THE FUNCTIONAL ANALYSIS OF THE OCEAN POUT (Macrozoarces americanus) TYPE III ANTIFREEZE PROTEIN GENE PROMOTER

TRINA MAXINE KIRBY

The functional analysis of the ocean pout (*Macrozoarces americanus*) type III antifreeze protein gene promoter

by

Trina Maxine Kirby, BSc (hon)

A thesis submitted to the Department of Biology in partial fulfillment of the requirements for the degree of Master of Science

Department of Biology

Memorial University of Newfoundland

July, 2005

St. John's, Newfoundland and Labrador

Abstract

The full length ocean pout antifreeze gene promoter and a series of promoter deletion constructs were fused to the luciferase reporter gene, and their transcriptional role was examined in a transient expression assay by transfection into three cell lines; a human epithelial (CRL-1578), a human fibroblast (HTB-92) and a Chinook salmon embryonic cell line (CRL-1691). The studies demonstrate that the ocean pout AFP gene promoter can function in both mammalian and fish cell lines and that they contain both positive as well as negative regulatory regions. This indicates that the AFP gene may be regulated at various sites.

In previous studies, a fish antifreeze/growth hormone chimeric gene construct opAFP-GHc2 was injected into Atlantic salmon eggs. Through PCR analysis it was found that this construct reorganized in the fish genome so that the beginning of the 5 ' ocean pout AFP promoter end was linked to the 3' end of the AFP polyadenylation sequence. In my study, this reorganized transgenic promoter was also ligated to the luciferase reporter gene and examined in a transient transfection assay, which determined that this reorganized promoter had significantly less activity than that of the full-length intact promoter. However, promoter activity was detected in all three cell lines which indicates that the transgene should be expected to express growth hormone *in vivo*.

To my parents

Acknowledgements

I would like to thank my supervisor, Dr. Garth Fletcher, for his guidance and support throughout this project. I would also like to thank all of the staff at Aqua bounty Canada for their help and friendship over the last two years. Particularly I would like to acknowledge Dr. Edward Yaskowiak for his expert advice on my project and Jim Houstan for his technical advice and assistance with my project. Also, to Dr Yogesh Mawal, Dr. Alka Mawal, and Madonna King who provided me with training in both molecular biology and cell culture techniques necessary to my study. I would also like to thank Dr. Brian Staveley at Memorial University of Newfoundland for his help and advice throughout my project, it was greatly appreciated.

Finally, I would like to acknowledge the financial support of AquaNet.

Table of Contents

List of figures

List of abbreviations

A	adenine
AFGP	antifreeze glycoprotein
AFP	antifreeze protein
Amp	ampicillian
Amp^r	ampicillian resistance
ATCC	American type culture collection
bp	base pair(s)
C	cytosine
CCLR	cell culture lysis reagent
cDNA	complementary DNA
DMEM	Dulbecco's minimum eagles medium
DPE	dpwnstream promoter elements
DTT	dithiothreitol
EDTA	ethylenediaminetetra-acetic acid
EtBr	ethidium bromide
EtOH	ethanol
FBS	fetal bovine serum
G	guanine
GH	growth hormone
HBS	Hank's balanced salt solution
Inr	cap site/initiator

kb	kilobase
L-15	Leibovitz's medium
LAR	luciferase assay reagent
LB	Luria broth
luc	luciferase
M	molar
mRNA	messenger RNA
MEM	mimimal eagles medium
min	minutes
op5a AFP	ocean pout 5a antifreeze protein
opAFPGHc(2)	ocean pout antifreeze protein growth hormone construct (two)
PAGE	polyacrilomide gel electrophoresis
PCR	polymerase chain reaction
RNA	ribonucleic acid
SDS	sodium dodecyl sulphate
T	thymine
TATA	Goldberg-Hogness
TBE	100 mM Tris-HCL, 100 mM borate, 2 mM EDTA
Tris	Tris-(hydroxymethyl) aminomethane
UTR	untranslated region
w/v	weight/volume

Introduction

I. Transcriptional control of gene expression

The regulation of gene expression in eukaryotes is essential for the control of many biological processes, including growth, differentiation, development, and response to environmental conditions. Gene expression is primarily regulated at the transcriptional level, however, it can be regulated at other levels including the post-transcriptional, translational, and posttranslational levels (Wu, 1994). The transcription of a eukaryotic protein coding gene is preceded by a number of actions including the decondensation of the locus, nucleosome remodelling, histone modification, binding of transcriptional activators and coactivators to promoters and enhancers, and recruitment of the basal transcription machinery to the core promoter (Smale and Kadonaga, 2003). Transcription by DNA-dependent RNA polymerases is a recurring process composed of four steps: 1) promoter binding and activation, 2) RNA chain initiation and promoter escape, 3) RNA transcript elongation, and 4) RNA transcript termination and release (Uptain *et al.*, 1997).

Transcription of a gene is influenced by *cis*-acting elements and *trans*-acting proteins. The *cis*-acting elements promoters, enhancers, and silencers, are composed of discrete DNA motifs, which serve as binding sites for the *trans*-acting factors. The *trans*-acting factors are sequence specific DNA- binding proteins, which control the location of transcription initiation. The interaction of these *cis*-and *trans*-acting elements control the transcription of a eukaryotic gene by enabling the gene to be expressed in a stage specific manner and/or making the gene responsive to external trigger signals (Gong *et al.*, 1991;

Chan *et al.*, 1993). There have been various studies preformed focusing on the identification of various *trans*-acting factors that help carry out the transcription process and the characterization of the *cis*-acting elements, such as promoter sequences, in recent years has gained much attention.

1. *Cis*-acting transcriptional regulatory regions

There are generally three types of *cis*-acting elements that control the activity of protein coding genes (Figure 1). The first element is the basal or core promoter sequence, which is the focal point of eukaryotic transcriptional regulation (Lewis *et al.*, 2000). This is roughly the 100 base pairs (bp) found directly upstream from the start site of transcription, which contains the binding site for both the RNA polymerase II and the basal transcriptional machinery that acts on most promoters. This sequence thus determines the location of transcriptional initiation (Conaway and Conaway, 1993; Arnosti, 2003). The second element is the enhancer, which can be located upstream, downstream or within the transcriptional unit. The third element is the boundary or insulator elements, which are regulatory regions that are usually several hundred base pairs in length.

A. The basal elements

The basal promoter region contains the Goldberg-Hogness box or TATA box (consensus TATAAA), which is an A + T rich sequence usually located about 25 to 30

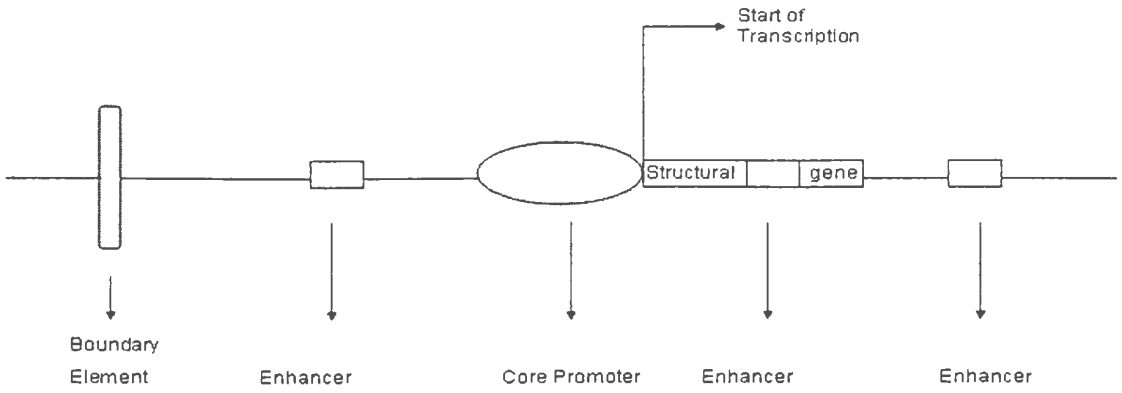


Figure 1. The *Cis*-acting transcriptional control elements. The core promoter element includes approximately 100 bp of sequence found directly upstream of the start site of transcription. The enhancer elements can be located upstream, downstream or within the structural gene unit. The boundary elements are usually several hundred base pairs in length and are found near some genes (adapted and modified from Smale and Kadonaga, 2003).

bp upstream from the start of transcription in higher eukaryotes (Zawel and Reinberg, 1995; Smale and Kadonaga, 2003). It is thought that the TATA box has many crucial and related activities. First, it nucleates the formation of a pre-initiation complex containing RNA polymerase and several general transcription factors. Second, its placement determines the location of initiation of RNA synthesis. It also aides in determining the direction of transcription in many promoters and finally, it is adequate for allowing activator proteins that are bound to the promoter's regulatory elements to stimulate high levels of specifically initiated transcription (Breathnach and Chambon, 1981; Zawel and Reinberg, 1992; Smale, 1994). Several studies have shown that all of these activities result from a transcription protein factor, TFIID, which recognizes and binds to the TATA box. This binding induces the binding of other factors and initiates a cascade of events that is essential for the formation of the stable preinitiation complex, the key event preceding the start of transcription.

Many steps in this cascade have been discovered and it is thought to require the action of at least five initiation factors and ATP (Figure 2). The binding of TFIID to the TATA box is assisted and stabilized by TFIIA (Samuels and Sharp, 1986; Moncollin *et al.*, 1994). There is also evidence linking a weak TFIID interaction to weak promoter strength, which is independent of the presence or absence of upstream promoter motifs (Nakajima *et al.*, 1988). The binding of TFIID and TFIIA is required for the selective binding of RNA polymerase II to the complex. At this stage, the complex is

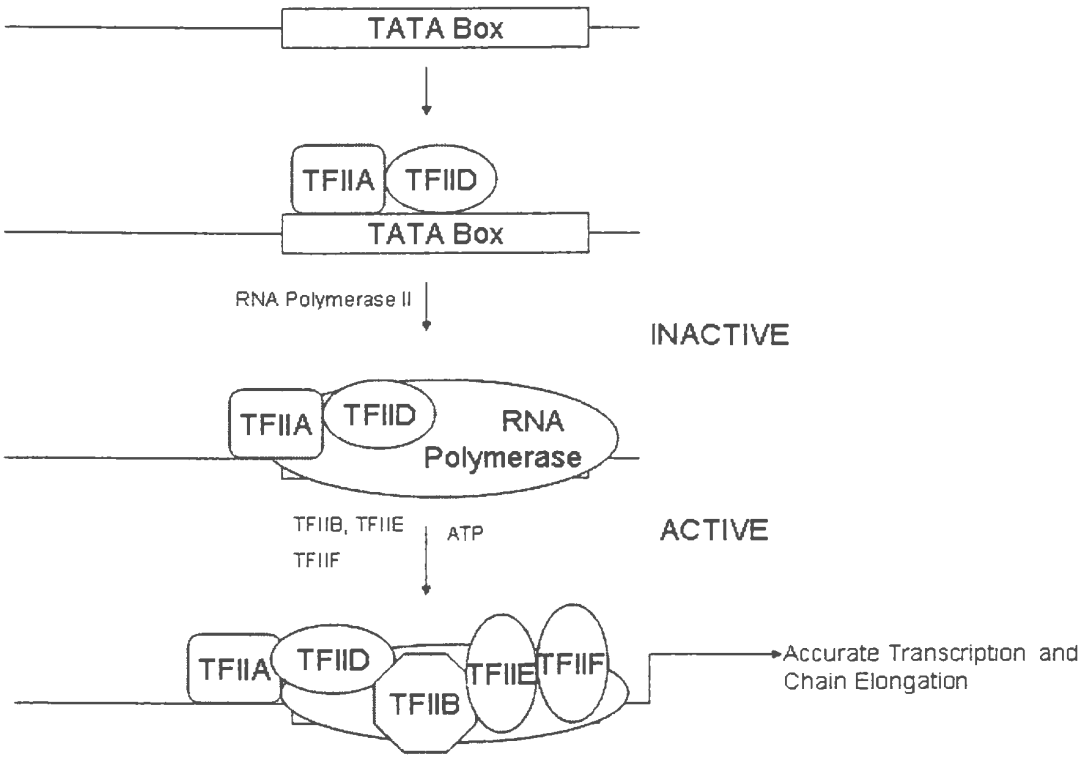


Figure 2. The cascade of events leading to the formation of the pre-initiation complex. This model shows the ordered events of general transcription factors into the pre-initiation complex. TFIID recognizes the TATA box with the assistance of TFIIA, and induces the binding of other factors. RNA can then assemble to the complex and in the presence of ATP the complex coverts to an active form. TFIIB, TFIIE and TFIIF are needed for accurate transcription and chain elongation (adapted and modified from Wu, 1994).

inactive, but in the presence of ATP, it is transformed into an active state. Another transcription factor termed TFIIB is required for accurate transcription while TFIIE and TFIIF, are required for transcription initiation and chain elongation (Conaway and Conaway, 1989; Flores *et al.*, 1989; Smale and Kadonaga, 2003; Sawadogo, 1990). Many genes lack an obvious TATA box, and it has been proposed that the formation of a pre-initiation complex with promoters lacking the TATA box may occur by an alternative pathway involving other sequences that replace the TATA box. However, studies in higher eukaryotic promoters which contain a TATA box have shown that a deletion in the TATA box results in a severe reduction or inactivation of promoter activity and thus gene transcription (Kraus *et al.*, 1995), indicating that when present the TATA box is essential in promoter activity.

A second sequence contained in the basal promoter found in higher eukaryotes is the 5′-CCAAT-3` binding motif (Ozer *et al.*, 1990), which is often found in a position 50 - 120 bp upstream from the start of transcription (Weimin *et al.*, 1997). It can operate independent of orientation (McKnight and Tjian, 1986), however deletion or mutational analysis in this motif have been shown to detrimentally affect levels of transcription from a variety of promoters (Ozer *et al.*, 1990). This indicates that the integrity of the CCAAT box is critical for optimal promoter strength (Hatamochi *et al.*, 1988). However, some promoters do not have a CCAAT motif and are still able to function efficiently (McKnight, 1982). Several CCAAT binding proteins have been isolated and

characterized such as CTF/NF1, c/EBP, YB-1, and CBF (also termed NF-Y and CP1) (Ozer *et al.*, 1990; Bi *et al.*, 1997) and it is this binding that may help stimulate gene transcription.

A third sequence that may be located in the basal promoter area of cellular and viral genes is the GC-rich motif (consensus GGGCCGG). The GC box may also be found in the distal promoter region and function in a similar way to an enhancer site (Courey *et al.*, 1989; Song *et al.*, 2001). A *trans*-element, termed the Sp1 factor recognizes and binds to the sequence. Sp1 has three carboxyl-terminal zinc finger motifs, which are used to facilitate the binding to the DNA and activate transcription by glutamine- rich domains. This selective binding activates mRNA synthesis from the gene, and aids in controlling transcription. The binding usually creates positive effects on promoter activity however a negative effect has also been documented (Kadonaga, 1987; De Clercq *et al.*, 1992).

A fourth sequence found in the basal promoter area is the cap-site or initiator (Inr) sequence (consensus TCAGT) and it is found centered around the transcription start site (Arnosti, 2003). This element is functionally similar to, but can function independently of the TATA box (Smale and Kadonaga, 2003). The Inr was originally described in a TATA-less promoter, where it established correct transcription initiation. However a variety of promoters have been found to contain both the Inr and the TATA box. The Inr is thought to play a direct role in promoter recognition and initiation by RNA polymerase. It has been reported that RNA polymerase can bind weakly to the Inr and initiate transcription (Kraus, 1995). In various deletion studies of model organisms such

as Drosophila, deletion of the Inr resulted in a decrease in transcription of various genes (Arkhipova, 1995). Although its role remains to be fully defined, it does contribute to the magnitude of transcription.

Sequences that are located downstream of the transcription start site of protein coding genes can also affect gene expression. The downstream promoter elements (DPE-consensus A/GGA/TC/TGT) (Arnosti, 2003) is located precisely 10 to 32 bp downstream relative to the Inr motif. This element has been found in both a TATA and TATA- less promoter, however in order to function in transcriptional control, this element requires the presence of an Inr. The DPE is thought to be involved in basal transcription by providing sites of interaction for the TFIID complex. This increases the TFIID complex formation and stability (He, 2002; Smale and Kadonaga, 2003).

B. The enhancer elements

The next element which will be discussed briefly is the enhancer sequences. Enhancers were originally found in viral genomes, however they are also known to regulate a variety of cellular genes (Chen *et al.* 1986; Ostapchuk *et al.*, 1986). These elements contain multiple binding sites for sequence-specific transcription activators and repressors which regulate the rate at which RNA polymerase II initiates new rounds of transcription from the core promoter, thus regulating levels of gene activity (Conaway and Conaway, 1993; Arnosti, 2003). Various enhancer sequences have been shown to display tissue and species specificity (Ostapchuk *et al.*, 1986). Enhancers act independent of orientation and location, and can be located both within a functional gene,

upstream from the start of transcription or downstream from the start of transcription. Enhancer sequences have been shown to effect transcription as far as 3000 bp away from the transcription start site (Latchman, 1998). Any functional cluster of transcription factor binding sites are usually considered as an enhancer regardless of the distance to the initiation site.

C. The boundary elements

The last element, which will be discussed, is the boundary elements or insulators. The eukaryotic genome is partitioned into independent functional domains which are separated by DNA regulatory sequences termed boundary or insulator elements (Ishihara and Sasaki, 2002). Insulators possess the capability to protect genes from inappropriate signals from their environment, which ensures correct gene expression. If the insulator is situated between the enhancer and the promoter, they act by blocking the action of the distal enhancer on a promoter. This would inhibit the enhancer from activating expression of an adjacent gene from which it is blocked, while leaving it free to stimulate expression of genes located on its unblocked side. Also, they may protect genes by acting as a barrier that prevents the advance of nearby condensed chromatin that might otherwise silence expression (Conaway and Conaway, 1993; West *et al.*, 2002).

All of these elements are interrelated and contribute to gene regulation. The term 'promoter' is often, and will be for this paper, used to indicate the DNA regulatory sequences that are required for accurate and efficient initiation and regulation of

transcription, including the core promoter, enhancer elements, and insulator or boundary sequences.

II. The functional analysis of a promoter

The analysis of a gene promoter will give valuable information on how a gene is regulated *in vivo*. The general transcriptional machinery is thought to be conserved in eukaryotes, however it is known that a promoter may function more efficiently in a homologous system as opposed to a heterologous system (Hildegard and Schart, 1990). Promoter regions from organisms such as mammals, *Drosophila,* and yeasts have been extensively characterized. Until recently, little has been known about fish promoters since only a few fish gene promoters had been studied. Among the initial fish promoters studied were promoters for the rainbow trout (*Oncorhynchus mykiss*) protamine (Dillon *et al.*, 1985; Jankowski and Dixon, 1987) and methallothionein β genes (Zafarullah *et al.*, 1988), the β- actin gene from carp (*Cyprinus sp.*) (Zhanjiang *et al.*, 1990) and several antifreeze protein (AFP) promoters from a number of cold water fish (Gong *et al.*, 1991, Gong and Hew, 1993). Recently, other fish promoters have been characterized such as the tunor necrosis factor associated factor 2 binding protein gene promoter from the grass carp (*Ctenopharyngodon idellus)* (Chang *et al,* 2005) the insulin-like growth factor binding protein-2 proximal promoter region of the zebrafish (*Danio rerio)* (Chen *et al.*, 2005), and the follicle-stimulating hormone β subunit gene promoter of the Chinook salmon (Chong *et al.,* 2005). Analysis of the fish gene promoters yielded results similar

to those obtained for the other organisms in that they possessed regions that could influence gene expression and that the genes could be regulated at multiple sites.

1. Promoter analysis for transgenic studies

The successful generation of transgenic animals has opened up exciting possibilities for applied research. In addition, transgenic organisms have also become useful tools with which to study gene regulation and function. Since the mid-1980s when the first transgenic fish were successfully created, gene transfer studies have been conducted on at least 35 fish species, half of which are important in aquaculture (Zbikowska, 2003). Most of the research on species being developed for or used in, aquaculture has been designed to produce superior strains of fish with improved growth rates, increased freeze resistance and increased disease resistance (Hew *et al.*, 1998). Although the research on transgenic fish has had some success, particularly with regards to the production of rapidly growing fish using chimeric growth hormone gene constructs in Atlantic salmon (*Salmo salar*) (Hew *et al.*, 1998), pacific salmon (*Oncorhynchus* sp.), rainbow trout (*Oncorhynchus mykiss*), cutthroat trout (*Oncorhynchus clarki clarki*) (Devlin *et al.*, 1995a), channel cat-fish (*Ictalurus punctatus*) (Dunham, 1999), loach (*Botia* sp.) (Nam *et al.*, 2001) and Tilapia (*Oreochromis* sp.) (Martinez *et al.*,1996; Maclean *et al.*, 2002) it is evident that in some cases the transgenes appear to have been too powerful, resulting in morphological irregularities (Devlin *et al.*, 1995b), and in other cases transgenic expression was insufficient to elicit the desired phenotype. A good example of the latter is the low level of antifreeze protein expression observed in Atlantic salmon (*Salmo*

salar) transgenic for a type I antifreeze protein gene (Hew *et al.*, 1999). One way to correct or avoid these potential difficulties is to better understand the promoters so that they can be tailored to optimize transgene expression and produce the desired phenotype.

In the past few years many stable transgenic zebrafish lines have been generated that detect water pollutants by changing colour. Zebra fish are black and silver in colour, however through genetic manipulation a few varieties that radiate green or red fluorescent colour have been produced. The fluorescent genes are extracted from jellyfish and then injected into zebra fish eggs, to allow the zebra fish to give off a fluorescent glow. Inducible gene promoters are used to act as control switches to activate different tissues in these transgenic fish and to drive the fluorescent colour genes. The fluorescent coloured transgenic fish are able to respond to the presence of chemicals like oestrogen through the estrogenic promoter and heavy metals and toxins through the stress-responsive promoter. The fish will immediately display the colour depending on the type of environment the colour has been specified for. (Ju *et al.*, 1999).

III. The ocean pout antifreeze protein gene promoter

Antifreeze proteins (AFPs) are a unique class of serum proteins that are produced by a variety of organisms including teleost fish that are associated with freezing environments (Gong and Hew, 1993). In fish, these proteins lower the freezing point of the body fluids and prevent them from freezing when they encounter ice at sub-zero temperatures (Goddard and Fletcher, 2002). Five distinct classes of protein antifreeze, abbreviated collectively as AF(G)P, have been isolated and characterized to date: the

antifreeze glycoproteins (AFGP) and four types of antifreeze proteins (AFP) classified as type I, type II, type III and type IV (Fletcher *et al.*, 2001). The classification of these proteins is primarily based on the primary and secondary structures and on the order in which they were first isolated. Despite their diversity of structure all of the AF(G)P function by binding to developing ice crystals thereby inhibiting their growth (Davies *et al.*, 2002).

The ocean pout (*Macrozoarces americanus)* produces a family of type III AFP and secretes them into the blood throughout the year (Fletcher *et al.*, 1985). This family of AFP is encoded by a multigene family and expressed in most tissues with the liver exhibiting the highest level of expression (Gong et al., 1992; Hew *et al.*, 1998). One of the ocean pout AFP genes (OP5a; Hew *et al.*, 1988) was selected for use in a number of studies on transgenic salmon. Wu (1994) transferred an OP5a AFP gene with a truncated promoter to Atlantic salmon in an attempt to improve their freeze resistance. Although the transgene successfully integrated into the salmon genome the level of expression was insufficient to confer any significant improvement in freeze resistance. Du *et al* (1992) generated an all fish gene cassette that contained the complete transcriptional unit from the OP5a AFP gene except for its coding region. This construct contained the promoter region, the 5' and 3' untranslated regions including the polyadenylation signal AATAAA, and the 3' flanking region comprising the transcriptional termination signal TTTTTCT. This cassette, which was designed to facilitate a wide variety of candidate genes for gene transfer, has been used in the production of rapidly growing Atlantic and pacific salmon, and rapidly growing tilapia and loach. All were transgenic for chimeric

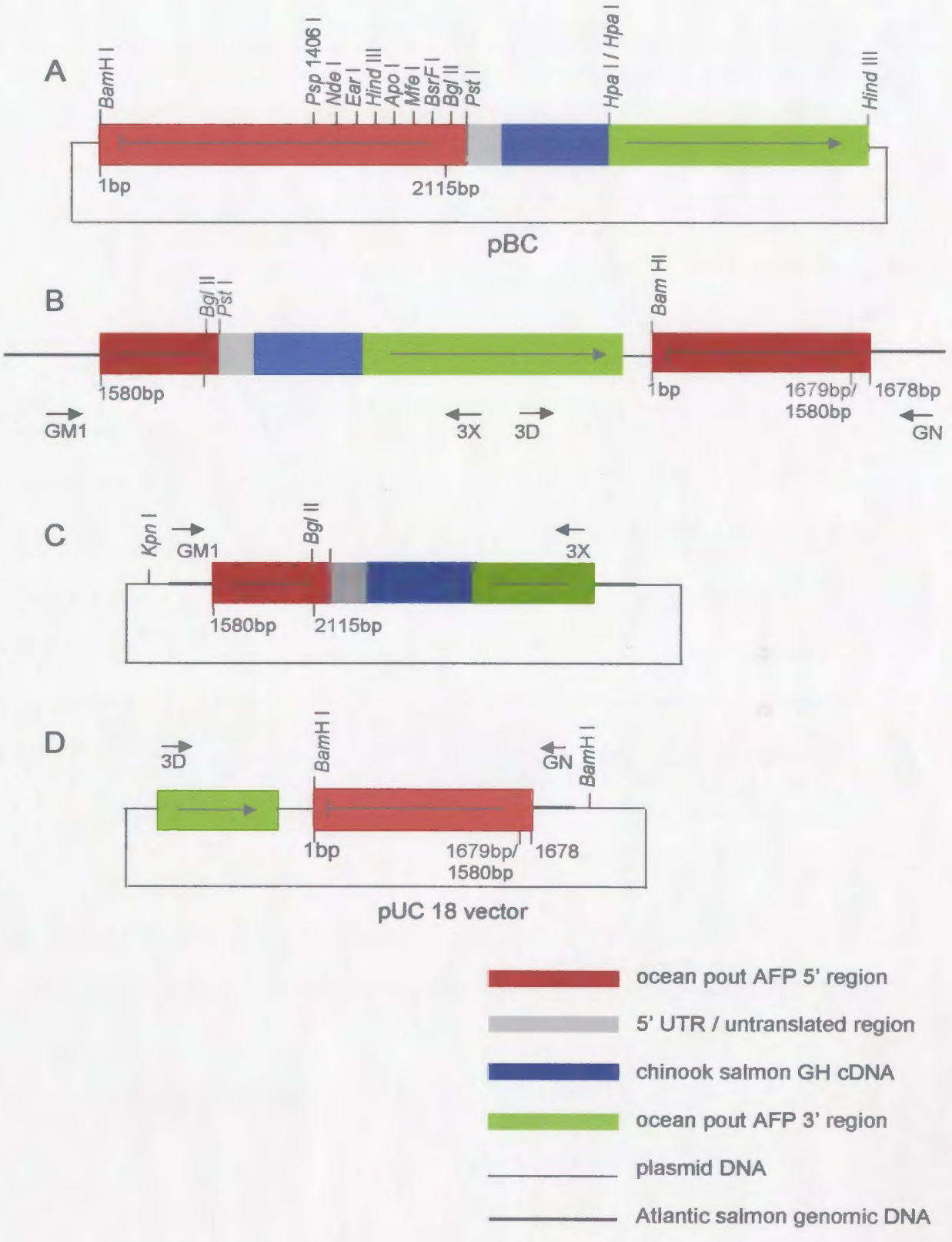
growth hormone gene constructs prepared using this casette (Devlin *et al.*, 1995a, Hew *et al.*, 1998; Nam *et al.*, 2001, Maclean *et al.*, 2002).

A detailed study of growth hormone (GH) transgenic Atlantic salmon containing one of Du's constructs, opAFP-GHc2 (Fig 3a), revealed that the genomically integrated GH transgene (EO-1α) was reorganized so that nucleotides 1580 bp to 2115 bp of the promoter remained upstream of the growth hormone gene. The first 1679 bp of the ocean pout AFP gene promoter plus a second copy of base pairs 1580 to 1678 reorganized downstream of the ocean pout 3' flanking region (personal communication, Yaskowiak and Fletcher). This reorganized structure raised a number of interesting and important questions, namely: 1) is the 5' truncated promoter able to drive expression of the GH transgene, and 2) does the relocation of the first 1580 bp 5' region downstream of the 3' flanking region influence this expression. This information along with a close examination of the OP5a AFP promoter sequence prompted a detailed examination of the promoter by taking advantage of the pGL3 family of luciferase reporter vectors.

IV. Research goals

For this thesis I addressed the two following questions; 1) Does the ocean pout AFP gene promoter contains regions that can positively and negatively regulate transcription and, if it does, where are they located in the promoter sequence? and 2) Is the 5' truncated promoter that is found in the transgenic salmon sufficiently functional to drive expression of the GH transgene, and does the relocation of the first 1580 bp 5' region downstream of the 3' flanking region influence the expression?

Figure 3. The ocean pout antifreeze promoter-chinook salmon growth hormone chimeric gene. A. Representation of the intact chimeric gene constructs opAFP-GHc or opAFP-GHc2. These constructs have the identical 2115 bp promoter sequence. The difference between these two constructs is the *Bgl* II to *Pst* I site which contains a 5' UTR site in opAFP-GHc2 and chinook salmon untranslated sequence in opAFP-GHc. B. The rearranged gene construct found integrated into the genome of the transgenic Atlantic salmon termed E0-1α. Nucleotides 1580 bp to 2115 bp (*Bgl* II site) of the promoter are present upstream of the growth hormone gene, while the first 1679 bp of the ocean pout AFP gene promoter plus a second copy of base pairs 1580 to 1678 is present downstream of the ocean pout 3' flanking region C. The promoter construct termed GM1-3X that contains the rearranged 5' promoter (1580 bp to 2115 bp). D. The promoter construct termed 3D-GN that harbours the rearranged promoter found downstream of the growth hormone gene (1 bp to 1678 bp).


A
BamH I
Psp 1406 I
Nde I
Ear I
Hind III
Apo I
Mfe I
BsrF I
Bgl II
Pst I
Hpa I / Hpa I
Hind III
1bp
2115bp
pBC
B
Bgl II
Pst I
Bam HI
1580bp
1bp
1679bp/
1580bp
1678bp
GM1
3X
3D
GN
C
Kpn I
GM1
Bgl II
3X
1580bp
2115bp
D
3D
BamH I
GN
BamH I
1bp
1679bp/
1580bp
1678
pUC 18 vector
ocean pout AFP 5' region
5' UTR / untranslated region
chinook salmon GH cDNA
ocean pout AFP 3' region
plasmid DNA
Atlantic salmon genomic DNA

Materials and Methods

1. Plasmids and constructs

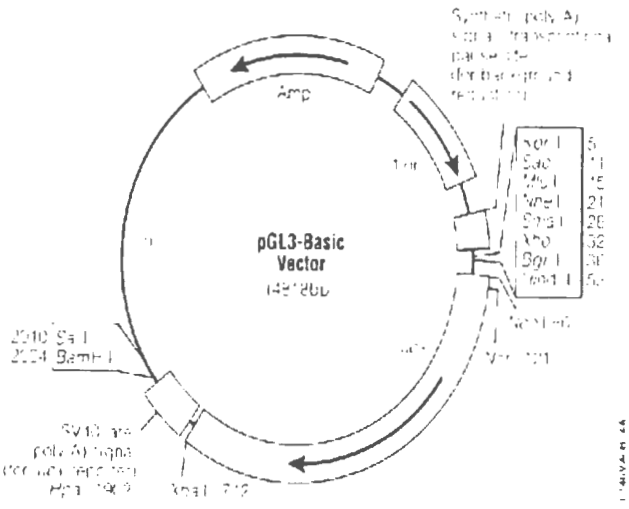
Escherichia coli DH5α (Invitrogen) was used as a host for the plasmid vectors. The pGL3 family of luciferase reporter vectors (pGL3 Basic, pGL3 Enhancer, pGL3 Promoter and pGL3 Control) were chosen for this present study. This family of vectors contain a modified firefly luciferase cDNA designated luc^+ and a redesigned vector backbone (Figure 4) that allows for increased luciferase expression, improved *in vivo* vector stability, and a greater flexibility in performing genetic manipulations over that of the earlier pGL2 vectors.

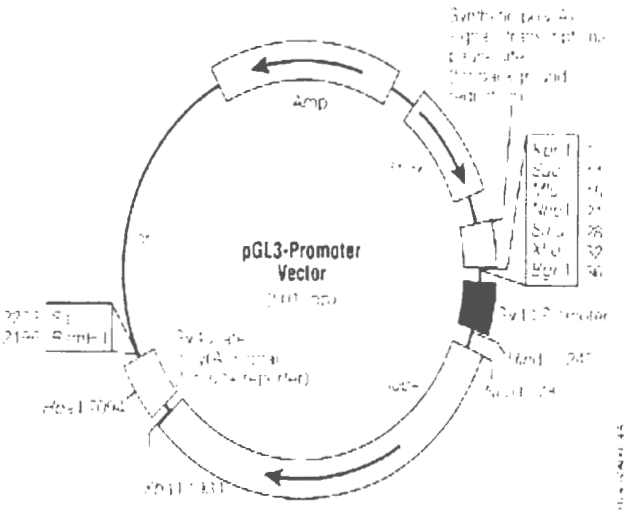
The opAFP-GHc was produced, sub cloned into the pBC vector (Stratagene) and sequenced by Du and colleagues (1992). The sequence data is presented in Genbank, Accession No. AY5946644. This construct contains the full length OP5a AFP gene promoter (2115bp) (Figure 5, construct number 1)

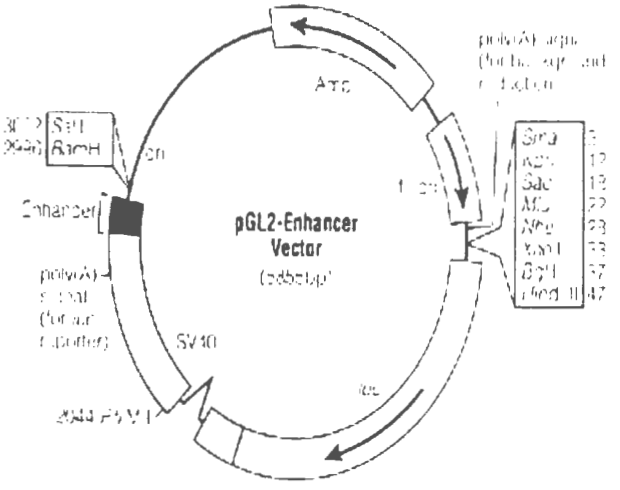
The construct termed pGL3 op # 2 was created using PCR strategies by Dr. Laurie Graham, (Department of Biochemistry, Queens University) (Figure 5, construct number 2). In this construct the first 331bp of the OP5α AFP promoter was eliminated prior to being cloned into the pGL3 Basic and pGL3 Enhancer vectors.

The constructs GM1-3X (Figure 3C) and 3D-GN (Figure 3D) were cloned into pUC18 vectors from segments of the EO-1α transgene(Figure 3B) that had integrated into the genome of Atlantic salmon. GM1-3X contains the OP5a AFP gene promoter

Figure 4. Structure of the pGL3 luciferase reporter vectors. These vectors all contain the cDNA encoding the modified firefly luciferase and the Amp^r gene conferring ampicillin resistance in *E. coli.* The pGL3 Control vector also contains the SV40 promoter and the SV40 enhancer. The pGL3 Promoter contains the SV40 promoter and the pGL3 Enhancer contains the SV40 enhancer. The full length promoter, the promoter constructs 2, 3 and 5 to 11 were ligated into the *Sma* I – *Bgl* II site of the pGL3 Basic and Enhancer vectors. The promoter deletion construct number 4 was created by ligating the *Bam*H I segment isolated from plasmid 3D – GN (Figure 3D) into the *Bam*H I site of the pGL3 Basic and pGL3 Enhancer plasmids that contained deletion construct number 3 (Figure 5) (Adapted from Wood, 1990).


pGL3-Basic
Vector
Amp
Xba I
Sac I
Xho I
Nhe I
Hind III
Sma I
Kpn I
Bgl II
Sal I
BamH I


pGL3-Promoter
Vector
Amp
Xba I
Sac I
Xho I
Nhe I
Hind III
Sma I
Kpn I
Bgl II
Sal I
BamH I


pGL2-Enhancer
Vector
Amp
Enhancer
SV40
Sma I
Kpn I
Sac I
Xho I
Nhe I
Hind III
Sal I
BamH I

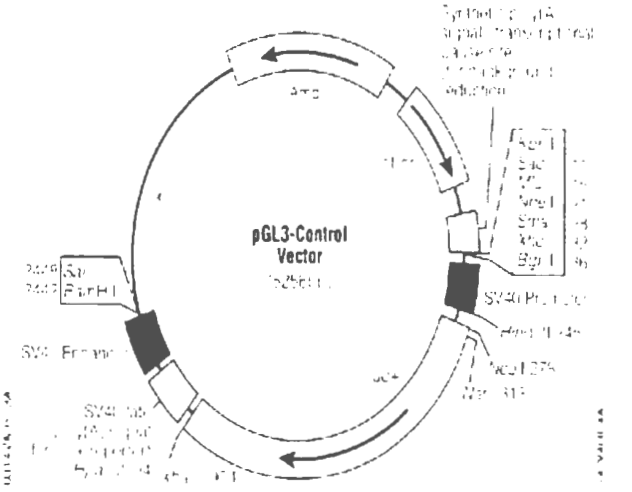

pGL3-Control
Vector
Amp
Xba I
Sac I
Xho I
Nhe I
Hind III
Sma I
Kpn I
Bgl II
Sal I
BamH I

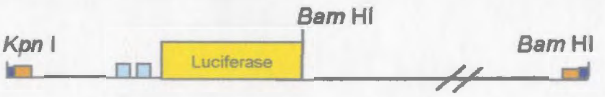
Figure 5. The opAFP- luciferase constructs transfected into the various cell lines. Each construct is denoted by the corresponding base pair numbers of each promoter sequence relative to the *Bgl* II site except for the downstream promoter sequence which is relative to the *Bam* HI restriction site. The restriction enzymes used to create each construct are noted. The solid line represents the promoter sequence. Each dark blue box represents Atlantic salmon genomic sequence and each orange box represents pUC18 sequence that is contained in the construct. Both of the CAAT and TATA boxes are represented by light blue boxes.

1. opAFP -2115 bp

2. opAFP -1784 bp

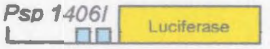
3. opAFP -536 bp

4. opAFP -536 bp + 1679 bp(+ second copy of 1580-1678bp) downstream promoter sequence



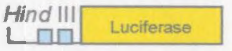
5. opAFP- 467 bp

6. opAFP - 363 bp



7. opAFP - 268 bp

8. opAFP - 216 bp



9. opAFP - 163 bp

10. opAFP -151bp

11. opAFP - 90 bp

that was truncated at a position 1580 bp. 3D-GN contains the remainder of the 5' end of the OP5a AFP gene promoter 1 bp to 1678 bp. These constructs were created by Dr. Edward S. Yaskowiak (Aqua Bounty Canada, Inc).

2. Cell lines

Three cell lines were used: human epithelial cells (CRL-1578, designation 293) (Figure 6), human fibroblast cells (HTB-92, designation SW 872) (Figure 7), and Chinook salmon cells (CRL 1681, designation CHSE-214) (Figure 8).

3. Deletion promoter and plasmid construction

A. Restriction Digestion and PCR of Promoter Constructs

In order to determine the potential regulatory role of the OP5a AFP gene promoter, the full length promoter (Figure 5, construct number 1) and promoter deletion constructs 5-11 (Figure 5) were created using restriction digestion of the opAFP-GHc construct (Figure 3A). The opAFP-GHc construct (1 μg) was first digested at the 5' end with one unit of the appropriate enzyme for an hour at the recommended temperature. The enzyme was then heat inactivated at the recommended temperature or inactivated by phenol: chloroform:isoamyl alcohol (25:24:1). When a 'sticky end' was created it was filled in with Klenow from *E.coli* DNA polymerase I (Fermentas) by incubating for 10 minutes at 37°C. The Klenow was heat inactivated at 75°C for 10 min and precipitated with ethanol.

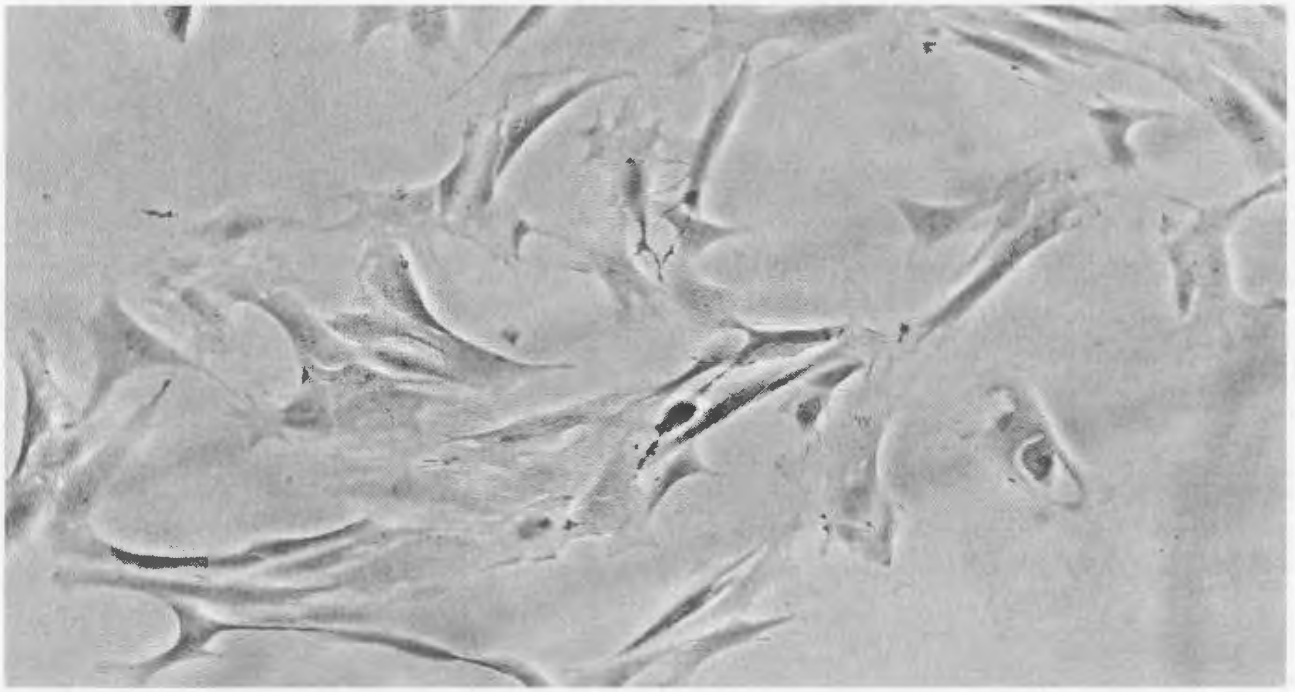


Figure 6. The 293 human epithelial cell line (CRL-1578) viewed at 100X. Source of cells are the human kidney that were transformed with adenovirus 5 DNA. These cells are propagated in Dulbecco's modified eagle medium (DMEM) containing 4.5 g/L D-glucose , L-glutamine, and pyridoxine HCl, supplemented with 10% fetal bovine serum (FBS) and 100 units/ ml of penicillin and streptomycin at 37^0C in the presence of 5% carbon dioxide.

Scale 1cm=10μm

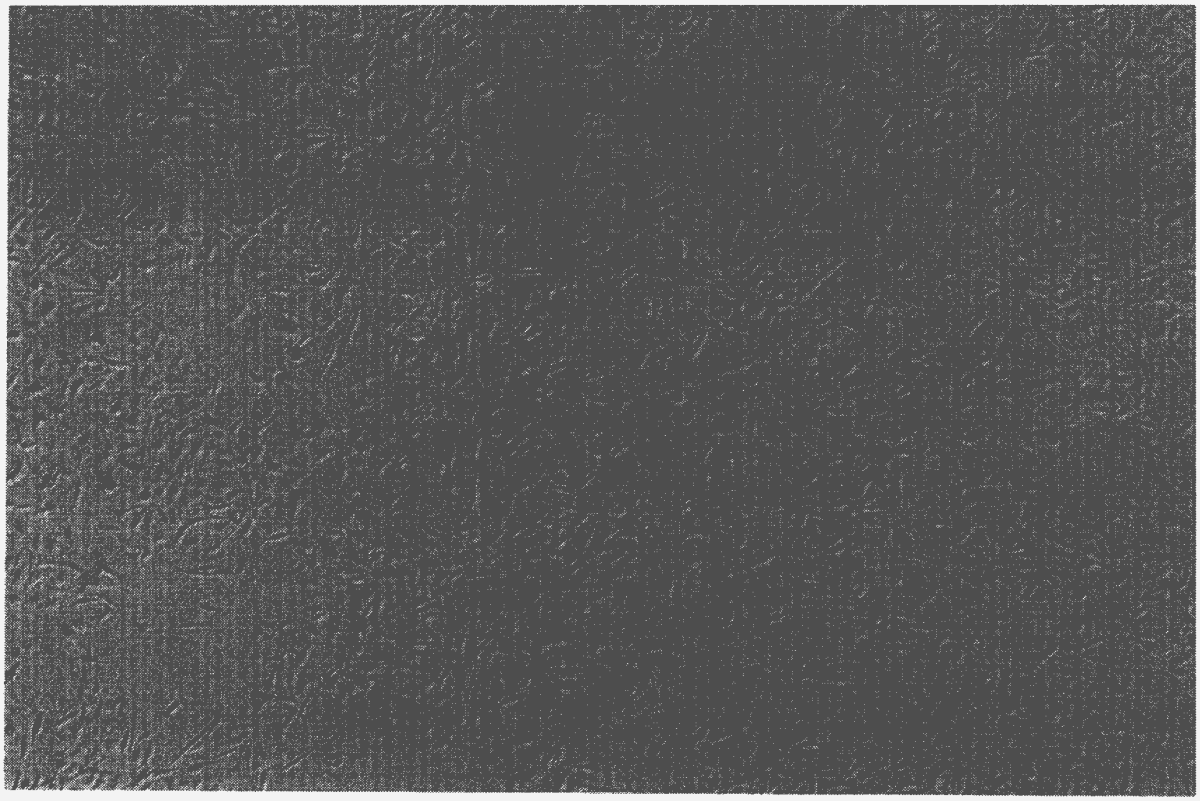


Figure 7. The SW 872 Human fibroblast cells (HTB-92) viewed at 40 X. Source of the cells are from human connective tissue. Cells were propagated in Leibovitz's L-15 medium containing L-glutamine and supplemented with 10% fetal bovine serum and 100 units/ ml of penicillin and streptomycin at 37°C.

Scale: 1cm=25μm

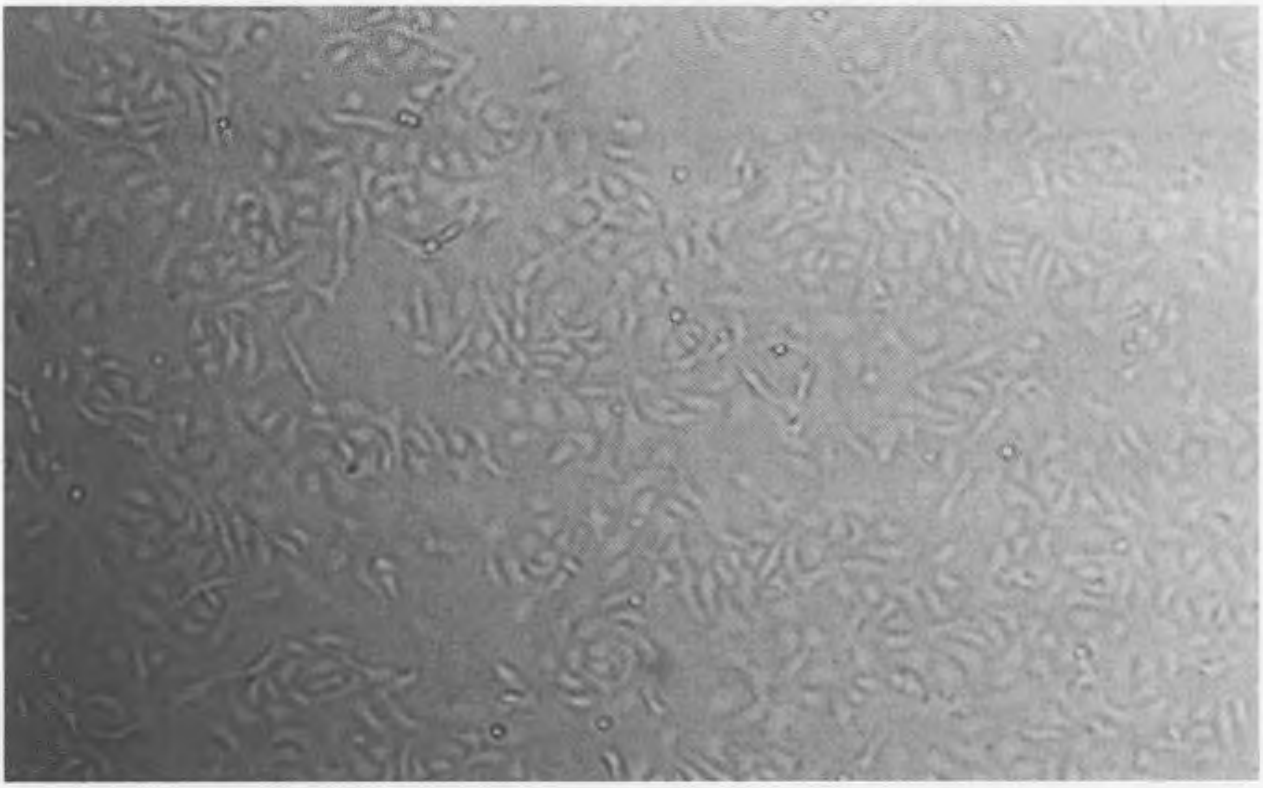


Figure 8. The CHSE-214 Chinook salmon cells (CRL 1681) viewed at 40 X. The cells were propagated in Minimum Essential Eagles medium (Invitrogen) containing Earle's salts and L-glutamine. The media was supplemented with 100 mM sodium pyruvate (Gibco) (final concentration of 10 mM) and 10% fetal bovine serum (ATCC). Scale: 1cm=20μm

Following this process the linearized constructs were digested with one unit of *Bgl* II, at 37°C for 1 hour, isolated using polyacrilomide gel electrophoresis (PAGE) and ligated into the *Sma* I – *Bgl* II site of the pGL3 Basic and Enhancer plasmids (Figure 4).

Promoter deletion construct number 3 (Figure 5) was prepared by digesting the plasmid GM1-3X (Fig 3C) with one unit of *Kpn* I at 37°C for an hour, followed by heat inactivation at 80 °C. The linearized plasmid was subsequently digested with 1 unit of *Bgl* II for 1 hour at 37 °C to release the deletion construct number 3 (Figure 5) plus an additional 6 bp of the pUC18 multiple cloning site and 112 bp of genomic DNA, purified on PAGE and ligated to the *Kpn* I – *Bgl* II site of the pGL3 Basic and pGL3 Enhancer plasmids (Fig 4).

Promoter deletion construct number 4 (Figure 5) was prepared by digesting plasmid 3D-GN (Fig 3D) with one unit of *Bam*H I for an hour at 37 °C to release the construct plus 7 bp of pUC18 multiple cloning site and 178 bp of genomic DNA, heat inactivated at 80 °C for 20 minutes, purified by PAGE and ligated into the *Bam*H I site (Figure 4) of the pGL3 Basic and pGL3 Enhancer plasmids that contained deletion construct 3 ligated into the *Kpn* I – *Bgl* I site. Restriction digestion was performed with various enzymes to ensure that the construct ligated into the *Bam*H I site of the pGL3 vectors in the correct orientation.

Sequence analysis of the OP5a AFP promoter revealed that the first 333 bp contained a reverse-complement sequence (125 – 333 bp) that was located at position 1143 to 1351 bp (see Figure 5). Inverted repeat DNA sequences are considered "at risk"

motifs for genetic instability that can result in both deletion and recombination events in the genome of eukaryotic organisms (Lobachev *et al.*, 1998). Promoter deletion contruct number 2 (pGL3 opAFP #2, Figure 5) was created to determine whether removal of the repeat located in the first 333 bp would alter regulatory aspects of the promoter.

B. Restriction digestion of pGL3 Plasmids

The pGL3 family of luciferase plasmids are widely used reporter vectors that contain the firefly luciferase gene. The pGL3 Basic and pGL3 Enhancer vectors were used as expression vectors into which the promoter constructs were cloned. The pGL3 Enhancer vector, containing a SV 40 eukaryotic enhancer, and the pGL3 Control vector, containing a SV 40 promoter and enhancer, were used as controls to monitor transfection efficiency.

Digests of the pGL3 Basic and pGL3 Enhancer plasmids (Figure 4) with restriction enzymes *Sma* I, *Kpn* I and *Bgl* II , which are located in the multiple cloning sites downstream of the luciferse gene, were performed as suggested by the supplying companies. Typically, 1 μg of plasmid DNA was digested with 1 unit of *Sma* I or *Kpn* I for 1 hour at 30°C or 37°C, inactivated at 65^0C for 20 minutes and DNA was precipitated using 0.1 volume of 3M sodium acetate and 2.5 volumes of cold 100% EtOH. The DNA pellet was washed using 75% ethanol and dissolved in deionized water (GIBCO). The plasmid was then digested with 1 unit of *Bgl* II at 37^0C for 1 hour, protein extracted using an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1), and centrifuged at

14,000 x g at room temperature for 5 minutes. The upper aqueous layer containing the plasmid was removed and the DNA precipitated. The desired fragment was separated on an agarose gel.

C. Isolation and purification of DNA fragments

The DNA fragments of interest were excised from the agarose gel and purified using a QIAquick gel extraction kit as recommended by the supplier (Qiagen). First, the desired DNA band was excised from the agarose gel, weighed, and placed in a 1.5 ml microcentrifuge tube. Then, the fragment was dissolved in 3 volumes (of the gel slice) of QG (binding) buffer for 10 minutes at 50^0C. This solution was then poured onto the silica spin column membrane, which is optimal for binding DNA, incubated for 1 minute at room temperature, and centrifuged for 1 minute at 13,000 x g in a microcentrifuge. Next, 750 µl of wash buffer was added, incubated for 2 to 5 minutes at room temperature and spun twice for 1 minute at 13,000x g in a microcentrifuge. The DNA was then eluted with 50 µl of deionzed water (GIBCO). The fragment was then run on a gel next to a DNA ladder of known concentration (100 kb ladder), so that the concentration of the DNA fragment could be approximated.

D. Ligation

The DNA promoter constructs were ligated to both the pGL3 Basic and pGL3 Enhancer plasmids. The molar ratio of insert DNA to vector was approximately 5:1, as estimated from PAGE. The ligation was carried out by adding insert DNA and vector to

deionized water (final volume of 15 µl). Then 4 µl of 5X Ligation reaction buffer (400 mM Tris-HCl, 100 mM $MgCl_2$, 100 mM DTT, 5 mM ATP) and 1 µl of T_4 DNA ligase (purified from an *Esherichia coli* strain) was added. This mixture was incubated at 16^0 C overnight.

E. Transformation

Transformations were performed by mixing 5 µl of the ligation mixture with 100 µl of the competent cell strain, *E. coli* DH5α. This mixture was incubated on ice for 30 minutes, heat-shocked at 37^0C for 40 seconds, and returned to ice for 2 to 5 minutes. Luria-Bertani (LB) broth, 990 µl, was added to the mixture and placed in a shaking incubator at 37^0C for 1 hour. A 250 µl aliquot of cells was then transferred to LB broth with ampicillin (amp) and incubated for an additional hour. The cells (50 to 200 µl) were spread onto LB-amp plates and incubated overnight at 37^0C. Individual colonies were then selected and analysed as described below to ensure that the insert had ligated to the plasmid.

F. Plasmid isolation and purification

Individual colonies on the LB-amp plates were screened for plasmids containing the insert. Plasmid DNA from these colonies was prepared using the QIAprep miniprep columns as recommended by the supplier. Briefly, each individual colony was placed in a 5.0 ml tube with 2.0 ml of LB-amp and incubated overnight in a 37^0C shaker. The next day 1.5 ml of the 2.0 ml sample was centrifuged for 3 minutes and the supernatant

discarded. The pellet was re-suspended in 250 µl of P1 (resuspension) buffer, to which was added 250 µl of lysis buffer and incubated at room temperature for a maximum of 5 minutes. A 350 µl aliquot of the neutralizing buffer was added to the resuspension and centrifuged for 10 minutes in a microcentrifuge. The supernatant was removed, loaded onto a QIAprep membrane and centrifuged for 1 minute. The membrane was then washed with 750 µl of wash buffer and spun twice for a minute. The DNA was then eluted with 50 µl of water. Various restriction enzymes that cut both the vector and insert DNA were used and subsequently run on PAGE to ensure that the insert was present in the proper orientation.

If the insert was present in the proper orientation, 250 µl of the original overnight culture was grown in 100 ml of LB-amp media at 37^0C overnight in a shaker. Plasmid DNA was isolated and purified from this culture using a large scale QIAgen Maxi prep kit following the procedure that was outlined by the supplier. Briefly, the 100 ml culture was spun at 4^0C at 6,000 x g for 15 minutes. The supernatant was then removed by aspiration and the pellet was re-suspended in 10 ml of P1 buffer. Next, 10 ml of lysis buffer was added, gently mixed and incubated at room temperature. After approximately 5 minutes, 10 ml of chilled neutralizer buffer was added to the lysate, mixed and poured into a QIAfilter cartridge. This mixture was incubated for 10 minutes to allow the proteins, genomic DNA and detergent to float to the top of the solution. Using a plunger, the solution was then loaded onto an equilibrated Qiagen tip membrane and allowed to pass through by gravity flow. The membrane, containing the bound DNA, was then washed with 60 ml of wash buffer. Finally, the DNA was eluted into a clean 50 ml tube

with 15 ml of QF-elution buffer and mixed with 0.7 volumes, 10.5 ml, of room temperature isopropanol. The solution was centrifuged for 10 minutes at 13,000 x g and the supernatant aspirated. The pellet was washed with 70% ethanol, centrifuged for an additional 5 minutes and re-suspended in 250 ml of water.

G. DNA Sequence of the ocean pout AFP gene promoter

The full length ocean pout AFP promoter construct and the various deletion constructs that were cloned in to the pGL3 Basic vector and pGL3 Enhancer vector were sequenced by Cortec DNA Service Laboratories Inc. (Queens University, Ontario) and analyzed to confirm the sequence of each construct.

4. Analysis of the promoter construct's activity using the luciferase reporter gene

A. Cell Culture

The human epithelial cells were grown in Dulbecco's modified Eagle medium (DMEM) (Invitrogen) containing 4.5 g/L D-glucose , L-glutamine, and pyridoxine HCl, supplemented with 10% fetal bovine serum (FBS)(ATCC) and 100 units / ml of penicillin and streptomycin at 37^0C in the presence of 5% carbon dioxide. The medium was changed every two days, and the cells were sub-cultured every 7-9 days when they reached 70 to 80% confluence. To sub-culture the cells they were first rinsed briefly with 2.0 ml of 0.25% (w/v) trypsin – EDTA containing 2.5 g of porcine trypsin and 0.2 g EDTA·4Na per liter of Hanks' Balanced Salt (HBS) solution with phenol red (GIBCO).

This removed all traces of the serum, which inhibits trypsin. Then 4.0 ml of the 0.25 % (w/v) trypsin – EDTA was added to cause the cell layer to disperse (1 to 2 minutes). After cell dispersal, 10 ml of complete media (media + serum + antibiotics) was added and the cells were resuspended by pipetting. The cells were spun in a table top centrifuge at 4^0C for 5 minutes at 2,000 x g. The supernatant was then removed and the cells were resuspended by vortex. Fresh medium was then added and the suspension of cells centrifuged again to ensure complete removal of the trypsin. After centrifugation, the media was removed by aspiration, the cells vortexed and were sub-cultured (1:4 ratio) into flasks containing complete media.

The human fibroblast cells were grown in Leibovitz's L-15 medium (Invitrogen) containing L-glutamine and supplemented with 10% fetal bovine serum (ATCC) and 100 units/ ml of penicillin and streptomycin at 37°C. The media was changed every two days, and the cells were sub-cultured every 6 to 10 days. Cells were sub-cultured in the manner described above.

The chinook salmon cells were grown in Minimum essential Eagle's medium (Invitrogen) containing Earle's salts and L-glutamine. The media was supplemented with 100 mM sodium pyruvate (GIBCO) (final concentration of 10 mM) and 10% fetal bovine serum (ATCC), however no antibiotics were added. The medium was changed twice weekly and the cells were sub-cultured every 3 to 4 weeks when they reached 70-80% confluence. The cells were sub-cultured in the same manner as the human epithelial and human fibroblast cells.

B. Optimization of transfection

To optimize transfection efficiency for each of the individual cell lines used in this study, various parameters were determined. Transfectam (Promega) is a commercial reagent used to facilitate the uptake of DNA in transfection. The first parameter to be determined was the Transfectam to DNA charge ratio. In each trial 2.0 μg of the pGL3 Basic vector was transfected into the individual cell lines at Transfectam to DNA charge ratios of 1:1 up to 10:1 while the duration of transfection was kept constant at 4.0 hours. The DNA to Transfectam ratio with the optimized transfection results were used in the actual study of promoter activity which will be described in detail in the next section.

The next parameter that was optimized for each cell line was the duration of transfection. To test this parameter all experiments were performed with a constant plasmid DNA ratio of 2.0 μg and a constant Transfectam to DNA ratio of 2:1. The range of duration of transfection was from 1 hour to sixteen hours. Once again the duration which was optimal for each cell line was used in the study.

C. Transfection of cell lines

For each cell line roughly 3 x 10^5 cells were plated per 35 mm culture dish, two days prior to the transfection experiment. This allowed the cells to reach approximately 70% confluence. The day of the transfection experiment, the cells were washed twice with 2 ml of the appropriate serum free media and incubated for 10 minutes at the appropriate temperatures, 37^0C for both of the human cell lines and 21^0C for the chinook salmon cell line. During this incubation period, each of the promoter - luciferase

constructs was thawed and 2.0 μg of the plasmid constructs or the four pGL3 plasmid controls were added to 500 μl of serum free media in sterile microcentrifuge tubes and vortexed. In separate microcentrifuge tubes, Transfectam (Promega) was added to 500 μl of serum free media and mixed. In the case of the human cell lines, 3.75 μl of Transfectam was used. For the transfection of the salmon cell line, 0.75 μl of Transfectam was used. The two solutions were then mixed and added to the correspondingly labelled plates containing the cells. The human cell lines were then incubated at 37°C for 16 hours, while the Chinook salmon cell line was incubated at 21 °C for 2 hours. After the incubation, 4 ml of complete media was added to the cells and incubated for a further 24 hours.

After incubation, the media was removed and the cells rinsed with serum free media. A 500 μl aliquot of cell culture lysis reagent - CCLR (Promega) was added to each culture well and allowed to incubate for ten minutes at room temperature while being rocked on a rotating shaker. The cells were then scraped from the dish and transferred, with the reagent, to a microcentrifuge tube. This was performed on every well containing cells that was transfected with a promoter construct. The microcentrifuge tubes were then vortexed for 10 seconds and centrifuged for 30 seconds at room temperature. The supernatant containing the cell lysate was then transferred to a fresh tube and placed on ice until analysis.

D. Luciferase assay

Luciferase assays were carried out using a simple luminometer (LabSystems, Finland) according to the guidelines supplied by the manufacturer. The procedure is summarized as follows: 100 µl of the luciferase assay reagent - LAR (Promega) and 20 µl of cell lysate was added to the luminometer tube containing the LAR and placed into the luminometer. The luminometer was then programmed to perform a two second measurement delay followed by a 10 second measurement read for luciferase activity. The relative intenstity was then recorded for each promoter construct and control DNA. This protocol was consistent for every cell line used.

Luciferase activity was corrected for protein measured by the Bradford (Bio-Rad) assay and calculated as luciferase per milligram of protein. All assays were performed in duplicate and the final results are the mean of four independent experiments.

Results

1. The optimization of gene transfection in cultured cells

The parameters that were optimized in order to maximize transfection efficiencies were the Transfectam to DNA charge ratio and the length of time that the cells were exposed to the Transfectam reagent. These parameters differed with each cell line and were essential for the success of the transfections.

To optimize the Transfectam to DNA charge ratio, the pGL3 Control plasmid was transfected into the three different cell lines using various Transfectam to DNA ratios. All experiments were performed with a constant DNA concentration of 2.0 μg and a constant duration of transfection of 2 hours. For both the human epithelial and fibroblast cell lines the optimal Transfectam:DNA ratio was 5:1 (Figure 9). In contrast, transfection of the chinook salmon embryonic cell line was optimized at a Transfectam:DNA ratio of 1:1. The higher Transfectam to DNA charge ratios in this fish cell line increased the instance of cell death.

The optimal duration of transfection was performed using a constant DNA concentration of 2.0 μg. A Transfectam to DNA ratio of 1:1 was kept constant for the salmon cell line, and a Transfectam:DNA ratio of 5:1 was used for both of the mammalian cell lines. In the case of the salmon cell line the highest level of luciferase expression was observed when the cells were exposed to Transfectam for one hour (Figure 10). Longer exposure time yielded lower levels of luciferase activity per plate of

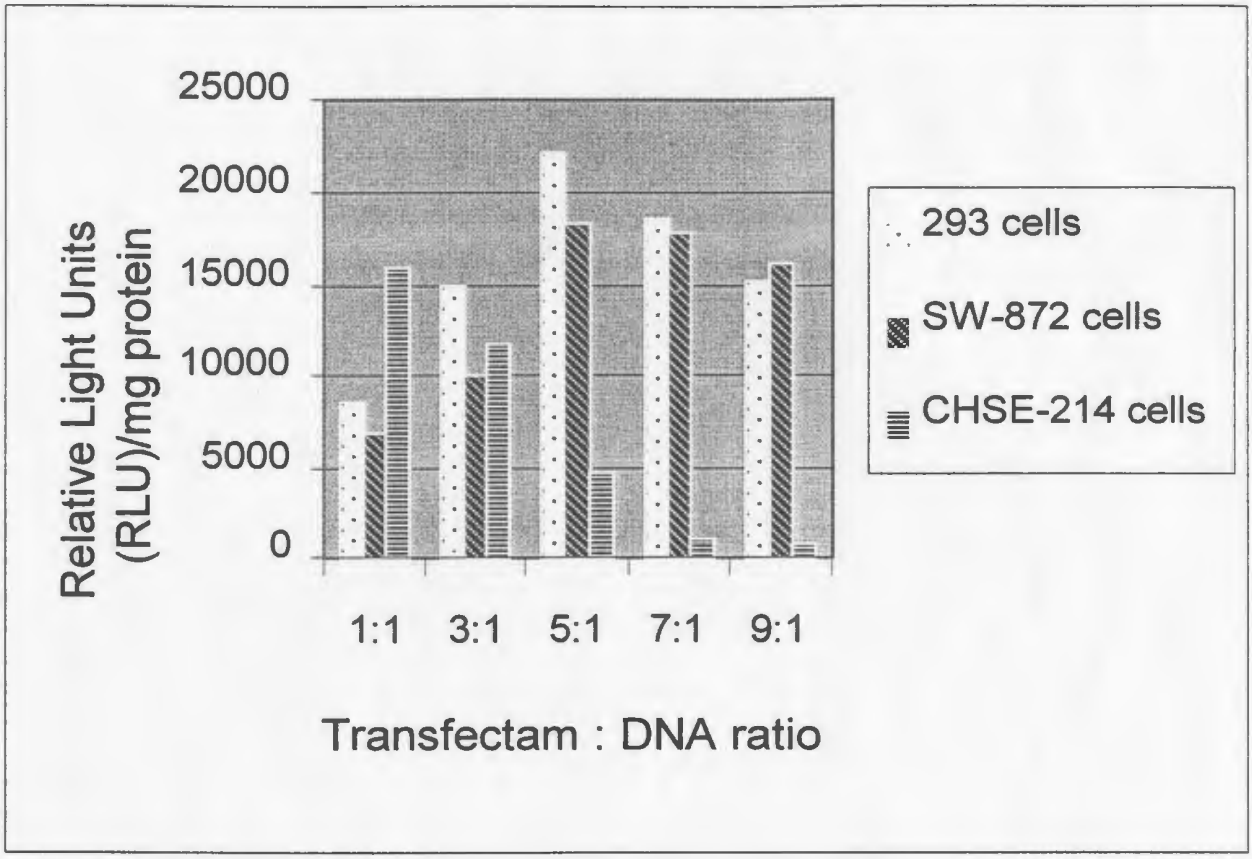


Figure 9. The relative levels of luciferase expression as a function of Transfectam:DNA charge ratio. The 293 cells, SW-872 cells, and CHSE-214 cells were plated at a density of 3 x 10^5 cells/ 35 mm well. Transfections were preformed in the absence of fetal bovine serum (FBS) and using pGL3 Control plasmid at various Transfectam:DNA ratios. All transfections were overlaid with complete media (containing FBS) after two hours, and cells were tested for luciferase activity after 48 hours. These results represent the mean of two replicate studies performed simultaneously.

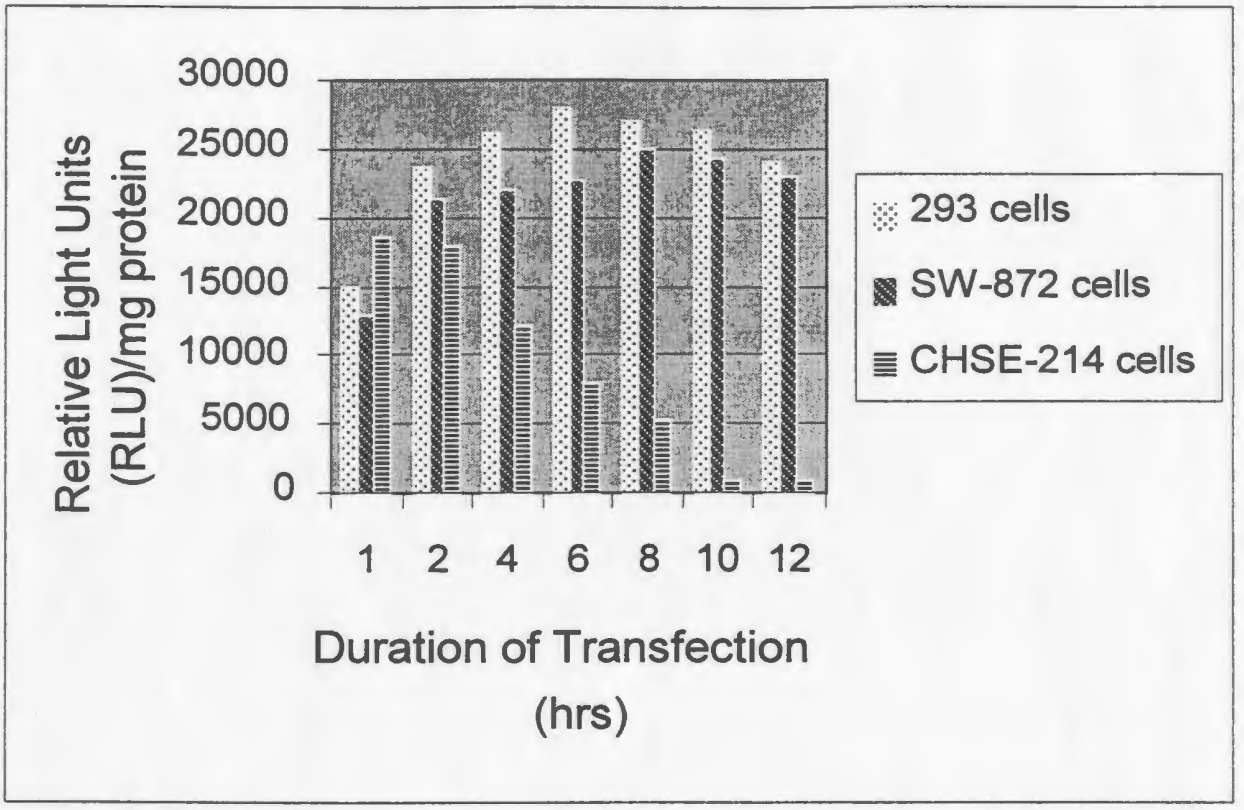


Figure 10. The relative levels of luciferase expression as a function of duration of transfection. The 293 cells, SW-872 cells, and CHSE-214 cells were plated at a density of 3 x 10^5 cells/ 35 mm well. Transfections were preformed using the optimal pGL3 Control plasmid: Transfectam charge ratio in the absence of fetal bovine serum (FBS) for various lengths of times up to 12 hours. All transfections were then overlaid with complete media (containing FBS) and cells were tested for luciferase activity after 48 hours. These results represent the mean of two replicate studies performed simultaneously.

cells. Longer exposure times altered cell morphology and increased the instance of cell death to a high degree. In the case of the mammalian cell lines the highest level of luciferase activity was obtained after 6 and 8 hours of exposure to the transfectam in the 293 and SW- 872 cells respectively (Figure 10).

2. The test of vector function

The pGL3 Basic, pGL3 Promoter, pGL3 Enhancer, and pGL3 Control were transfected into the three different cell lines and assayed for luciferase activity. As expected the pGL3 vector did not direct luciferase expression in the absence of the SV40 enhancer or promoter (Figure 11). The luminance emitted from the pGL3 Basic vector was comparable to the luminance emitted from the cell lysate that had not been transfected with either of the pGL3 vectors (No DNA, Figure 11) and was therefore considered as background luminance. The pGL3 Control vector, which possesses the SV40 promoter and enhancer exhibited the highest level of luciferase activity with all three cell lines and was thus designated as 100% luminescence. The level of expression within the pGL3 Control vector was 3 to 5 fold higher than that of the pGL3 Promoter, depending on the cell line. The pGL3 Enhancer vector exhibited low levels of expression that was only slightly greater than that exhibited by the pGL3 Basic vector.

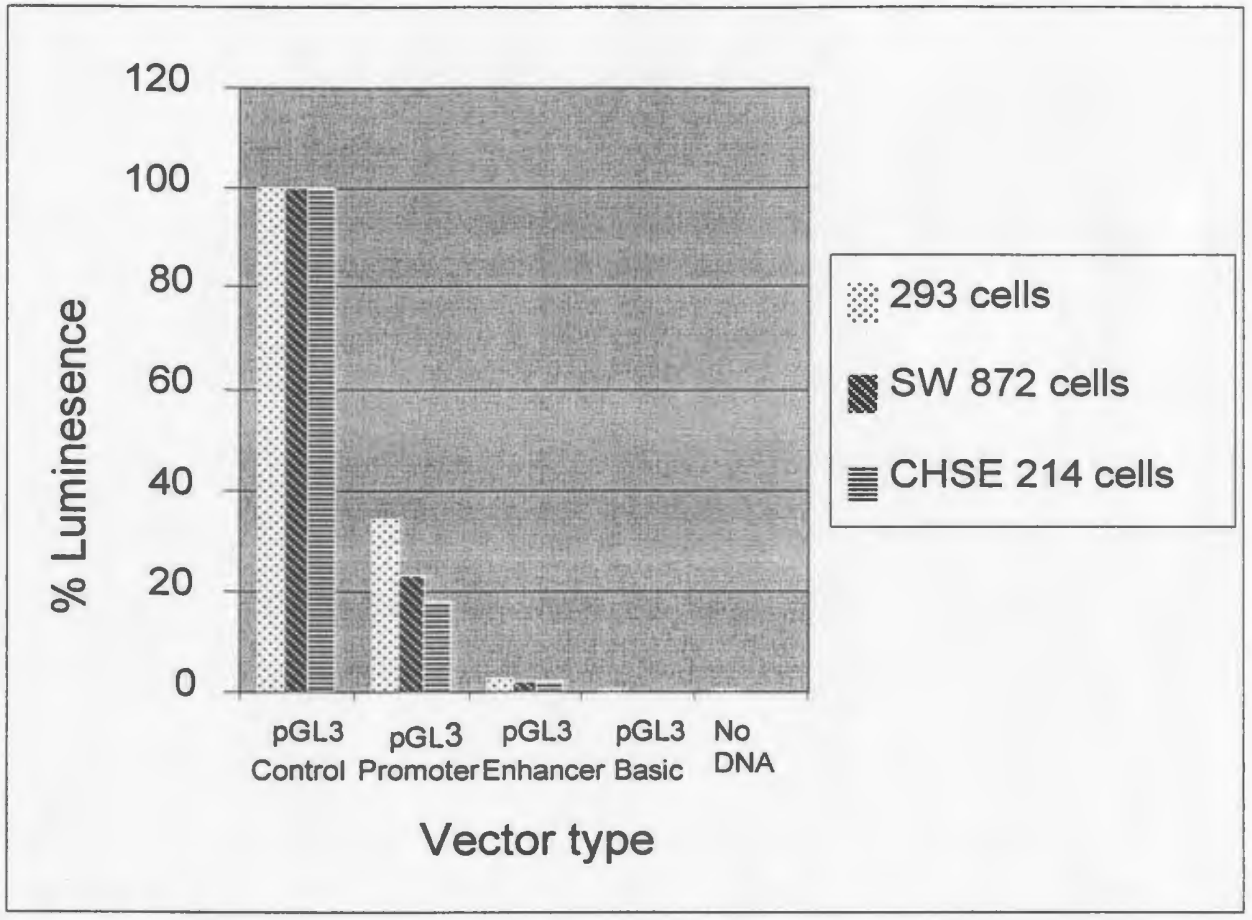


Figure 11. The relative percentage of luminescence as a test of vector function. The 293 cells, SW-872 cells, and CHSE-214 cells were plated at a density of 3 x 10^5 cells/ 35 mm well. Transfections were preformed under the optimal conditions. Results were expressed in comparison with the pGL3 Control plasmid. These results represent the mean of two replicate studies performed simultaneously.

3. Analysis of the ocean pout AFP gene promoter

Upon analysis of the ocean pout AFP gene promoter, it was discovered that the region 125 - 333 and 1143 - 1351 base pairs upstream of the transcriptional start site contain complementary repeat sequences that only differ by 8 bp (Figure 12). The first 331 bp of sequence was removed by PCR to provide the truncated promoter construct number 2 (Figure 5).

The AFP gene promoter sequence was also analyzed for the various *cis*-acting elements including the Goldberg-Hogness box (TATA box), the CAAT binding motif, the GC domain, the initiator (Inr) sequence, the DPE sequence and others. Like most other eukaryotic class II genes, the proximal 5` sequence upstream of this promoter contains the TATA box and the CAAT box. However there were no other sequences showing apparent homology to any other of the known *cis*-acting regulatory regions. The TATA box was found 29 bp upstream from the start of transcription and the CAAT box was found 73 bp upstream from the transcription start site. The transcription start site (+1) was also identified (Figure 12).

4. Functional analysis of the AFP promoter deletion constructs in human and fish cell lines

The results of the promoter analyses using the pCL3 Basic vector are presented numerically in Figure 13 and graphically in Figure 14. The highest luciferase expression levels were exhibited by the full length promoter in the mammalian cell lines and by the opAFP-163 bp deletion construct (construct 9) in the fish cell line. Expression levels for

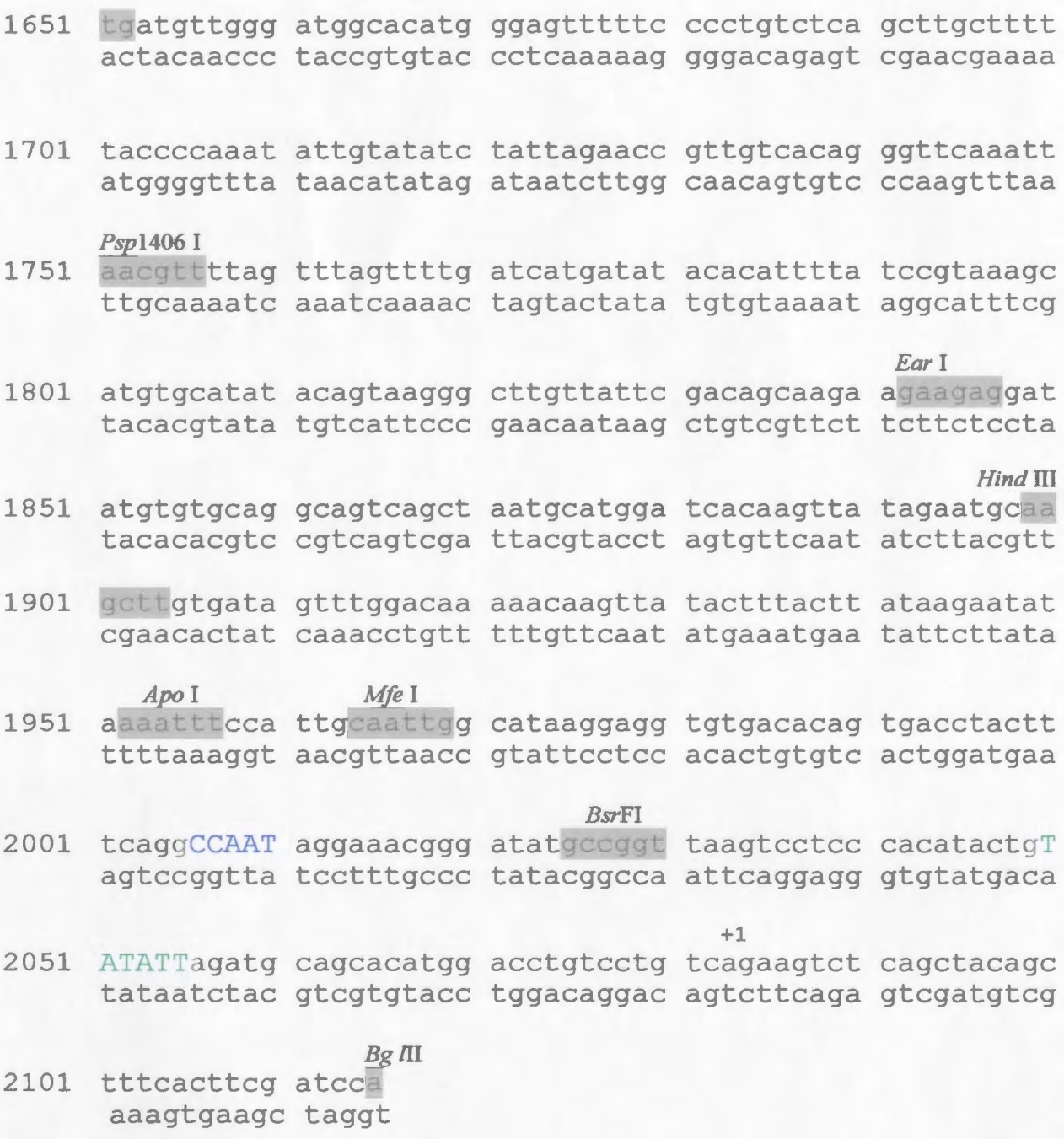
Figure 12: The ocean pout antifreeze protein gene promoter sequence. This figure shows the 2115 bp gene promoter sequence including the TATA box and CAAT regulatory regions. The sequence in red is the reverse complement repeat regions, out of which there are 8 mismatched base pairs shown in black. The restriction sites used to create the deletion constructs are highlighted. The last nucleotide a was involved in the creation of a *Bgl* II site when the promoter was ligated to the 5'UTR region in the chimeric gene construct opAFP-GHc(2). The enzymes *Bam* HI and *Bgl* II were used to release the promoter from the chimeric gene construct.

Bam HI

```
1     ggatcccca gaatgagctg gaacatgttg cggggagagg gaagtctggg
      cctaggggt cttactcgac cttgtacaac gcccctctcc cttcagaccc

51    tcagcctgct tggcctgctg ccaccgtgac ccgacctcag ataagcggag
      agtcggacga accggacgac ggtggcactg ggctggagtc tattcgcctc

101   gaaaatggat ggatggattg aatcacagaa tgtttctgaa gacagatatc
      cttttaccta cctacctaac ttagtgtctt acaaagactt ctgtctatag

151   accttcgctt caaagaggtg cgcacctggg caggcaccca acagccaca
      tggaagcgaa gtttctccac gcgtggaccc gtccgtgggt tgtcggtgt

201   caaatggcat atgaatcaac caagaagacg gttggaactg gtcaaaacct
      gtttaccgta tacttagttg gttcttctgc caaccttgac cagttttgga

251   tcactatacc atgtgtgaca gttgtttgtc acagtgtata aaagacaggg
      agtgatatgg tacacactgt caacaaacag tgtcacatat tttctgtccc

301   acttagagac agagctctga gcagctatga gattgtagtt tggccaggat
      tgaatctctg tctcgagact cgtcgatact ctaacatcaa accggtccta

351   gcgcttaaga cctttgtgat gaaaagttat caaattcgtg agttttcatg
      cgcgaattct ggaaacacta cttttcaata gtttaagcac tcaaaagtac

401   gaagaacctt gacgtggcgt ggtggccatt ttgcgtcatt cggcatggaa
      cttcttggaa ctgcaccgca ccaccggtaa aacgcagtaa gccgtacctt

451   aaggaagtcg ttataactcc caggtacatt atcttatcta cacaaaatgt
      ttccttcagc aatattgagg gtccatgtaa tagaatagat gtgttttaca

501   ctaatgcatg atactactta aagcctgagc atatttcaag gccagcactt
      gattacgtac tatgatgaat ttcggactcg tataaagttc cggtcgtgaa
```

```
551   ttcaataact cataggccac ctgctggcaa aaggaaatgc cacattttat
      aagttattga gtatccggtg gacgaccgtt ttcctttacg gtgtaaaata

601   acttttattt actcctagac agttgacctg atcagtctca aatttggtaa
      tgaaaataaa tgaggatctg tcaactggac tagtcagagt ttaaaccatt

651   ggatagcctt aagacaatga agatgcttca tcaggaatat tgtgagttgt
      cctatcggaa ttctgttact tctacgaagt agtccttata acactcaaca

701   cgttgaacgt tgttgccgtg gcaacgcatc attcgccatg aaaaagaagc
      gcaacttgca acaacggcac cgttgcgtag taagcggtac tttttcttcg

751   tgatggttca gtggcttggg atgctcaaaa agtcatggaa ctttgtacat
      actaccaagt caccgaaccc tacgagtttt tcagtacctt gaaacatgta

801   gtgtcataat tgatgggaag ttgtatgggt ttttggcttg cttgttataa
      cacagtatta actacccttc aacataccca aaaaccgaac gaacaatatt

851   attgtctcca tagcgccccc tacaatattt caaaagagca gccccagtgc
      taacagaggt atcgcggggg atgttataaa gttttctcgt cggggtcacg

901   tacgtacatg tatgaaactt agtagccaga tgtaccatat agagacttac
      atgcatgtac atactttgaa tcatcggtca tcatggtata tctctgaatg

951   aaaaaggtat cttggccatg ctctcaaccg tactggaagt cggccatttt
      tttttccata gaaccggtac gagagttggc atgaccttca gccggtaaaa

1001  gatttttgca taatttttca atagattttt gcacatttgt aatcgctata
      ctaaaaacgt attaaaaagt tatctaaaaa cgtgtaaaca ttagcgatat

1051  ctttaacgaa ctcctccaag gaactttgtc taatcaattt caaattttgt
      gaaattgctt gaggaggttc cttgaaacag attagttaaa gtttaaaaca
```

```
1101  cagtacaatc tcagtactac agtaccaaat ctacagttct gcatctcgta
      gtcatgttag agtcatgatg tcatggttta gatgtcaaga cgtagagcat

1151  gctgctcaga ggtctgtctc taagtccctg tcttttatac actgtgacaa
      cgacgagtct ccagacagag attcagggac agaaaatatg tgacactgtt

1201  acaactgtca cacatggtat agtgaaggtt ttgaccagtt ccaaccgtct
      tgttgacagt gtgtaccata tcacttccaa aactggtcaa ggttggcaga

1251  tgttggttga ttcatatgcc attcgtgtgg ctgtgtgggt gcctacccag
      acaaccaaca aagtatacgg taagcacacc gacacaccca cggatgggtc

1301  atgcgcacct ctttgaagcg aatgtgatat ctgtcttcat aaacattctg
      tacgcgtgga gaaacttcgc ttacactata gacagaagta tttgtaagac

1351  ttattagcaa gttcatatga gattgaaggc tgtatgcaaa caggtgcaca
      aataatcgtt caagtatact ctaacttccg acatacgttt gtccacgtgt

1401  gtctgtttct aagcatcatg gaaaagtaca agcaatttgc acaaatcatt
      cagacaaaga ttcgtagtac cttttcatgt tcgttaaacg tgtttagtaa

1451  ctgtattttt ccaatagcta acaatgtcac cgggacattg tgctattgga
      gtcataaaaa ggttatcgat tgttacagtg gccctgtaac acgataacct

1501  tagaagagac cagctgatct agacagttga tatcatgatc aacagcccca
      atcttctctg gtcgactaga tctgtcaact atagtactag ttgtcggggt

1551  aacaacaagt gtgcatgcgc gaggagtgat tggcagatgt atgagaacta
      ttgttgttca cacgtacgcg ctcctcacta accgtctaca tactcttgat

                                                    Nde I
1601  aaccactgac tgaacttgca ctagaggcat ctattttgtc ttttctcata
      ttggtgactg acttgaacgt gatctccgta gataaaacag aaaagagtat
```

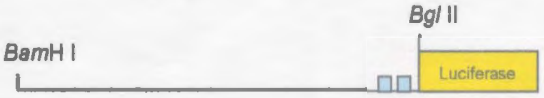
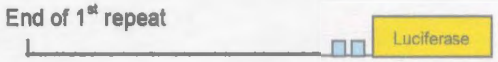
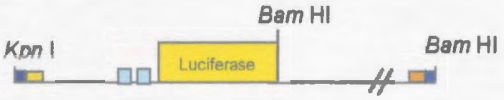
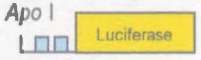
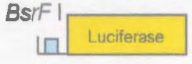

1651 tgatgttggg atggcacatg ggagtttttc ccctgtctca gcttgctttt
actacaaccc taccgtgtac cctcaaaaag gggacagagt cgaacgaaaa
1701 taccccaaat attgtatatc tattagaacc gttgtcacag ggttcaaatt
atggggttta taacatatag ataatcttgg caacagtgtc ccaagtttaa
Psp1406 I
1751 aacgtttag tttagttttg atcatgatat acacatttta tccgtaaagc
ttgcaaaatc aaatcaaaac tagtactata tgtgtaaaat aggcatttcg
Ear I
1801 atgtgcatat acagtaaggg cttgttattc gacagcaaga agaagaggat
tacacgtata tgtcattccc gaacaataag ctgtcgttct tcttctccta
Hind III
1851 atgtgtgcag gcagtcagct aatgcatgga tcacaagtta tagaatgcaa
tacacacgtc cgtcagtcga ttacgtacct agtgttcaat atcttacgtt
1901 gcttgtgata gtttggacaa aaacaagtta tactttactt ataagaatat
cgaacactat caaacctgtt tttgttcaat atgaaatgaa tattcttata
Apo I
Mfe I
1951 aaaatttcca ttgcaattgg cataaggagg tgtgacacag tgacctactt
ttttaaaggt aacgttaacc gtattcctcc acactgtgtc actggatgaa
BsrFI
2001 tcaggCCAAT aggaaacggg atatgccggt taagtcctcc cacatactgT
agtccggtta tcctttgccc tatacggcca attcaggagg gtgtatgaca
+1
2051 ATATTagatg cagcacatgg acctgtcctg tcagaagtct cagctacagc
tataatctac gtcgtgtacc tggacaggac agtcttcaga gtcgatgtcg
Bg lII
2101 tttcacttcg atcca
aaagtgaagc taggt

all other constructs are expressed as percentages of these maximal values for each of the cell lines.

The overall relationship between luciferase expression and ocean pout AFP promoter deletion construct length is consistent among all three cell lines (Figures 13 and 14).

Removal of the first reverse complement repeat (-2115 to 1784, construct 2) results in a slight reduction in luciferase expression in all three cell lines. However shortening the promoter to 536 bp (construct 3, which also contains pUC and genomic sequence) resulted in a dramatic reduction in luciferse expression to approximately 30% of the activity of the full length promoter. Ligating the deleted 1678 bp 5' promoter sequence downstream of the luciferase reporter gene in the pGL3 vector (construct 4) resulted in a modest increase in luciferase expression that was far less than the expression exhibited by the full length promoter. Further nucleotide deletions of the promoter to produce 467 bp and 363 bp construct (constructs 5 and 6) resulted in little or no change in luciferase expression in the mammalian cell lines, and 50 to 70% increased expression in the salmon cell line. When the deletions were extended through to a 163 bp construct (constructs 7, 8 and 9) luciferase expression in the mammalian cell lines increased to values approaching that of the full length promoter and, in the fish cell line, to double that of the full length promoter. Further deletions resulted in reduced expression. The shortest construct examined (90 bp, construct 11) resulted in expression amounting to approximately 25% of that exhibited by the strongest promoter in each of the cell lines (Figure 13 and 14).

Figure 13. Relative luciferase expression of opAFP gene promoter constructs in three cell lines. Each construct is denoted by the corresponding base pair numbers of each promoter sequence relative to the *Bgl* II site except for the downstream promoter sequence which is relative to the *Bam* HI restriction site. The restriction enzymes used to create each construct are noted. The solid line represents the promoter sequence. Each dark blue box represents Atlantic salmon genomic sequence and each orange box represents pUC18 sequence that is contained in the construct. The TATA and CAAT sequences are represented by the light blue boxes. The luciferase activity is shown as percentages relative to the full length promoter in the mammalian cells and to opAFP-163 in the fish cells.

		Luciferase Activity		
		293	SW 872	CHSE 214
1. opAFP -2115 bp		100	100	48
2. opAFP -1784 bp		96	97	45
3. opAFP -536 bp		34	31	14
4. opAFP -536 bp + 1679 bp (+ second copy of 1580-1678 bp) downstream promoter sequence		41	38	20
5. opAFP- 467 bp		34	28	28
6. opAFP - 363 bp		27	23	21
7. opAFP - 268 bp		44	47	55
8. opAFP - 216 bp		73	59	69
9. opAFP - 163 bp		82	78	100
10. opAFP -151bp		80	71	85
11. opAFP - 90 bp		29	26	25

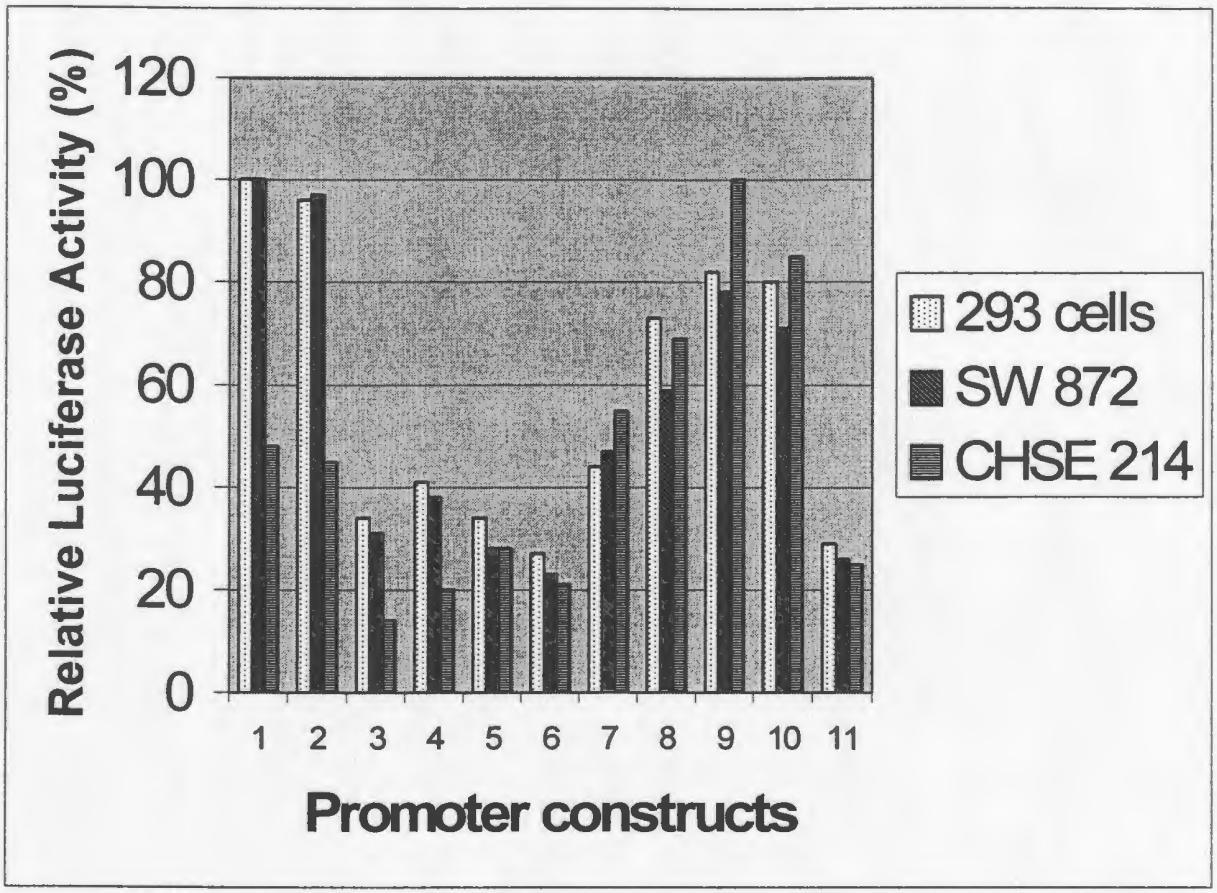


Figure 14. Relative luciferase expression of opAFP gene promoter constructs in three cell lines. Note that the promoter construct number corresponds to the numbered constructs in figure 13.

Discussion

1. The luciferase expression vector for promoter analysis

The pGL3 family of reporter vectors provided the basis for the quantitative analysis of the *cis*-acting promoter region regulating gene expression. The pGL3 reporter vectors contain the firefly luciferase gene which encodes an enzyme that catalyzes the light producing, ATP - dependent oxidation of luciferin (Ow *et al.*, 1986). The firefly luciferase gene is widely used as a reporter gene for many reasons. The reporter activity is immediately available upon translation since the protein does not require post - translational processing. The assay is extremely sensitive because its light production has the highest quantum efficiency known for any chemiluminescent reaction, and no background luminescence is found in the host cells or the assay chemistry. Of importance, also, is the fact that the assay is rapid, requiring only a few seconds per sample (Wood, 1990).

The pGL3 vectors also provided the appropriate multiple cloning sites for the ligation of the promoter constructs. There was a *Sma* I restriction site located at the beginning of the multiple cloning site. *Sma* I is a blunt end restriction site, which enabled other restriction sites that had been filled in with Klenow to be easily cloned into the site. Also, downstream of the *Sma* I site there was a *Bgl* II site. This was convenient, since at the 5' end of the ocean pout AFP promoter sequence there is a *Bgl* II site as well which provided an opportunity for a compatible ends to be easily ligated.

2. Optimization of transfection

It is essential to optimize specific transfection conditions to gain optimal transfection efficiencies. The optimization studies indicate that the mammalian cell lines are similar to each other but differ with from salmon fish cell line with respect to the optimal Transfectam to DNA charge ratio and the duration of transfection. As found here, and in other studies, fish cells seem to be more sensitive to reagents that are used to render mammalian cells competent for the uptake of DNA. There seemed to be a dramatic increase in morphological change and cell death in the fish cell line as Transfectam to DNA ratio and the duration of transfection increased. Transfectam is a synthetic cationic lipopolyamine molecule, containing a lipid head-group, which combines with DNA converting it into a DNA carrier complex mimicking a cationic lipid layer. This structure makes contact with the cell membrane through a cooperative ionic interaction and promotes the endocytosis of DNA. Unlike other commercial reagents, Transfectam has been successful in transfecting both cultured mammalian and fish cell cultures (Villalobos *et al.*, 1999).

3. DNA sequence analysis of the ocean pout AFP gene promoter

Populations of ocean pout off the Newfoundland coast have approximately 150 copies of the type III AFP genes present in the genome (Hew *et al.*, 1988). This makes it difficult to select an individual gene for promoter analysis and be confident that its activity is typical for the majority of the type III AFP genes in ocean pout. Previous work on protein and DNA sequence analysis has established that the ocean pout AFP can be

fractionated into five distinct groups based on their binding to ion exchange resins. One group binds to QAE-Sephadex (QAE-1) and four to SP-Sephadex (SP-1 to SP-4). Although there is only 55% sequence identity among the SP and QAE components, their 5'flanking DNA regions are highly conserved. In fact, it has been reported that there is over 90% sequence identity among five different AFP genes in a stretch of about 400 bp immediately upstream of their translation start site (Hew et al., 1984; Hew et al., 1988). This high degree of sequence conservation within the promoter region of the AFP gene allows one to be confident that the analysis of an individual gene can provide insight into the activity and regulation of the majority of the AFP genes.

Similar to the majority of eukaryotic class II genes, the ocean pout AFP gene promoter contains a TATA box and a CAAT box. However, I found no other sequence motifs in the promoter that shows any apparent homology to other known regulatory elements. Previous analysis of the 5' flanking sequences of types I, II and II AFP did not reveal any common regulatory elements among them (Gong and Hew, 1993). Also, I did not look for the presence of any boundary or insulator elements in the sequence. The identification of insulator elements in the ocean pout antifreeze gene promoter has not been looked at in previous studies. It could, however, be an important study in promoter analysis, since boundary elements are known to regulate promoter and enhancer sequences, thus affecting promoter strength. It has been proposed previously that insulators organize the eukaryotic genome into domains of gene activity, so that regulatory sequences in one domain cannot interact with promoters in a different one (Gdula *et al.* 1996). There are not a lot of studies regarding the identification of insulator

sequences, however, a few sequences have been identified in *Drosophila* and vertebrates using the test vector system. This system uses a pNI vector containing the Neomycin-resistance gene driven by the human γ-globin promoter and stimulated by the mouse HS2 enhancer of the ß-globin LCR. Briefly, constructs to be tested are placed between the promoter and enhancer ('enhancer-blocking or insulator site') in the test vector, and transfected into a human K562 cell line in the presence of G418. The number of G418-resistant colonies are measured and insulator activity is indicated by a decrease in the number of G418-resistant colonies. (Gerasimova and Corces, 1998 and Du *et a*l., 2003). G418 is an aminoglycoside used as a selective agent of transfected mammalian, yeast, plant and bacteria cells.

The present study did reveal the presence of a reverse – complement sequence within the promoter. Such inverted DNA repeats are considered "at risk" motifs for genetic instability because they can induce both deletion and recombination events within the genome of eukaryotic organisms (Lobachev *et al.*, 1998). Therefore, if possible they should be avoided in the creation of gene constructs for the production of transgenic organisms.

4. The ocean pout AFP gene promoter can function in various cell types and eukaryotic systems

The functional ability of the OP5a AFP gene promoter was clearly demonstrated in the present study by its ability to drive expression of the luciferase reporter gene in three diverse cell lines: 293 human epithelial cell, SW-872 human fibroblast cells, and CHSE-214 chinook salmon embryonic cells. These 293 epithelial cell line and SW- 872

cell line was chosen due to the fact that this promoter had not been previously tested in a mammalian cell line. While the CHSE-214 chinook salmon cell line was chosen to confirm results in a previous study undertook by Gong and Hew, 1993. These results confirm previous reports indicating that the OP5a AFP promoter can function in a wide range of fish cell types and in transiently transgenic medaka (*Oryzias latipes*) embryos (Gong *et al.*, 1991, Gong and Hew, 1993) as well as mammalian cell lines. Further evidence indicating that the OP5a AFP promoter can drive expression in a wide range of fish tissues comes from a recent study by Hobbs (2005) where it was demonstrated that this promoter drives expression of a type III AFP gene and a chimeric GH gene in most tissues of stable lines of transgenic salmon. This apparent lack of tissue specific expression may be because the factors interacting with the *cis*-acting regions exist generally in most cell types or, it may be that the specificity of the regulatory regions may be low with regard to the requirements of the *trans*-acting factors (Gong *et al.*, 1991).

Transcription factors isolated from rat and human cell lines have been shown to be capable of binding to fish AFP gene promoters, even though AFP genes are restricted to certain teleost fish and are not present in mammals (Wu, 1994). Two possible explanations for this capability are as follows. One possibility is that an ancestral gene from which the AFP gene evolved may be present and still function in mammals, but has not been identified to date. This may be the case for the type II AFP genes which are homologous to the carbohydrate recognition domain of mammalian C-type lectins. In the sea raven (*Hemitripterus americanus*), the type II AFP are synthesized in the liver for

export to the blood. It has been suggested that a regulatory system involved in liver-specific gene expression has been conserved during the course of evolution.

Another possible explanation for the ability of mammalian transcription factors to bind to fish AFP promoters is the fact that transcription factors are highly conserved in eukaryotic systems. Evidence in support of this comes from the discovery that the mammalian HSP70 TATA, CAAAT, Sp1 binding, ATF binding and heat shock elements were functional in the fission yeast *Schizosaccharomyces pombe* (Prentice and Kingston, 1992). Other studies have defined analogous factors in rats, *Drosophila,* and various plant species suggesting that transcriptional factors are indeed "general" factors, required for expression of most, perhaps all, class II genes. Thus, the process of transcription initiation by RNA polymerase II is highly conserved among eukaryotic organisms, allowing for the experimental advantages offered by different organisms to be exploited to identify and define these factors (Hampsey, 1998, Masaki et al, 2001).

5. The ocean pout AFP gene promoter contains regions that can positively and negatively regulate transcription

The results of this study clearly demonstrate the presence of positive and negative regulatory elements within the OP5a AFP gene promoter that function in several cell lines including the fish cell line, thus confirming and extending the results of earlier studies by Gong et al. (1991) and Gong and Hew (1993).

In summary, the regions of the op5a AFP promoter -2115 to -1784, -1784 to -536, -467 to -363, -163 to -151, and -151 to -90 were found to have positive effects on

promoter activity in all three cell lines, suggesting that these sequences contain one or more *cis*-acting elements that can enhance gene expression. Region -536 to -467 had little or no effect on promoter activity in the mammalian cell lines. However it had a clear negative effect on the promoter activity in the CHSE-214 fish cell line. Promoter regions -363 to -268, -268 to -216, -216 to -163 appeared to contain *cis*-acting elements that repress gene expression because they had a negative effect on promoter activity in all three cell lines.

These results, for the most part, show that the ocean pout AFP gene promoter and its various elements can regulate gene expression similarly in both mammalian and fish cells. However the results obtained from the fish cell line clearly differ from those observed using the mammalian cells with regard to the region promoting the greatest expression of luciferase. These results suggest that the transcription factors in the mammalian and fish cells differ in their ability to bind to and control the various regions of the AFP gene promoter.

The earlier studies of the ocean pout OP5a AFP gene promoter demonstrated that the presence of the TATA box was essential for the promoter to function (Gong *et al.*, 1991, Gong and Hew, 1993). In addition, those authors found evidence for the presence of a strong repressor element between base pairs -500 and -185. However, the detailed analysis of this region conducted in the present study reveals that, in addition to repressor elements, there is an enhancer element located between base pairs -467 and -363, thus further defining the regulatory regions of the promoter.

Although a number of transgenic strains of fish have been generated using gene constructs that are driven by the OP5a AFP promoter, a detailed analysis of the genomically integrated transgene has only been carried out on a line of transgenic salmon that was created using the opAFP-GHc2 gene construct (Figure 3A). This line of salmon, which are now in their sixth generation, grow up to 10 times faster than non-transgenics during their first 6 month of feeding and reach a market size of 4 to 6 kg a year earlier than non-transgenics cultured by commercial ventures in Atlantic Canada (Fletcher *et al.*, 2001, 2004).

The genomically integrated transgene (EO-1α) believed to be responsible for this rapid growth has been fully characterized to reveal that the opAFP-GHc2 construct was reorganized so that the first 1579 bp were removed from the 5' region of the promoter and effectively relocated downstream of the ocean pout AFP gene 3' flanking region (see Figure 3B) (personal communication Yaskowiak and Fletcher). However, despite the loss of 1579 bp from the 5' region of the promoter, it is apparent that the remaining 536 bp regions is capable of driving gene expression. In addition, it is also apparent that the presence of the same 1579 plus a second copy of base pairs 1580 to 1678 downstream of the 3' flanking did not increase the luciferase expression to that of the intact promoter. This result indicates that not all of the positive elements that were evident in this region when it was located upstream of the promoter were able to function when placed downstream of the functional gene. Since enhancers function independently of orientation and location, and can be located within a functional gene, upstream or downstream from the start of transcription (Latchman, 1998), it is apparent that there are

no major enhancer elements within the first 1678 bp of the promoter. This suggests that in order to function effectively, this region of the promoter must be located upstream of the functional gene. Another explanation for the low luciferase activity in the reorganized promoter construct could be that there is a repressor region within base pairs 1580 and 1678. Since there are two copies of this region in the reorganized promoter construct this could account for the lowered luciferase expression.

One of objectives of this study was to determine the efficacy of the OP5a AFP promoter in a wide range of cells and to determine whether or not it contained elements that would facilitate tailoring to drive relatively strong or weak expression of a transgene. It is evident from the results that the OP5a promoter can be used to drive gene expression in mammalian as well as fish cells. In addition, it is clear that the removal of a large section of the 5' end of the promoter greatly reduces its functional ability and results in a relatively weak promoter. When the structural gene is a hormone, a weak promoter would appear to be the desirable outcome. Over-expression of a protein such as a somatotropin is likely to produce undesirable effects resulting in severe bone irregularities that accompany growth (Devlin *et al.*, 1995b). Therefore, in order to reduce the possibility of gene over-expression it may be advisable to truncate the OP5a AFP promoter prior to gene transfer. In other situations, it may be desirable to design a stronger promoter to maximize expression. In the case of antifreeze protein gene transfer, it is essential to have a strong promoter in order to produce the high blood plasma concentration of AFP that is necessary to protect the fish from freezing (Fletcher et al., 1985). The results of this study indicate that in order to maximize expression it

would be preferable to transfer a full-length promoter, or a promoter that is truncated down to approximately 160 bp. Given the evidence that the full length promoter was considerably truncated in the EO-1α growth hormone transgenic salmon, it may be of value to test the *in vivo* efficacy of the more powerful short promoter.

Conclusion

The major objectives of this study were to characterize the type III ocean pout AFP gene promoter and to assess the activity of the rearranged opAFP-GHc2 using various cultured mammalian and fish cell lines *in vitro*. My data demonstrated that the ocean pout antifreeze gene promoter is functional in both mammalian and fish cell lines. The dissection of the ocean pout AFP gene promoter determined that this promoter consisted of both positive and negative regulatory regions. In summary regions -2115 to -1784, -1784 to -536, -467 to -363, -163 to -151, and -151 to -90 were found to have positive effects on promoter activity while promoter regions -363 to -268, -268 to -216, -216 to -163 had a negative effect on promoter activity in all three cell lines. It was also determined that the rearranged promoter found in the transgenic Atlantic salmon did retain promoter activity, however at a much lower level than the full length intact promoter sequence.

This study has defined a moderately strong promoter with several regulatory sequences. One of the applications of this study would be to design a stronger/weaker promoter to drive gene expression. It is sometimes beneficial to down-regulate or up-regulate gene expression and the analysis of a gene promoter to locate the various

positive and negative elements may allow one to manipulate the promoter to reduce or enhance gene expression.

References

Arnosti, D. 2003. Analysis and function of transcriptional regulatory elements: insights from *Drosophila*. *Annual Review of Entomology* 48:579-602.

Arkhipova, I. 1995. Promoter elements in *Drosophila melanogaster* revealed by sequence analysis. *Genetics* 139:1359-1369.

Bi, W., Wu, L., Coustry, F., de Crombrugghe, B. and Maity, S. 1997. DNA binding specificity of the CCAAT binding factor CBF/NF-Y. *Journal of Biochemistry* 272; 26562-26572.

Breathnach, R. and Chambon, P. 1981. Organization and expression of eukaryotic split genes coding for proteins. *Annual Review of Biochemistry* 50:349-383.

Chan, SL., Fletcher, GL. and Hew, CL. 1993. Control of antifreeze protein gene expression in winter flounder. In *Biochemistry and Molecular Biology of Fishes*, edited by Hochachka, PW and Mommsen, TP. Amsterdam:Elsevier 2: 293-305.

Chang, M., Nie, P., Sun, B. and Yao, W. 2005. Molecular cloning of TRAF2 binding protein gene and its promoter region from the grass carp *Ctenopharyngodon idellus*. Veterinary Immunology and Immunopathology 105: 105-113.

Chen, Z., Schuler, M. and Beachy, R. 1986. Functional analysis of regulatory elements in a plant embryo-specific gene. *Proceedings of the National Academy of Sciences of the USA* 83: 8560-8564.

Chen, J., Chou, M., Gong, H., Huang, T., Wu, J. and Kuo, C. 2005. Cloning and Biological analysis of the zebrafish (*Danio rerio)* insulin-like growth factor Binding protein-2 proximal promoter reigon. *DNA Cell Biology 24: 199-208.*

Chong, K., Wang, S. and Melamed, P. 2004. Isolation and characterization of the follicle-stimulating hormone β subunit gene and 5' flanking regions of the chinook salmon. *Neuroendocrinology* 80:158-170.

Conaway, JW. and Conaway, RC. 1989. Initiation of eukaryotic messenger RNA Synthesis. *Journal of Biological Chemistry* 266: 17721-17724.

Conaway, RC. and Conaway, JW. 1993. General initiation factors for RNA polymerase II. *Annual Review of Biochemistry* 62: 161-190.

Courey, AJ., Holtzman, DA., Jackson, SP. and Tjian, R. 1989. Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. *Cell* 59: 827-836.

Davies, P., Baardsness, J., Kuiper, M. and Walker, V. 2002. Structure and function of antifreeze proteins. *Philosophical Transactions of the Royal Society of London.* 357: 927-935.

De Clercq, N., Hemschoote, K., Devos, A., Peters, B., Heyns, W. and Rombauts, W. 1992. The 4.4-Kilodalton Proline-rich Polypeptides of the Rat Ventral Prostate are the proteolytic products of a 637- kilodalton Protein Displaying Highly repetitive sequences and encoded in a Single Exon. *Journal of Biological Chemistry* 267: 9884-9894.

Devlin, R., Yesaki, T., Donaldson, E., Du, S. and Hew, C. 1995a. Production of germline transgenic Pacific salmonids with dramatically increased growth performance. *Canadian Journal of Fisheries and Aquatic Sciences* 52: 1376-1384.

Devlin, R., Yesaki, T., Donaldson, E. and Hew, C. 1995b. Transmission and phenotypic Effects of an antifreeze/GH gene construct in coho salmon (Oncorhynchus kisutch). *Aquaculture* 137: 161-169.

Dillon, N., Spencer, V. and Butterworth, P. 1985. Transcription of a cloned rainbow trout protamine gene is accurately initiated following transfection into HeLa cells but the majority of the transcripts fail to polyadenylate at the correct site. *Nucleic Acids Research* 13: 8715-8727.

Du, SJ., Gong, Z., Fletcher, GL., Shears, MA. and Hew, CL. 1992. Growth hormone gene transfer in Atlantic Salmon: Use of fish antifreeze/ growth hormone chimeric gene construct. In *Transgenic Fish.* Edited by Hew, CL. and Fletcher, GL. World Scientific Publishing Co. Singapore. 1: 190-208.

Du, M., Beatty, L., Zhou, W., Lew, J., Schoenherr C, Weksberg, R and Sadowski, R. 2003. Insulator and silencer sequences in the imprinted region of human chromosome 11p15.5. *Human Molecular Genetics.* 12:1927-1939.

Dunham, R., Chitmanat, C., Nichols, A., Argue, B., Powers, D. and Chen, T. 1999. Predator avoidance of transgenic channel catfish containing salmonid growth hormone genes. *Marine Biotechnology* 1: 545-551.

Fletcher, GL., Hew, CL., Li, X., Haya, K. and Kao, M. 1985. Year round presence of high levels of plasma antifreeze peptides in a temperate fish: ocean pout *Macrozoarces americanus. Canadian Journal of Zoology* 63: 488-493.

Fletcher, GL., Hew, CL. and Davies, PL. 2001. Antifreeze proteins of teleost fishes. *Annual Review of Physiology* 63: 359-390.

Fletcher, GL., Shears, MA., Yaskowiak., ES. 2004. Gene transfer: potential to enhance The genome of Atlantic salmon for aquaculture. *Australian Journal of Experimental Agriculture* 44: 1095-1100.

Flores, D., Maldonado, E. and Reinberg, D. 1989. Factors involved in specific transcription by mammalian RNA polymerase II. Factors IIE and IIF independently interact with RNA polymerase II. *Journal of Biological Chemistry* 264: 8913-8921.

Goddard, SV. and Fletcher, GL. 2002. Physiological ecology of antifreeze proteins- A northern perspective. In *Molecular Aspects of Fish and Marine Biology: Fish Antifreeze Proteins*. Edited by Ewart, K. and Hew, CL. World Scientific Publishing Co. Singapore. 1: 17-60.

Gong, Z., Hew, CL. and Vielkind, J. 1991. Functional analysis and temporal expression of promoter regions from fish antifreeze protein genes in transgenic Japanese medaka embryos. *Molecular Marine Biology and Biochemistry* 1: 64-72.

Gong, Z., Fletcher, GL. and Hew, CL. 1992. Tissue distribution of fish antifreeze protein mRNAs. *Canadian Journal of Zoology* 70: 810-814.

Gong, Z. and Hew, CL. 1993. Promoter analysis of fish antifreeze protein genes. In *Biochemistry and molecular biology of fishes*. Edited by Hochachka, PW and Mommsen, TP. Elsevier Science Publishers B.V. Amsterdam; New York. 2: 307-324.

Gdula, D., Gerasimova, T. 1989. Novel genes influencing the expresión of the *yellow* locus and mdg^4 (gypsy) in *Drospholia melanogaste. Molecular Genetics.* 220: 121-126.

Hampsey, M. 1998. Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbial Molecular Biology Reviews* 62: 465- 503.

Hatamochi, A., Golumbek, P., Van Schaftingen, E. and de Crombrugghe, B. 1988. A CCAAT DNA binding factor consisting of two different components that both required for DNA binding. *Journal of Biological Chemistry* 263: 5940-5947

He, X., Futerer, J. and Hohn, T. 2002. Contribution of downstream promoter elements to transcriptional regulation of the rice tungro bacilliform virus promoter. *Nucleic Acid Research* 30: 497-506.

Hew, CL., Slaughter, D., Joshi, SB., Fletcher, GL. and Ananthanarayanan, VA. 1984. Antifreeze polypeptides from the Newfoundland ocean pout, *Macrozoarces americanus*: presence of multiple and compositionally diverse components. *Journal of Comparative Physiology* 155: 81-88.

Hew, CL., Wang, N., Joshi, S., Fletcher, GL., Scott, G., Hayes, P., Buettner, B. and Davies, PL. 1988. Multiple genes provide the basis for Antifreeze protein diversity and dosage in the ocean pout, *Macrozoaces americanus*. *Journal of Biological Chemistry* 263: 12049-12055.

Hew, CL., Du, SJ., Gong, Z., Shears, MA., King, MK., Fletcher, GL., Davies, PL. and Saunders, R. 1998. Use of the fish antifreeze protein gene promoter in the production of growth hormone-transgenic salmon with enhanced growth performance. In *Agricultural Biotechnology*. Edited by Altman, A. Marcel Dekker, New York. 549-561.

Hew, CL., Poon, R., Xiong, F., Gauthier, S., Shears, MA., King, MK., Davies, PL. and Fletcher, GL. 1999. Liver specific and seasonal expression of transgenic Atlantic salmon harbouring the winter flounder antifreeze protein gene. *Transgenic Research* 8: 405-414.

Hobbs, R. 2005. The ocean pout (*Macrozoarces americanus)* antifreeze protein gene Promoter drives the expression of different genes in transgenic Atlantic salmon (*Salmo salar)*. Masters thesis, Department of Biology, Memorial University of Newfoundland.

Ishihara, K. and Sasaki, H. 2002. An evolutionary conserved putative insulator element near the 3' boundary of the imprinted Igf2/H19 domain. *Human Molecular Genetics* 11: 1627-1636.

Jankowski, J. and Dixon, G. 1987. The GC as a silencer. *Bioscience Reports*. 7: 955-963.

Ju, B., Xu, Y., He, J., Liao, J., Yan, T., Hew, C., Lam, T. and Gong, Z. 1999. Faithful expression of green fluorescent protein (GFP) in transgenic zebrafish embryos under zebrafish gene promoters. *Developmental Genetics* 25:158-167.

Kadonaga, J., Carner, K., Masiarz, F. and Tian, R. 1987. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* 51: 1079-1090.

Kraus, R., Murray, E., Wiley, S., Zink, N,. Loritz, K., Gelembiuk., G. and Mertz, J. 1995. Experimentally determined weight matrix definitions of the initiator and TBP binding site elements of promoter. *Nucleic Acids Research* 24: 1531- 1540.

Latchman, D. 1998. Eukaryotic transcription factors 4th edition. London. Elsevier.

Lewis, B., Kim, T. and Orkin, S. 2000. A downstream element in the human β-globin promoter: Evidence of extended sequence-specific tramscription factor IID contacts. *Proceedings of the National Academy of Sciences of the USA*. 97: 7172-7177.

Lobachev, K., Shor, B., Tran, H., Taylor, W., Keen, J., Resnick, M. and Gordenin, D. 1998. Factors affecting inverted repeat stimulation of recombination and deletion in *Saccharomyces cerevisiae*. *Genetics* 148: 1507-1524.

Maclean, N., Rahman, M., Sohm, F., Hwang, G., Lyenger, A., Ayad, H., Smith, A. and Farahmand, H. 2002. Transgenic tilapia and the tilapia genome. *Gene* 295: 265-277

Martinez, R., Estrada M. and Berlanga, J. 1996. Growth enhancement in transgenic tilapia by ectopic expression of tilapia growth hormone. *Molecular marine biology and biotechnology* 5: 62-70.

Masaki, I., Satoshi, A., Sachihiro, M., Takashi I., Ryuichi, N., Yasunori, M., John D. and Akira, W. 2001. G2/M-Phase–Specific Transcription during the Plant Cell Cycle Is Mediated by c-Myb–Like Transcription Factors. *Plant Cell.* 13: 1891-1906.

McKnight, S. 1982. The nucleotide sequence and transcript map of the herpes simplex virus thymidine kinase. *Nucleic Acids Research* 8: 5949-5964.

McKnight, S. and Tjian, R. 1986. Transcriptional selectivity of viral genes in mammalian cells. *Cell* 46: 795-805.

Moncollin, V., Roy, RDW. and Egly, JM. 1994. The TATA saga: structure and function of the TATA-binding factor. In *Transcription Mechanisms and Regulation*. Edited by Conaway, RC and Conaway, JW. Raven Press 3: 45-62.

Nakajima, N., Horikoshi, M. and Roeder, R.G. 1988. Factors involved in specific Transcription by mammalian RNA polymerase II: purification, genetic specificity, and TATA box- promoter interactions of TFIID. *Molecular and Cellular Biology* 8: 4028-4040.

Nam, Y., Noh, J., Cho, Y., Cho, H., Cho, K., Kim, C. and Kim, D. 2001. Dramatically accelerated growth and extraordinary gigantism of transgenic mud loach *Misgurnus mizolepis*. *Transgenic Research* 10: 353-362.

Ostapchuk, P., Diffley, F., Bruder, J., Stillman, B., Levine, A. and Hearing. 1986. Interaction of a nuclear factor with the polyomavirus enhancer region. *Proceedings of the National Academy of Science of the USA.*. 83: 8550-8554.

Ow, D., Wood, K., DeLuca, M., De Wet, J., Helinski, D. and Howell, S. 1986. Transient and stable expression of the firefly luciferase gene in plant cells and transgenic plants. *Science* 234: 856-859.

Ozer, J., Faber, M., Chalkley, R. and Sealy, L. 1990. Isolation and characterization of a cDNA for the CCAAT Transcription Factor EFI_A reveals a novel structural motif. *Journal of Biological Chemistry* 265: 22143-22152.

Prentice, H. and Kingston, R. 1992. Mammalian promoter element function in the fission yeast *Schizosaccharomyces pombe*. *Nucleic Acids Research* 20: 3383-3390.

Samuels, M. and Sharp, P. 1986. Purification and characterization of a specific RNA polymerase II transcription factor. *Journal of Biological Chemistry* 261: 2003 2013.

Sawadogo, M. 1990. RNA polymerase B (II) and general transcription factors. *Annual Review of Biochemistry* 59: 711-54 .

Smale, S. 1994. Core promoter architecture for eukaryotic protein-coding genes. In *Transcription Mechanisms and Regulation*. Edited by Conaway, RC and Conaway, JW. Raven Press. 3: 63-81.

Smale, S. and Kadonaga, J. 2003. The RNA polymerase II core promoter. *Annual Review of Biochemistry* 72: 449-479.

Song, J., Ugai, H., Wang, Y., Sarai, A., Obataa, Y., Kanazawa, I., Sun, Kailai, Itakura, K. and Yokoyama, K. 2001. Two consecutive zinc fingers for interactions with *cis*- elements. *Journal of Biological Chemistry* 276: 30429-30434.

Gerasimova, T. and Corces, V. 1998. Polycomb and Trithorax Group Proteins Mediate the function of a chromatin insulator. *Cell.* 92: 511-521.

Uptain, S., Kane, C. and Chamberlin, M. 1997. Basic mechanisms of transcript elongation and its regulation. *Annual Review of Biochemistry* 66: 117-172.

Villalobos, P., Rogas, M., Conejeros, P., Marshall, S. 1999. Lipopolyamine-mediated transfection of reporter plasmids into a fish cell line. *Electronic Journal of Biotechnology* 2

Weimin, B., Coustry, F., Crombrugghe, B. and Maity, S. 1997. DNA binding specificity Of the CCAAT- binding factor CBF/NF-Y. *Journal of Biological Chemistry* 272: 26562-26572.

West, A., Gaszner, M. and Felsenfeld, G. 2002. Insulators: many functions, many mechanisms. *Genes and Development* 16: 271-288.

Wood, K. 1990. Firefly luciferase: a new tool for molecular biologists. Promega Notes 28: 1.

Wu, Y. 1994. Characterization of an antifreeze protein gene promoter for use in transgenic fish. PhD thesis Department of Biochemistry, Queens University

Zafarullah, M., Bonham, K. and Gedamu, L. 1988. Structure of the rainbow trout metallothionein B gene and characterization of its metal-responsive region. *Molecular and Cellular Biology* 8: 4469-4476.

Zawel, L. and Reinberg, D. 1992. Advances in RNA polymerase II transcription. *Current Opinions in Cell Biology* 4: 488-495.

Zawel, L. and Reinberg, D. 1995. Common themes in assembly and function of eukarytic transcription complexes. *Annual Review of Biochemistry* 64: 533-561.

Zbikowska, HM. 2003. Fish can be first- advances to fish transgenesis for applications. *Transgenic Research* 12: 379-389.

Zhanjiang, L., Moav, B., Faras, A., Guise, K., Kapuscinski, A. and Hacktt, P. 1990. Functional analysis of elements affecting expression of the β-actin gene of carp. *Molecular and Cellular Biology* 10: 3432-3440.