. (1994). Thioredoxin: a multifunctional regulatory protein with a bright future in technology and medicine. *Archives of Biochemistry & Biophysics* **314,** 257-60.

Burns, K., Duggan, B., Atkinson, E. A., Famulski, K. S., Nemer, M., Bleackley, C. & Michalak, M. (1994). Modulation of gene expression by calreticulin binding to the glucocorticoid receptor. *Nature* **367,** 476-80.

Calberg, C. & Wiesenberg, I. (1995). The orphan receptor family RZR/ROR, melatonin and 5-lipoxygenase: An unexpected relationship. *Journal of Pineal Research* **18,** 171-78.

Carey, M. (1994). Simplifying the complex. *Nature* **368,** 402-3.

Carey, M., Lin, Y. S., Green, M. R. & Ptashne, M. (1990). A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. *Nature* **345,** 361-4.

Caserta, M., Camilloni, G., Venditti, S., Venditti, P. & Di Mauro, E. (1994). Conformational information in DNA: its role in the interaction with DNA topoisomerase I and nucleosomes. *Journal of Cellular Biochemistry* **55,** 93-7.

Chan, S. L. (1995). Transcriptional control of the gene encoding the winterflounder antifreeze protein: University of Toronto).

Chellappan, S., Hiebert, S., Mudryj, M., Horowitz, J. & Nevins, J. (1991). The E2F transcription factor is a cellular target for the RB protein. *Cell* **65**, 1053-61.

Chen, J. D. & Evans, R. M. (1995). A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature* **377**, 454-7.

Chiang, C. & Roeder, R. (1995). Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators. *Science* **267**, 531-6.

Choder, M. (1991). A general topoisomerase I-dependent transcriptional repression in the stationary phase in yeast. *Genes & Development* **5**, 2315-26.

Christensen, M. & Pusey, N. (1994). Binding of nuclear proteins to transcription regulatory elements in selenium deficiency. *Biochimica et Biophysica Acta* **1225**, 338-41.

Clark, M. E., Lieberman, P. M., Berk, A. J. & Dasgupta, A. (1993). Direct cleavage of human TATA-binding protein by poliovirus protease 3C *in vivo* and *in vitro*. *Molecular & Cellular Biology* **13**, 1232-7.

Collins, J. F. & Coulson, A. F. W. (1987). Molecular sequencing comparison and alignment. In Nucleic Acid and Protein Sequence Analysis: a Practical Approach, M. J. Bishop and C. J. Rawlings, eds. (Oxford: IRL Press).

Cooper, C., Henderson, A., Artandi, S., Avitahl, S. & Calame, K. (1995). Ig/EBP(C/EBPg) is transdominant negative inhibitor of C/EBP family transcriptional activators. *Nucleic Acids Resear*h **23**, 4371-7.

Cowell, I. G. (1994). Repression versus activation in the control of gene transcription. *Trends in Biochemical Sciences* **19**, 38-42.

Davies, P. L. & Hew, C. L. (1990). Biochemistry of fish antifreeze proteins. *FASEB Journal* **4**, 2460-8.

Davies, P. L., Hew, C. L. & Fletcher, G. L. (1988). Fish antifreeze proteins: physiology and evolutionary biology. *Canadian Journal of Zoology* **66**, 2611-7.

de Groot, R. P., Auwerx, J., Bourouis, M. & Sassone-Corsi, P. (1992). Negative regulation of Jun/AP-1: conserved function of glycogen synthase kinase 3 and the *Drosophila* kinase *shaggy*. *Oncogene* **7**, 841-7.

DeCaprio, J. A., Ludlow, J. W., Lynch, D., Furukawa, Y., Griffin, J., Piwnica-Worms, H., Huang, C. M. & Livingston, D. M. (1989). The product of the retinoblastoma susceptibility gene has properties of a cell cycle regulatory element. *Cell* **58**, 1085-95.

Dedhar, S., Rennie, P. S., Shago, M., Hagestejn, C. L., Yang, H., Filmus, J., Hawley, R. G., Bruchowsky, N. & Giugère, V. (1994). Inhibition of nuclear hormone receptor activity by calreticulin. *Nature* **367**, 480-3.

Descombes, P. & Schibler, U. (1991). A liver-enriched transcriptional activator protein, LAP, and a transcriptional inhibitory protein, LIP, are translated from the same mRNA. *Cell* **67**, 569-9.

DeVries, A. L. (1980). Biological antifreezes and the survival in freezing environments. In Animals and Environmental Fitness, R. Gilles, ed. (Oxford: Pergamon Press), pp. 583-607.

DeVries, A. L. (1982). Biological antifreeze agents in coldwater fishes. *Comparative Biochemical Physiology* **73A,** 627-40.

DeVries, A. L., Komatsu, S. K. & Feeney, R. E. (1970). Chemical and physical properties of freezing-point depressing glycoproteins from Antarctic fishes. *Journal of Biological Chemistry* **245,** 2901-8.

DeVries, A. L. & Wohlschlag, D. E. (1969). Freezing resistance in some Antarctic fish. *Science* **163,** 1073-5.

Dikstein, R., Ruppert, S. & Tjian, R. (1996). TAFII250 is a bipartite protein kinase that phosphorylates the basal transcription factor RAP74. *Cell* **84**, 781-90.

Dillin, A. & Rine, J. (1995). On origin of a silencer. *Trends in Biochemical Sciences* **20**, 231-5.

Dostatni, N., Lambert, P. F., Sousa, R., Ham, J., Howley, P. M. & Yaniv, M. (1991). The functional BPV-1 E2 trans-activating protein can act as a repressor by preventing formation of the initiation complex. *Genes & Development* **5**, 1657-71.

Drew, P. D., Franzoso, G., Carlson, L. M., Biddison, W. E., Siebenlist, U. & Ozato, K. (1995). Interferon regulatory factor-2 physically interacts with NF-κB in vitro and inhibits NF-κB induction of major histocompatibility class I and β2-microglobulin gene expression in transfected human neuroblastoma cells. *Journal of Neuroimmunology* **63,** 157-62.

Dröge, W., Schulze-Osthoff, K., Mihm, S., Galter, D., Schenk, H., Eck, H. P., Roth, S. & Gmünder, H. (1994). Functions of glutathione and glutathione disulfide in immunology and immunopathology. *FASEB Journal* **8**, 1131-8.

Du, S. J., Gong, Z., Hew, C. L., Tan, C. H. & Fletcher, G. L. (1992a). Development of an all-fish gene cassette for gene transfer in aquaculture. *Molecular Marine Biology & Biotechnology* **1,** 290-300.

Du, S. J., Gong, Z. Y., Fletcher, G. L., Shears, M. A., King, M. J., Idler, D. R. & Hew, C. L. (1992b). Growth enhancement in transgenic Atlantic salmon by the

use of an "all fish" chimeric growth hormone gene construct. *Biot/Technology* **10,** 176-81.

Duman, J. G. & DeVries, A. L. (1974). Freezing resistance in winterflounder, *Pseudopleuronectus americanus*. *Nature* **247,** 237-8.

Durand, B., Saunders, M., Leroy, P., Leid, M. & Chambon, P. (1992). All-trans and 9-cis retinoic acid induction of CRABPII transcription is mediated by RAR-RXR heterodimer bound to DR1 and DR2 repeated motifs. *Cell* **71,** 73-85.

Durnan, J. G. & Olsen, T. M. (1993). Thermal hysteresis protein activity in bacteria, fungi, and phylogenetically diverse plants. *Cryobiology* **30,** 322-8.

Evans, R. M. (1988). The steroid and thyroid hormone receptor superfamily. *Science* **240**,, 889-95

Ewart, K. V. & Fletcher, G. L. (1993). Herring antifreeze protein: primary structure and evidence for a C-type lectin evolutionary origin. *Molecular Marine Biology & Biotechnology* **2,** 20-7.

Ewart, K. V., Rubinsky, B. & Fletcher, G. L. (1992). Structural and functional similarity between fish antifreeze proteins and calcium-dependent lectins. *Biochemical & Biophysical Research Communications* **185,** 335-40.

Ewart, K. V., Yang, D. S. C., Ananthanarayanan, V. S., Fletcher, G. L. & Hew, C. L. (1996). Ca^{2+} -dependent antifreeze proteins. *Journal of Biochemistry* **271,** 16627-32.

Faisst, S. & Meyer, S. (1992). Compilation of vertebrate-encoded transcription factors. *Nucleic Acid Research* **20,** 3-26.

Fanning, E. & Knippers, R. (1992). Structure and function of Simian Virus 40 large Tumor Antigen. *Annual Review of Biochemistry* **61,** 55-85.

Fattaey, A., Helin, K. & Harlow, E. (1993). Transcriptional inhibition by the retinoblastoma protein. *Philosophical transactions of the Royal Society of London - Series B: Biological Sciences* **340,** 333-6.

Fletcher, G. L. (1979). The effects of hypophysectomy and pituitary replacement on the plasma freezing point depression, Cl^-, glucose and protein antifreeze in the winterflounder (*Pseudopleuronectus americanus*). *Comparative Biochemical Physiology* **63,** 535-7.

Fletcher, G. L. (1981). Effects of temperature and photoperiod on the plasma freezing point depression Cl^-, concentrations, and protein antifreeze in the winter flounder. *Canadian Journal of Zoology* **59,** 193-201.

Fletcher, G. L., Campbell, C. M. & Hew, C. L. (1978). The effects of hypophysectomy on seasonal changes in plasma freezing-point depression.

protein 'antifreeze,' and Na+ and Cl- concentrations of winter flounder (*Pseudopleuronectes americanus*). *Canadian Journal of Zoology* **56,** 109-13.

Fletcher, G. L., King, M. J. & Hew, C. L. (1984). How does the brain control the brain pituitary's release of antifreeze synthesis inhibitor. *Canadian Journal of Zoology* **62,** 839-44.

Fletcher, G. L., King, M. J. & Kao, M. H. (1987). Low temperature regulation of antifreeze glycopeptide levels in Atlantic cod (*Gadus morhua*). *Canadian Journal of Zoology* **65,** 227-33.

Fourney, R. M., Fletcher, G. L. & Hew, C. L. (1984a). Accumulation of winter flounder antifreeze messenger RNA after hypophysectomy. *General & Comparative Endocrinology* **54,** 392-401.

Fourney, R. M., Fletcher, G. L. & Hew, C. L. (1984b). The effects on long daylight on liver antifreeze mRNA in the winter flounder, *Pseudopleuronectus americanus*. *Canadian Journal of Zoology* **62,** 1456-60

Galang, C. & Hauser, C. (1993). Cooperative DNA binding of the human HoxB5 (Hox-2.1) protein is under redox regulation in vitro. *Molecular & Cellular Biology* **13,** 4609-17.

Garcia, J. A., Ou, S. H., Wu, F., Lusis, A. J., Sparkes, R. S. & Gaynor, R. B. (1992). Cloning and chromosomal mapping of a human immunodeficiency virus 1 "TATA" element modulatory factor. *Proceedings of the National Academy of Sciences of the United States of America* **89,** 9372-6.

Ge, H. & Roeder, R. (1994). The high mobility group protein HMG1 can reversibly inhibit class II gene transcription by interaction with the TATA-binding protein. *Journal of Biological Chemistry* **269,** 17136-40.

Geisler, R., Bergmann, A., Hiromi, Y. & Nüsslein-Volhard, C. (1992). *cactus*, a gene involved in dorsoventral pattern formation of *Drosophila*, is related to the IκB gene family of vertebrates. *Cell* **71,** 613-21.

George, M. F. & Burke, M. J. (1984). Supercooling of tissue water to extreme low temperatures in overwintering plants. *Trends in Biochemiccal Sciences* **9,** 211-4.

Girling, R., Partridge, J. F., Bandara, L. R., Burden, N., Totty, N. F., Hsuan, J. J. & La Thangue, N. B. (1993). A new component of the transcription factor DRTF1/E2F. *Nature* **362,** 83-7.

Glover, J. M. & Harrison, S. C. (1995). Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. *Nature* **373,** 257-61.

Goll, D. E., Thompson, V. F., Taylor, R. G. & Zalewska, T. (1992). Is calpein activity regulated by membranes and autolysis or by calcium calpstatin? *BioEssays* **14,** 549-56.

Gong, Z., Ewart, K. V., Hu, Z., Fletcher, G. L. & Hew, C. L. (1996). Skin antifreeze protein genes of the winter flounder, *Pleuronectus americanus*, encode distinct and active polypeptides without the secretory signal and prosequences. *Journal of Biochemical Chemistry* **271**, 4106-12.

Gong, Z., Fletcher, G. L. & Hew, C. L. (1992). Tissue distribution of antifreeze protein mRNAs. *Canadian Journal of Zoology* **70**, 810-4.

Gong, Z. & Hew, C. L. (1993). Promoter analysis of antifreeze protein genes. In Biochemistry and Molecular Biology of Fishes, Hochachka and Mommsen, eds.: Elsevier Science), pp. 307-24.

Gong, Z., Hew, C. L. & Vielkind, J. R. (1991). Functional analysis and temporal expression of promoter regions from fish antifreeze protein genes in transgenic Japanese medaka embryos. *Molecular Marine Biology & Biotechnology* **1**, 64-72.

Graham, G. (1995). Tandem genes and clustered genes. *Journal of Theoretical Biology* **175**, 71-87.

Grimm, S. & Baeuerle, P. A. (1993). The inducible transcription factor NF-κB: structure-function relationship of its protein subunits. *Biochemical Journal* **290**, 297-308.

Guehmann, S., Vorbrueggen, G., Kalkbrenner, F. & Moelling, K. (1992). Reduction of a conserved Cys is essential for Myb DNA-binding. *Nucleic Acids Research* **20**, 2279-86.

Hagemeier, C., Cook, A. & Kouzarides, T. (1993). The retinoblastoma protein binds E2F residues required for activation in vivo and TBP binding in vitro. *Nucleic Acids Research* **21**, 4998-5004.

Hagen, G., Muller, S., Beato, M. & Suske, G. (1994). Sp1-mediated transcriptional activation is repressed by Sp3. *EMBO Journal* **13**, 3843-51.

Haglund, B., Cnattingius, S. & Otterblad-Olausson, P. (1995). Sudden infant death syndrome in Sweden, 1983-1990: season at death, age at death, and maternal smoking. *American Journal of Epidemiology* **142**, 619-24.

Hainaut, P. & Milner, J. (1993). Redox modulation of p53 conformation and sequence-specific DNA binding in vitro. *Cancer Research* **53**, 4469-73.

Hanna-Rose, W. & Hansen, U. (1996). Active repression mechanism of eukaryotic transcription repressors. *Trends in Genetics* **12**, 229-34.

Harrison, S. C. (1991). A structural taxonomy of DNA-binding domains. *Nature* **353**, 715-9.

Hayashi, T., Ueno, Y. & Okamoto, T. (1993). Oxidoreductive regulation of nuclear factor κ B. Involvement of a cellular reducing catalyst thioredoxin. *Journal of Biological Chemistry* **268**, 11380-8.

Hayes, P. H., Scott, G. K., Ng, N. F., Hew, C. L. & Davies, P. L. (1989). Cystine-rich type II antifreeze protein precursor is initiated from the third AUG codon of its mRNA. *Journal of Biological Chemistry* **264,** 18761-7.

He, G., Mulse, A., Li, A. & Ro, H. (1995). A eukaryotic transcriptional repressor with carboxypeptidase activity. *Nature* **378,** 92-6.

Henninghausen, L. & Lubon, H. (1987). Interaction of proteins with DNA *in vitro*. In Guide to Molecular Cloning Techniques, S. L. Berger and A. R. Kimmel, eds. (San Diego: Academic Press Inc.), pp. 721-34.

Herschbach, B. M. & Johnson A. D. (1993). Transcriptional repression in eukaryotes. Annual Review of Cell Biology **9,** 479-509

Hernandez, N. (1993). TBP, a universal eukaryotic transcription factor? *Genes & Development* 7, 1291-308.

Hew, C. L. & Fletcher, G. L. (1979). The role of pituitary in regulating antifreeze protein synthesis in the winter flounder. *FEBS Letters* **99,** 337-9.

Hew, C. L., Fletcher, G. L. & S., A. V. (1980). Antifreeze proteins from the shorthorn sculpin, Myoxocephalus scorpius: isolation and characterization. *Canadian Journal of Biochemistry* **58,** 377-83.

Hew, C. L., Wang, N. C., Joshi, S., Fletcher, G. L., Scott, G. K., Hayes, P. H., Buettner, B. & Davies, P. L. (1988). Multiple genes provide the basis for antifreeze protein diversity and dosage in the ocean pout, *Macrozoarces americanus*. *Journal of Biological Chemistry* **263,** 12049-55.

Hew, C. L., Wang, N. C., Yan, S., Cai, H., Sclater, A. & Fletcher, G. L. (1986). Biosynthesis of antifreeze polypeptides in the winter flounder. Characterization and seasonal occurrence of precursor polypeptides. *European Journal of Biochemistry* **160,** 267-72.

Hill, C. S. & Treisman, R. (1995). Transcriptional regulation by extracellular signals; mechanism and specificity. *Cell* **80,** 199-211.

Hoeijmakers, J. H. J. (1990). Cryptic initiation sequence revealed. *Nature* **343,** 417-418.

Hoey, T., Dynlacht, B. D., Peterson, M. G., Pugh, B. F. & Tjian, R. (1990). Isolation and characterization of the *Drosophila* gene encoding the TATA box binding protein, TFIID. *Cell* **61,** 1179-86.

Hoffman, M. (1991). New role found for a common protein "motif". *Science* **253,** 742.

Holdaway, S., Hearn, R. A., Lindsay, S. & Hoke, T. (1994). Claris Works™ (Santa Clare: Claris Corporation ©).

Holmes, W. N. & Donaldson, E. M. (1969). The body compartments and distribution of electrolytes. In Fish Physiology, W. S. Hoar and D. J. Randall, eds. (New York: Academic Press).

Hon, W., Griffith, M., Chong, P. & Yang, D. S. (1994). Extraction and isolation of antifreeze proteins from winter rye (Secale cereale L.) leaves. *Plant Physiology* **104,** 971-80.

Horikoshi, M., Wang, C., Fujii, H., Cromlish, J., Weil, P. & Roeder, R. (1989). Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box. *Nature* **341,** 299-303.

Hsiao, K., Cheng, C., Fernandes, I., Detrich, H. & DeVries, A. (1990). An antifreeze glycopeptide gene from the antarctic cod Notothenia coriiceps neglecta encodes a polyprotein of high peptide copy number *Proceedings of the National Academy of Sciences of the United States of America* **87,** 9265-9.

Hunter, T. & Karin, M. (1992). The regulation of transcription by phosphorylation. *Cell* **70,** 375-87.

Hörlein, A. J., Näär, A. M., Heinzel, T., Torchia, J., Gloss, B., Kurokawa, R., Ryan, A., Kamei, Y., Söderström, M., Glass, C. K. & Rosenfeld, M. G. (1995). Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature* **377,** 397-404.

Idler, D. R., Fletcher, G. L., Belkhode, S., King, M. J. & Hwang, S. J. (1989). Regulation of antifreeze protein production in winter flounder: a unique function for growth hormone. *General & Comparative Endocrinology* **74,** 327-34.

Imataka, H., Sogawa, K., Yasumoto, K., Kobayashi, A., Hayami, M. & Fujii-Kuriyama, Y. (1992). Two regulatory proteins that bind to the basic transcription element (BTE), a GC box sequence in the promoter region of the rat P4501A1 gene. *EMBO Journal* **11,** 3663-71.

Inostroza, J. A., Mermelstein, F. H., Ha, I., Lane, W. S. & Reinberg, D. (1992). Dr1, a TATA-binding protein-associated phosphoprotein and inhibitor of class II gene transcription. *Cell* **70,** 477-89.

Jackson, S. & Tjian, R. (1988). O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation. *Cell* **55,** 125-33.

Jackson, S. & Tjian, R. (1989). Purification and analysis of RNA polymerase II transcription factors by using wheat germ agglutinin affinity chromatography. *Proceedings of the National Academy of Sciences of the United States of America* **86,** 1781-5.

Jia, Z., Delucia, C. I., Chao, H. & Davies, P. L. (1996). Structural basis for the binding of a globular antifreeze protein to ice. *Nature* **384,** 285-8.

Johnson, A. D. (1995). The price of repression. *Cell* **81,** 655-8.

Johnson, D. G., Schwarz, J. K., Cress, W. D. & Nevins, J. R. (1993). Expression of transcription factor E2F1 induces quiescent cells to enter S phase. *Nature* **365,** 349-52.

Johnson, F. B. & Krasnow, M. A. (1992). Differential regulation of transcription preinitiation complex assembly by activator and repressor homeo domain proteins. *Genes & Development* **6,** 2177-89.

Jupp, R., Hoffmann, S., Depto, A., Stenberg, R. M., Ghazal, P. & Nelson, J. A. (1993). Direct interaction of the human cytomegalovirus IE86 protein with the cis repression signal does not preclude TBP from binding to the TATA box. *Journal of Virology* **67,** 5595-604.

Kääriäinen, L. & Ranki, M. (1984). Inhibition of cell functions by RNA-virus infections. *Annual Review of Microbiology* **38,** 91-109.

Kadonaga, J. T. (1990). Assembly and disassembly of the *Drosophila* RNA polymerase II complex during transcription. *Journal of Biological Chemistry* **265,** 2624-31.

Kadonaga, J. T., Carner, K. R., Masiarz, F. R. & Tjian, R. (1987). Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* **51,** 1079-90.

Kao, C. K., Lieberman, P. M., Schmidt, M. C., Zhou, Q., Pei, R. & Berk, A. (1990). Cloning of a transcriptionally active human TATA binding factor. *Science* **248,** 1646-9.

Kao, M. H. & Fletcher, G. L. (1988). Juvenile cod (*Gadus morhua*) can be more freeze resistant than adults. *Canadian Journal of Aquatic Science* **45,** 902-5.

Kapiloff, M. S., Farkash, Y., Wegner, M. & Rosenfeld, M. G. (1991). Variable effects of phosphorylation of Pit-1 dictated by the DNA response element. *Science* **253,** 786-9.

Kaufman, P. D. & Rio, D. C. (1991). *Drosophila* P-element transposase is a transcriptional repressor in vitro. *Proceedings of the National Academy of Sciences of the United States of America* **88,** 2613-7.

Kilduff, T. S., Krilowicz, B., Milsom, W. K., Trachsel, L. & Wang, L. C. H. (1993). Sleep and mammalian hibernation: Homologous adaptations and processes. *Sleep* **16,** 372-86.

Kim, J. L., Nikolov, D. B. & Burley, S. K. (1993a). Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365,** 520-7.

Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. B. (1993b). Crystal structure of a yeast TBP/TATA-box complex. *Nature* **365,** 512-20.

Kingsley, C. & Winoto, A. (1992). Cloning of GT box-binding proteins: a novel Sp1 multigene family regulating T-cell receptor gene expression. *Molecular & Cellular Biology* **12,** 4251-61.

Kliewer, S. & Dasgupta, A. (1988). An RNA polymerase II transcription factor inactivated in poliovirus-infected cells copurifies with transcription factor TFIID. *Molecular & Cellular Biology* **8,** 3175-82.

Koleske, A. J. & Young, R. A. (1994). An RNA polymerase II holoenzyme responsive to activators. *Nature* **368,** 466-9.

Komatsu, S., DeVries, A. L. & Feeney, R. E. (1970). Studies of the structure of freezing point-depressing glycoproteins from an Antarctic fish. *Journal of Biological Chemistry* **245,** 2909-13.

Korf, H. W. (1994). The pineal organ as a component of the biological clock. Phylogenetic and ontogenetic considerations. *Annals of the New York Academy of Sciences* **719,** 13-42.

Kunzler, M., Braus, G. H., Georgiev, O., Seipel, K. & Schaffner, W. (1994). Functional differences between mammalian transcription activation domains at the yeast GAL1 promoter. *EMBO Journal* **13,** 641-5.

Kurokawa, R., DiRenzo, J., Boehm, M., Sugarman, J., Gloss, B., Rosenfeld, M. G., Heyman, R. A. & Glass, C. K. (1994). Regulation of retinoid signalling by receptor poalarity and allosteric control of ligand binding. *Nature* **371,** 528-30.

Kurokawa, R., Söderström, M., Hörlein, A., Halachmi, S., Brown, M., Rosenfeld, M. G. & Glass, C. K. (1995). Polarity-specific activities of retinoic acid receptors determined by a co-repressor. *Nature* **377,** 451-4.

Lawson, M. A. & Semler, B. L. (1990). Picornavirus protein processing--enzymes, substrates, and genetic regulation. *Current Topics in Microbiology & Immunology* **161,** 49-87.

Lee, G., Wu, J., Luu, P., Ghazal, P. & Flores, O. (1996). Inhibition of the association of RNA polymerase II with the preinitiation complex by a viral transscriptional repressor. *Proceedings of the National Academy of Sciences of the United States of America* **93,** 2570-5.

Lesnikov, V. & Pierpaoli, W. (1994). Pineal cross-transplantation (old-to-young and vice versa) as evidence for an endogenous "aging clock". *Annals of the New York Academy of Sciences* **719,** 456-60.

Leuther, K. K. & Johnston, S. A. (1992). Nondissociation of GAL4 and GAL80 in vivo after galactose induction. *Science* **256,** 1333-5.

Leuther, K. K., Salmeron, J. M. & Jonston, S. A. (1993). Genetic evidence that an activation domain of the GAL4 does not require acidity and may form a β sheet. *Cell* **72,** 575-585.

Lewin, B. (1990). Genes IV (Oxford: Oxford University Press).

Levine, M. & Manley, S. J. (1989). Ttranscriptional repression of eukaryotic promoters. *Cell* **59,** 405-8.

Li, Y., Slansky, J., Myers, D., Drinkwater, N., Kaelin, W. & Farnham, P. (1861). Cloning, chromosomal location, and characterization of mouse E2F1. *Molecular & Cellular Biology* **14,** 1861-9.

Licht, J. D., Grossel, M. J., Figge, J. & Hansen, U. M. (1990). *Drosophila* Krüppel protein is a transcriptional repressor. *Nature*, 76-9.

Lin, A., Frost, J., Deng, T., Smeal, T., al-Alawi, N., Kikkawa, U., Hunter, T., Brenner, D. & Karin, M. (1992). Casein kinase II is a negative regulator of c-Jun DNA binding and AP-1 activity *Cell* **70,** 777-89.

Liu, X., Miller, C. W., Koeffler, P. H. & Berk, A. J. (1993). The p53 activation domain binds the TATA box-binding polypeptide in Holo-TFIID, and a neighboring p53 domain inhibits transcription. *Molecular & Cellular Biology* **13,** 3291-300.

Luckow, B. & Schütz, G. (1987). CAT constructions with multiple unique restriction sites for the functional analysis of eukaryotic promoter and regulatory elements. *Nucleic Acids Research* **15,** 5490.

Lüscher, B., Christenson, E., Lithcfield, D. W., Krebs, E. G. & Eisenman, R. N. (1990). Myb DNA binding inhibited by phosphorylation at a site deleted during oncogenic activation. *Nature* **344,** 517-22.

MacGregor, G. R. & Caskey, C. T. (1989). Construction of plasmids that express *E. coli* β-galactose in mammalian cells. *Nucleic Acid Research* **17,** 2365.

Maes, M., Scharpe, S., Verkerk, R., D'Hondt, P., Peeters, D., Cosyns, P., Thompson, P., De, M. F., Wauters, A. & Neels, H. (1995). Seasonal variation in plasma L-tryptophan availability in healthy volunteers. Relationships to violent suicide occurrence. *Archives of General Psychiatry* **52,** 937-46.

Mangelsdorf, D. J. & Evans, R. M. (1995). The RXR heterodimers and orphan receptors. *Cell* **83,** 841-50.

Mangelsdorf, D. J., Umesono, K., Kliewer, S. A., Borgmeyer, U., Ong, E. S. & Evans, R. M. (1991). A direct repeat in the cellular retinol-binding protein type II gene confers differntial regulation by RXR and RAR. *Cell* **66,** 555-561.

Marmorstein, R., Carey, M., Pthasne, M. & Harrison, S. C. (1992). DNA recognition by GAL4: structure of a protein-DNA complex. *Nature* **356,** 408-14.

Matthews, J. R. & Hay, R. T. (1995). Regulation of the DNA binding activity of NF-κB. *International Journal of Biochemistry & Cell Biology* **27,** 865-79.

Matthews, J. R., Wakasugi, N., Virelizier, J. L., Yodoi, J. & Hay, R. T. (1992). Thioredoxin regulates the DNA binding activity of NF-κB by reduction of a disulphide bond involving cysteine 62. *Nucleic Acids Research* **20,** 3821-30.

Merino, A., Madden, K. R., Lane, W. S., Champoux, J. J. & Reinberg, D. (1993). DNA topoisomerase I is involved in both repression and activation of transcription. *Nature* **365,** 227-32.

Miller, H. (1987). Practical aspects of preparing phage and plasmid DNA: Growth maintenance and storage of bacteria and bacteriophage. In Guide to Molecular Cloning Techniques, S. Berger and A. Kimmel, eds. (San Diego: Academic Press Inc.), pp. 145-2.

Mitchell, P. J. & Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245,** 371-8.

Moqtaderi, Z., Bai, Y., Poon, D., Weil, P. A. & Struhl, K. (1996). TBP-associated factors are not generally required for transcriptional activation in yeast. *Nature* **383,** 188-91.

Mylin, L. M., Bhat, J. P. & Hopper, J. E. (1989). Regulated phosphorylation and dephosphorylation of GAL4, a transcriptional activator. *Genes and Development* **3,** 1157-65.

Ng, L., Forrest, D. & Curran, T. (1993). Differential roles for Fos and Jun in DNA-binding: redox-dependent and independent functions. *Nucleic Acids Research* **21,** 5831-7.

Nürnberger, F. (1995). The neuroendocrine system in hibernating mammals: present knowledge and open questions. *Cell Tissue Research* **281,** 391-412.

Ohkuma, Y., Horikoshi, M., Roeder, R. & Desplan, C. (1990). Engrailed, a homeodomain protein, can repress in vitro transcription by competition with the TATA box-binding protein transcription factor IID. *Proceedings of the National Academy of Sciences of the United States of America* **87,** 2289-93.

Ohkuma, Y. & Roeder, R. (1994). Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature* **368,** 160-3.

Okayama, H. & Chen, C. (1991). Calcium phosphate mediated gene transfer into established cell lines. In Gene Transfer and Expression Protocols, E. J. Murray, ed. (Clifton, New Jersey: Human Press), pp. 15-21.

Ossipow, V., Tassan, J., Nigg, E. & Schibler, U. (1995). A mammalian RNA polymerase II holoenzyme containing all components required for promoter-specific transcription initiation. *Cell* **83,** 137-46.

Pankratz, M. J., Seifer, E., Gerwin, N., Billi, B., Nauber, U. & Jäckel, H. (1990). Gradients of *Krüppel* and *knirps* gene products direct pair-rule gene stripe patterning in the posterior region of the *Drosophila* embryo. *Cell* **61,** 309-17.

Parvin, J. D. & Sharp, P. A. (1993). DNA topology and a minimal set of basal factors for transcription by RNA polymerase II. *Cell* **73,** 533-40.

Pognonec, P., Kato, H. & Roeder, R. G. (1992). The helix-loop-helix/leucine repeat transcription factor USF can be functionally regulated in a redox-dependent manner. *Journal of Biological Chemistry* **267,** 24563-7.

Poon, D., Campbell, A. A., Bai, Y. & Weil, P. A. (1994). Yeast taf170 is encoded by *MOT1* and exists in a TATA box binding protein(TBB)-TBP-associated complex distinct from transcription factor IID. *Journal of Biological Chemistry* **269,** 23135-40.

Price, J. F., Gourlie, B. B., Lin, Y. & Huang, R. C. C. (1986). Induction of winter flounder messenger RNA at 4 °C *in vivo* and *in vitro*. *Physiol. Zool.* **59, 689-95**.

Ptashne, M. (1986). A Genetic Switch Gene Control and Phage λ (Palo Alto: Cell press & Blackwell Scientific Publications).

Rand-Weaver, M., Pottinger, T. & Sumpter, J. (1995). Pronounced seasonal rhythms in plasma somatolactin levels in rainbow trout. *Journal of Endocrinology* **146,** 113-9.

Ranish, J. A., Lane, W. S. & Hahn, S. (1992). Isolation of two genes that encode subunits of the yeast transcription factor IIA. *Science* **255,** 1127-9.

Reese, J. C., Apone, L., Walker, S. S., Griffin, L. A. & Green, M. R. (1994). Yeast TAFIIS in a multisubunit complex required for activated transcription. *Nature* **371,** 523-7.

Reiter, R. (1993). The melatonin rhythm: both a clock and a calendar. *Experientia* **49,** 654-64.

Renkawitz, R. (1990). Transcriptional repression in eukaryotes. *Trends in Biochemical Scincess* **6,** 192-7.

Renkawitz, R. (1993). Repression mechanisms of v-ERBA and other members of the steroid receptor superfamily. *Annals of the New York Academy of Sciences* **684,** 1-10.

Roberts, S. G. E., Ha, I., Maldonado, E., Reinberg, D. & Green, M. R. (1993). Interaction between an acidic activator and transcription factor TFIIB is required for transcriptional activation. *Nature* **363,** 741-4.

Ron, D. & Habener, J. F. (1992). CHOP, a novel developmentally regulated nuclear protein that dimerizes with transcription factors C/EBP and LAP and functions as a dominant-negative inhibitor of gene transcription. *Genes & Development* **6,** 439-53.

Rosenfeld, M. G. (1991). POU-domain transcription factors: pou-er-ful developmental regulators. *Genes and Development* **5,** 897-907.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). Molecular Cloning, A Laboratory Manual (New York: Cold Spring Harbor Press).

Santoro, C., Mermod, N., Andrews, P. & 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.

Sassone-Corsi, P. & Borrelli, E. (1986). Transcriptional regulation by *trans*-acting factor. *Trends in Genetics*, 215-9.

Sauer, F., Fondell, J. D., Ohkuma, Y., Roeder, R. G. & Jäckle, H. (1995). Control of transcription by Krüppel through interactions with TFIIB and TFIIEβ. *Nature* **375,** 162-4.

Sauer, F. & Jäckel, H. (1993). Dimerization and the control of transcription by *Krüppel*. *Nature* **364,** 454-7.

Sauer, F. & Jäckle, H. (1991). Concentration-dependent transcriptional activation or repression by Krüppel from a single binding site. *Nature* **353,** 563-6.

Schimke, R. T. (1984). Gene amplification in cultured animal cells. *Cell* **37,** 705-13.

Schmid, W., Strähle, U., Schütz, G., Schmitt, J. & Stunnenberg, H. (1989). Glucocorticoid receptor binds cooperatively to adjacent recognition sites. *EMBO Journal* **8,** 2257-63.

Schmid, W. D. (1982). Survival of frogs at low temperatures. *Science* **215,** 697-8.

Scholander, P. F., Van Dam, L., Kanwisher, J. W., Hammel, H. T. & Gordon, M. S. (1957). Supercooling and osmoregulation in Arctic fish. *Journal of Cell Comp* **49,** 5-24.

Schule, R. (1990). Jun-Fos and receptors for vitamin A and D recognize a common response element in the human osteocalsin gene. *Cell* **61,** 497-504.

Scott, G. K., Fletcher, G. L. & Davies, P. L. (1986). Fish antifreeze proteins: Recent gene evolution. *Canadian Journal of Fisheries and Aquatic Sciences* **43,** 1028-34.

Scott, G. K., Hayes, P. H., Fletcher, G. L. & Davies, P. L. (1988). Wolffish antifreeze protein genes are primarily organized as tandem repeats that each contain two genes in inverted orientation. *Molecular & Cellular Biology* **8,** 3670-5.

Segall, J., Matsui, T. & Roeder, R. (1986). Multiple factors are required for the accurate transcription of purified genes by RNA polymerase III. *Journal of Biological Chemistry* **255,** 11986-91.

Sheets, M. D., Ogg, S. C. & Wickens, M. P. (1990). Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation *in vitro*. *Nucleic Acids Research* **18,** 5799-805.

Sicheri, F. & Yang, D. (1995). Ice-binding structure and mechanism of an antifreeze protein from winter flounder. *Nature* **375,** 427-31.

Sierra, F. (1990). A Laboratory Guide to *In Vitro* Transcription. In Biomethods, A. Azzi, J. M. Polak and H. P. Saluz, eds. (Basel: Birkhäuser Verlag).

Silver, S. C. & Humt, S. W. (1993). Techniques for cloning cDNAs encoding interactive transcriptional regulatory proteins. *Molecular Biology Reports* **17,** 155-65.

Singh, H., LeBowitz, J., Baldwin, A. S. & Sharp, P. A. (1988). Molecular cloning of an enhancer binding protein: Isolation by screening of an expression library with recognition site DNA. *cell* **52,** 415-23.

Slaughter, D., Fletcher, G. L., Ananthanarayanan, V. S. & Hew, C. L. (1981). Antifreeze proteins from the sea raven, *Hemitripterus americanus*. Further evidence for diversity among fish polypeptide antifreezes. *Journal of Biological Chemistry* **256,** 2022-6.

Sonnichsen, F. D., Sykes, B. D., Chao, H. & Davies, P. L. (1993). The nonhelical structure of antifreeze protein type III. *Science* **259,** 1154-7.

Sonnichsen, F. D., Sykes, B. D. & Davies, P. L. (1995). Comparative modeling of the three-dimensional structure of type II antifreeze protein. *Protein Science* **4,** 460-71.

Stargell, L. & Struhl, K. (1995). The TBP-TFIIA interaction in the response to acidic activators in vivo. *Science* **269,** 75-8.

Stelzer, G., Goppelt, A., Lottspeich, F. & Meisteremst, M. (1994). Repression of basal transcription by HMG2 is counteracted by TFIIH-associated factors in an ATP-dependent process. *Molecular & Cellular Biology* **14,** 4712-21.

Stoscheck, C. M. (1990). Quantitation of protein. In Guide to Protein Purification. M. P. Deutzcher, ed. (San Diego: Academic Press Inc.), pp. 50-68.

Szabo, C. & Blanche, M. (1995). Seasonal variation in mood disorder presentation: further evidence of this phenomenon in a South African sample. *Journal of Affective Disorders* **33,** 209-14.

Sømme, L. (1982). Supercooling and winter survival in the terrestial arthorpods. *Comparative Biochemical Physiology* **73A,** 519-43.

TenHarmsel, A., Austin, R., Savenelli, N. & Biggin, M. (1993). Cooperative binding at a distance by even-skipped protein correlates with repression and suggests a mechanism of silencing. *Molecular & Cellular Biology* **13,** 2742-52.

Urrutia, M. E., Duman, J. G. & Knight, C. A. (1992). Plant thermal hysteresis protein. *Biochimica et Biophyisica Acta* **1121,** 199-206.

Urzainqui, A. & Carrasco, L. (1989). Degradation of cellular proteins during poliovirus infection: studies by two-dimensional gel electrophoresis. *Journal of Virology* **63,** 4729-35.

Usheva, A., Maldonado, E., Goldring, A., Lu, H., Houbavi, C., Reinberg, D. & Aloni, Y. (1992). Specific interaction between the nonphosphorylated form of RNA polymerase II and the TATA-binding protein. *Cell* **69,** 871-81.

Vaisius, A., Martin-Kearley, J. & Fletcher, G. L. (1989). Antifreeze protein gene transcription in winter flounder is not responsive to temperature. *Cellular & Molecular Biology* **35,** 547-54.

Van Voorhies, W. V., Raymond, J. A. & DeVries, A. L. (1978). Glycoproteins as a biological antifreeze agents in the cod *Gadus ogac* (Richardson). *Physiological Zoology* **51,** 347-53.

Verma, I. M., Stevenson, J. K., Schwarz, E. M., Van, A. D. & Miyamoto, S. (1995). Rel/NF-κB/I κB family: intimate tales of association and dissociation. *Genes & Development* **9,** 2723-35.

Visvader, J. & Begley, G. (1991). Helix-loop-helix genes translocated in lymphoid leukemia. *Trends in Biochemical Sciences* **16,** 330-3.

Walker, S. S., Reese, J. C., Apone, L. M. & Green, M. R. (1996). Transcription activation in cells lacking TAFIIs. *Nature* **383,** 185-8.

Wang, X., DeVries, A. L. & Cheng, C. H. (1995a). Antifreeze peptide heterogeneity in an antarctic eel pout includes an unusually large major variant comprised of two 7 kDa type III AFPs linked in tandem. *Biochimica et Biophysica Acta* **1247,** 163-72.

Wang, X., DeVries, A. L. & Cheng, C. H. (1995b). Genomic basis for antifreeze peptide heterogeneity and abundance in an Antarctic eel pout: gene structures and organization. *Molecular Marine Biology & Biotechnology* **4,** 135-47.

Wang, Z. Y., Qiu, Q. Q. & Deuel, T. F. (1993). The Wilms' tumor gene product WT1 activates or suppresses transcription through separate functional domains. *Journal of Biological Chemistry* **268,** 9172-5.

Wasserman, S. A. (1993). A conserved signal transduction pathway regulating the activity of the rel-like proteins dorsal and NF-κ B. *Molecular Biology of the Cell* **4,** 767-71.

Watt, F. & Molly, P. L. (1993). Specific cleavage of transcription factors by thiol protease, m-calpein. *Nucleic Acid Research* **21,** 5092-100.

Weinzierl, R. O., Dynlacht, B. D. & Tjian, R. (1993). Largest subunit of *Drosophila* transcription factor IID directs assembly of a complex containing TBP and a coactivator. *Nature* **362,** 511-7.

White, R. J., Khoo, B. C., Inostroza, J. A., Reinberg, D. & Jackson, S. P. (1994). Differential regulation of RNA polymerases I, II, and III by the TBP-binding repressor Dr1. *Science* **266,** 448-50.

Witt, I., Kwart, M., Groz, T. & Käufer, N. F. (1995). The tandem repeat AGGGTAGGGT is, in the fission yeast, a proximal activation sequence and activates basal transcription mediated by the sequence TGTGACTG. *Nucleic Acid Research* **23,** 4296-302.

Witt, I., Strauz, N., Käufer, N. F. & Gross, T. (1993). The CAGTCACA box in the fission yeast *Schizosaccharomyces pombe* functions like a TATA element and binds a novel factor. *EMBO Journal* **12,** 1201-08.

Wu, J., Jupp, R., Stenberg, R., Nelson, J. & Ghazal, P. (1993). Site-specific inhibition of RNA polymerase II preinitiation complex assembly by human cytomegalovirus IE86 protein. *Journal of Virology* **67,** 7547-55.

Wulczyn, F. G., Naumann, M. & Scheidereit, C. (1992). Candidate proto-oncogene *bcl*-3 encodes a subunit-specific inhibitor of transcription factor NF-κB. *Nature* **358,** 597-9.

Xanthoudakis, S. & Curran, T. (1992). Identification and characterization of Ref-1, a nuclear protein that facilitates AP-1 DNA-binding activity. *EMBO Journal* **11,** 653-65.

Xanthoudakis, S., Miao, G., Wang, F., Pan, Y. C. & Curran, T. (1992). Redox activation of Fos-Jun DNA binding activity is mediated by a DNA repair enzyme. *EMBO Journal* **11,** 3323-35.

Yamamoto, T., Horikoshi, M., Wang, J., Hasegawa, S., Weil, P. A. & Roeder, R. G. (1992). A bipartite DNA binding domain composed of direct repeats in the TATA box binding factor TFIID. *Proceedings of the National Academy of Sciences of the United States of America* **89,** 2844-8.

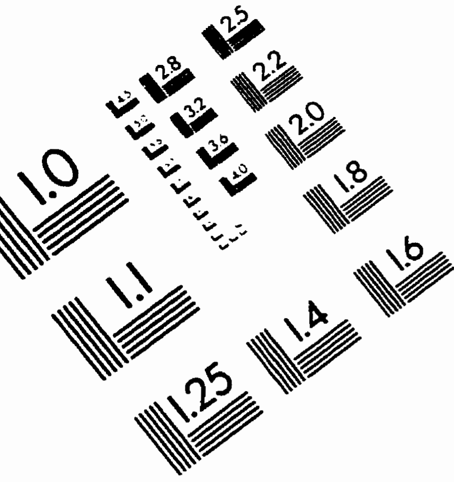
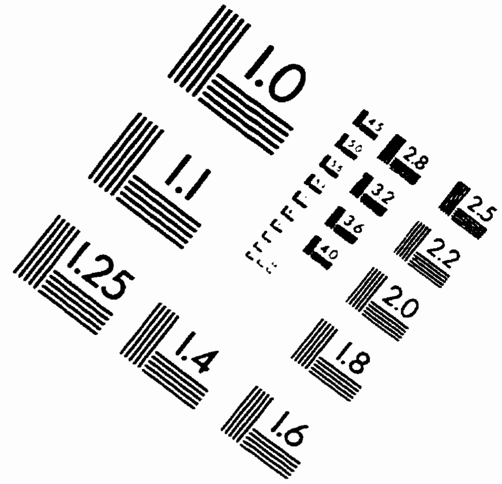
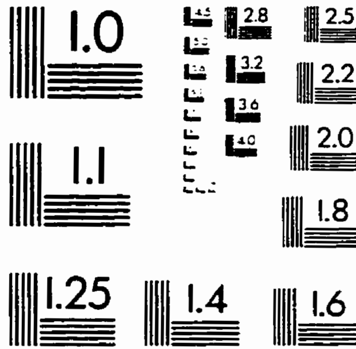
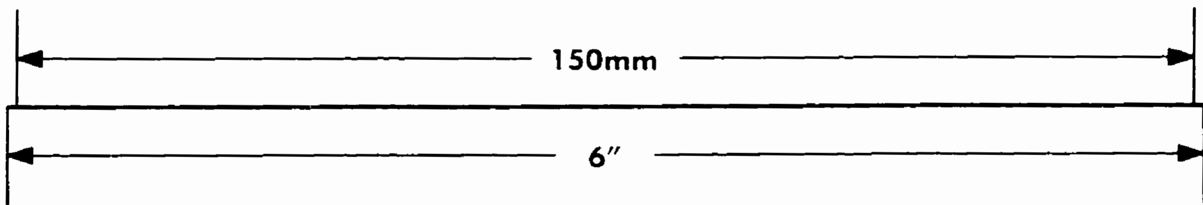
Yang, D. S. C., Sax, M., Chakrabartty, A. & Hew, C. L. (1988). Crystal structure of an antifreeze polypeptide and its mechanistic implications. *Nature* **333,** 232-7.

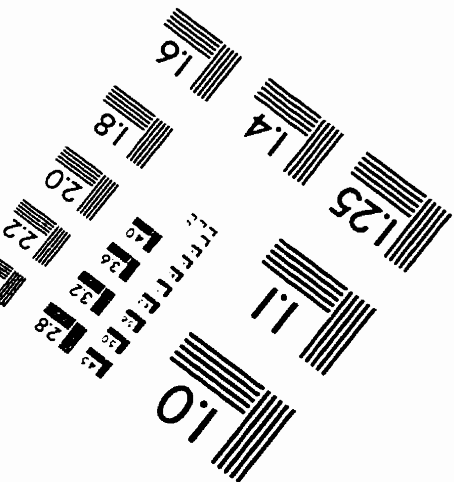
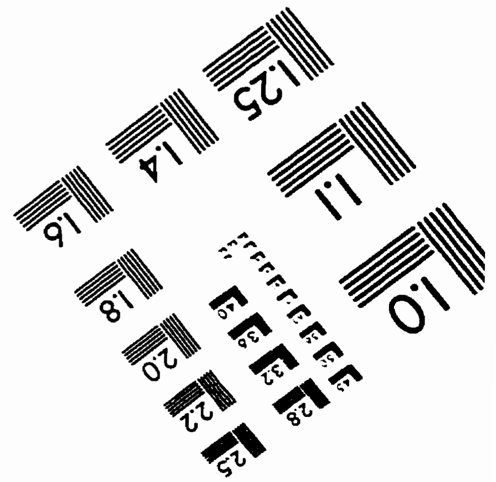
Yeung, K. C., Inostroza, J. A., Mermelstein, F. H., Kannabiran, C. & Reinberg, D. (1994). Structure-function analysis of the TBP-binding protein Dr1 reveals a mechanism for repression of class II gene transcription. *Genes & Development* **8,** 2097-109.

Young, R. A. (1991). RNA Polymerase II. *Annual Review of Biochemistry* **60,** 689-715.

Zhang, W. & Yaneva, M. (1993). Reduced sulphydryl groups are required for DNA binding of Ku protein. *Biochemical Journal* **293**, 769-74.

IMAGE EVALUATION TEST TARGET (QA-3)

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved