Expression and Function of Heat Shock Factors in Zebrafish (*Danio rerio*)

A Thesis Submitted to the College of
Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in the Department of Anatomy and Cell Biology
University of Saskatchewan
Saskatoon

By

Cynthia Lynn Swan

PERMISSION TO USE

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Anatomy and Cell Biology
University of Saskatchewan
Saskatoon, Saskatchewan
S7N 5E5 CANADA

ABSTRACT

Heat shock proteins (hsp) and heat shock transcription factors (HSF) have important roles in the development of the eye lens. Our lab previously demonstrated that knockdown of *hsp70* gene expression using morpholino antisense technology (MO) resulted in a small lens phenotype in zebrafish (*Danio rerio*) embryos. A less severe phenotype was seen with knockdown of *hsf*1, suggesting other factors that regulate *hsp70* are involved during lens formation. Both HSF1 and HSF4 are known to play a role in mammalian lens development. An expressed sequence tag encoding zebrafish HSF4, named *hsf4a*, has been identified and a second splice variant, *hsf4b*, has been predicted in the Ensembl database. The objectives of this thesis were to characterize the zebrafish HSF4 and compare its expression to other HSFs as well as investigate its role in lens development. Analysis of zebrafish HSF4 sequence was performed using standard *in silico* analytical software. The deduced amino acid sequence of HSF4a shares structural similarities with mammalian HSF4 including the lack of an HR-C domain. This domain is absent due to a C-terminal truncation within zebrafish HSF4a relative to the mammalian protein. HSF4b is identical to the HSF4a sequence with the exception of an additional 155 amino acids at the carboxyl end of the protein which contains an HR-C domain, unlike mammalian HSF4. Surprisingly, electrophoretic mobility shift assays (EMSA) demonstrated that the binding affinity of zebrafish HSF4 to discontinuous HSEs is more similar to HSF1 than to other HSF4 proteins. The amino acid sequence of zebrafish HSF4 DNA binding domain was also more similar to HSF1 than other HSF4 proteins. These results, along with a phylogenetic analysis of HSF proteins from eleven species, suggest that HSF1 was an evolutionary precursor of HSF4 and that functions of this protein may differ between zebrafish and mammals. The expression level for each of the three zebrafish HSFs was determined in adult tissues and in developing embryos by quantitative reverse transcription polymerase chain reaction (qPCR) analysis. Expression of both *hsf4* transcripts was observed predominantly in the eye but only observed in developing embryonic tissue at 60 hours post fertilization or later. This, together with the lack of an observable phenotype following MO knockdown of *hsf4*, suggests that HSF4 likely has a role in later stages of lens development. Additionally, *hsf1* and *hsf2* expression were detected in all tissues and in all stages of development as well as being present as maternal transcripts in zebrafish eggs. The results presented in this thesis demonstrate that while zebrafish HSFs share

some similarity with HSF proteins from other species, they also have structural characteristics and expression patterns unique to the zebrafish.

ACKNOWLEDGEMENTS

There are many people who I would like to acknowledge for their assistance during my PhD program and ultimately in the production of this thesis. Many thanks go to the members of my committee, Dr. Bill Roesler, Dr. Bill Kulyk, Dr. Nick Ovsenek, and Dr. Helen Nichol. Whether it was providing equipment, lab space, reference material or advice and encouragement, all of my committee members provided assistance at some point in my program and I very grateful for their support. I would also like to thank Dr. Marc Ekker for agreeing to be my external examiner and accommodating the defense date into his busy travel schedule. I would like to thank all of the above for their thoughtful comments and insights into this dissertation.

I wish to thank all my fellow students and colleagues for advice, camaraderie, and generally being sounding boards when problems arose – Amy McKay, Zach Belak, Carlyn Matz, Lindsay Jacobi, Kate Martelli, Tracy MacDonald, Ashley James, Gosia Korbas and Nicole Sylvain – thanks so much.

I would also like to extend my thanks to the members of the Department of Anatomy and Cell Biology, including Dr. David Schreyer, all of whom were supportive during my program especially during all the bumps and hiccups that occurred along the way.

My sincerest thanks go to my supervisor and mentor Dr. Patrick Krone. I would like to thank him for introducing me to the zebrafish, for sharing his knowledge of all things heat shock with me, as well allowing me latitude to pursue some of my own ideas and engaging me in discussions about them. I appreciate his support, encouragement and especially his patience and understanding throughout my program, which I am sure he sometimes thought would never come to an end!

Of course, I would also like to thank my husband, mom, sister, and the rest of my family and friends for their support and understanding throughout the years.

DEDICATION

I would like dedicate this thesis to my husband Gordon Gray.
The path to achieving my goal of obtaining a PhD was not direct and included several unexpected twists and turns that affected both our lives. Together, we endured and overcame many obstacles to allow me to reach this point. You supported and encouraged me through the entire journey (whether I liked it or not) and I love you for that.
Thank you for all your inspiration, love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

°C	degrees Celsius
µCi	microCurrie
µg	microgram
µL	microlitre
µM	micromolar
5bpmm	5 base pair mismatch
ATP	adenosine triphosphate
APZ	anterior proliferative zone
AZF	azoospermic factor interval
BCIP	5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt
Bfsp	lens specific beaded filaments
BLAST	Basic Local Alignment Search Tool
bp	base pair
Brg1	Brahma-related gene
BSA	bovine serum albumin
CBF	CCAAT box binding factor
cDNA	complementary DNA
CDY2	testis-specific chromodomain protein Y
ChIP	chromatin immunoprecipitation
cps	counts per second
C_T	critical threshold
CTP	cytidine triphosphate
CV	coefficient of variation
Da	dalton
dATP	deoxyadenosine triphosphate
DAXX	FAS death-domain associated protein
DBD	DNA binding domain
dCTP	deoxycytidine triphosphate
DEPC	diethylpyrocarbonate

dGTP	deoxyguanosine triphosphate
dI-dC	deoxyinosine-deoxycytosine
DIG	digoxigenin
DLAD	DNAse 2β
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
dpf	days post fertilization
ds	double stranded
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
EDTA	ethylenediamine tetraacetic acid
EGTA	ethylene glycol tetraacetic acid
EMSA	electrophoretic mobility shift assay
ERK	extracellular signal-regulated kinase
EST	expressed sequence tag
EtOH	ethanol
fgf	fibroblastic growth factor
GTP	guanosine triphosphate
HEPES	2-[4-(2-hydroxyethyl) piperazin-1-yl] ethanesulfonic acid
His	histidine
hpf	hours post fertilization
HR-A/B	heptad repeat (A/B)
HR-C	heptad repeat (C)
HS	heat shock
HSE	heat shock response element
HSF	heat shock factor
hsp	heat shock protein
I.M.A.G.E.	Integrated Molecular Analysis of Genomes and their Expression
IMAC	immobilized metal ion affinity chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
KCl	potassium chloride

kDa	kilodalton
KH_2PO_4	monopotassium phosphate
LB	Luria-Bertani
LiCl	lithium chloride
Lif	leukemia inhibitory factor
LW-1	alternate acronym for heat shock factor X
MAP	mitogen activated protein kinase
MEF	mouse embryo fibroblasts
MeOH	methanol
mg	milligram
MGC	Mammalian Gene Collection
$MgCl_2$	magnesium chloride
$MgSO_4$	magnesium sulphate
MIQE	*M*inimum *I*nformation for publication of *Q*uantitative real-time PCR *E*xperiments
mL	millilitre
mm	millimetre
mM	millimolar
MO	morpholino
MOPS	3-(N-morpholino) propanesulfonic acid
mRNA	messenger RNA
MSY	male specific region of the Y chromosome
Na_2HPO_4	disodium hydrogen phosphate
NaAc	sodium acetate
NaCl	sodium chloride
NBT	nitrotetrazolium blue chloride
NCBI	National Center for Biotechnology Information
ng	nanogram
nL	nanolitre
nmol	nanomole
nM	nanomolar

nm	nanometre
NRC	National Research Council
NTC	non-template control
nt	nucleotides
OD	optical density
OFZ	organelle free zone
ORF	open reading frame
p	probability
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PBST	phosphate buffered saline + Tween20
PCR	polymerase chain reaction
PFA	paraformaldehyde
pmol	picomole
PMSF	phenylmethylsulfonyl fluoride
PROM2	prominin membrane glycoprotein
psi	pounds per square inch
PTM	post translational modification
PVDF	polyvinylidene difluoride
qPCR	quantitative real time reverse transcription polymerase chain reaction
REST	Relative Expression Software Tool
RNA	ribonucleic acid
RNAi	RNA interference
ROX	5-carboxy-X-rhodamine, succinimidyl ester
rpm	revolution per minute
RT-PCR	reverse transcription polymerase chain reaction
S	Svedberg unit
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SIRT1	sirtuin 1

SKAP2	src kinase-associated phosphoprotein 2
SMART	Simple Modular Architecture Research Tool
SOB	super optimal broth
ss	single stranded
SYBR	N',N'-dimethyl-N-[4-[(E)-(3-methyl-1,3-benzothiazol-2-ylidene) methyl]-1-phenylquinolin-1-ium-2-yl]-N-propylpropane-1,3-diamine
TAE	Tris-acetate-EDTA
TEMED	tetramethylethylenediamine
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
tRNA	transfer RNA
TBS	Tris-buffered saline
TTBS	Tris-buffered saline + Tween 20
TUNEL	terminal deoxynucleotidyl transferase dUTP nick end labeling
Tween 20	polyoxyethylene (20) sorbitan monolaurate
U	international unit
UTP	uridine triphosphate
UTR	untranslated region
×g	times gravity
Yp	short arm of chromosome Y
Yq	long arm of chromosome Y
ZFIN	the zebrafish model organism database

1.0 INTRODUCTION AND BACKGROUND

1.1 Introduction to the heat shock response

Heat shock proteins (hsps) were initially discovered as proteins produced by an organism in response to heat stress and were therefore named accordingly. Investigations into the heat shock response began in 1962 when Ferruccio Ritossa discovered an unusual pattern of puffs on the salivary gland chromosomes of *Drosophila* (Ritossa 1962). It was not until 1973 that molecular analysis of these chromosomal puffs revealed that the synthesis of new proteins, later determined to be heat shock proteins, occurred at the same time as the puffs were observed. By the late 1970s and early 1980s research on several organisms had shown that this response to heat stress was conserved in a large variety of organisms from prokaryotes such as *Escherichia coli* (*E. coli*), to all types of plants, *Drosophila* and in higher eukaryotic animals (Lindquist 1986). It was also demonstrated that these proteins were synthesized in response to other cellular stresses including exposure to ethanol, heavy metals and inflammation (Lindquist and Craig 1988). Heat shock proteins are one of the most highly conserved families of proteins identified to date (Lindquist and Craig 1988; Rupik et al. 2011). Hsps have been grouped into gene families, including Hsp70, Hsp90, Hsp47 and the small hsps, based on their sequence homology and molecular mass which can range from 10-100 kDa (Rupik et al. 2011).

While expression of several hsp gene families is upregulated in response to stresses, other members of hsp families, such as heat shock cognate proteins, are also constitutively expressed in tissues. These proteins act as molecular chaperones which assist with the folding of newly synthesized proteins within the cells of an organism (Walsh et al. 1997). However, when heat shock or other stresses occur, cellular proteins begin to denature and misfolding of proteins causes them to aggregate and become inactive. Under stress conditions the amount of misfolded proteins that accumulate in cells can overwhelm the constitutive hsps. To overcome this, cells upregulate the expression of inducible heat shock genes. These hsps can bind to denatured proteins and prevent aggregation of the protein as well as assist in refolding them back into their native functional state (Walsh et al. 1997; Mathew and Morimoto 1998; Rupnik et al. 2011). In addition, hsps are involved in the packaging and transport of damaged proteins which have been targeted for degradation (Walsh et al. 1997; Rupnik et al. 2011). Binding of hsps to their target proteins occurs through the C-terminal portion of the chaperone with the assistance of co-chaperones and other cofactors. The refolding of target proteins requires ATP hydrolysis which,

for several families of hsps, is performed using the ATPase activity at the N-terminal domain of the hsp (Rupnik et al. 2011).

The genes encoding for hsps are regulated by transcription factors called heat shock factors (HSFs). These proteins are also highly conserved and have been identified in all plant and animal species studied. While over 20 HSFs exist in plants, only four have been characterized in the animal kingdom. Yeast and invertebrates such *Caenorhabditis elegans* and *Drosophila* have only a single HSF while multiple HSFs are found in vertebrates (Pirkkala et al. 2001; Åkerfelt et al. 2007; von Koskull-Döring et al. 2007). The first HSF identified in vertebrates is named HSF1 and functions as the primary regulator of heat shock genes during the heat shock response as well as having regulatory roles in development. Another HSF with a role in regulating heat shock genes in response to stress is HSF3, although so far this protein has only been identified in two vertebrate species (Björk and Sistonen 2010). The functions of HSF2 and HSF4 have been shown to be primarily related to development and do not directly regulate heat shock genes in response to stress (Pirkkala et al. 2001; Åkerfelt et al. 2007). Two other members of the HSF family have been identified, HSFY and HSF5, however very little is known about the function of these HSFs. The HSFs activate gene transcription by binding to a conserved DNA binding domain known as the heat shock response element (HSE) usually located in the 5' regulatory regions of the genes (Morimoto 1998; Åkerfelt et al. 2007). Several genes including *hsp70* have more than one HSE in the 5' regulatory regions (Åkerfelt et al. 2007). Activation of HSFs is complex, involves several different protein complexes, and requires multiple steps (Voellmy 2004; Björk and Sistonen 2010). Several of these steps and the proteins involved have yet to be elucidated. Regulation of HSF activation is discussed in more detail in section 1.3.6.

1.2 Heat shock protein families and development

While the hsp family of proteins was initially identified and characterized based on their roles during the cellular stress response, hsps also have many housekeeping roles within the cell that are unrelated to the stress response. These include assisting in the folding of newly synthesized proteins, protein degradation, cell signaling including steroid hormone activation, and antigen presentation. Several hsps are involved in mitotic spindle stabilization as well as cytoskeleton remodeling (Saito et al. 2009). Activation of regulatory proteins such as transcription factors (including HSFs), replication proteins and kinases involve hsps (Evans and Krone 2005). Hsps also have important roles in developmental processes such as tissue formation, growth, and survival. Many hsps have elevated expression levels during embryogenesis and patterns of hsp protein expression have been shown to be stage and tissue specific during development (Abane and Mezger 2010). The expression pattern of different hsp proteins during vertebrate development has been found to be very similar in several developmental models including mouse, *Xenopus* and zebrafish. During development, the formation of many embryonic structures, including the neural tube and cardiac tissue, depends on a regulated series of cell proliferation, differentiation and apoptotic events and the regulation of these processes has been found to be correlated to distinct expression patterns of hsp families (Rupik et al. 2011).

The Hsp70 family of proteins has been the most extensively studied and most organisms have several genes that encode members of this protein family. Hsp70 can be found in almost all tissue types at various developmental stages after heat stress but it also has clearly defined developmental roles in vertebrates in the absence of stress. For example, knockdown of *hsp70* disrupts blastocyst development in pre-implantation mouse embryos (Dix et al. 1998). Hsp70 has also been demonstrated to be critical for normal mammalian spermatogenesis as *hsp70* null mice lack post-meiotic spermatids and mature sperm (Dix et al. 1996). Other studies have demonstrated that Hsp70 has roles in apoptosis (Mosser et al. 2000; Awasthi and Wagner 2005), cell cycle regulation (Walsh et al. 1993) and chondrocyte differentiation (Trichilis and Wroblewski 1997; Mosser et al. 2000; Marshall and Harley 2001). In zebrafish embryos, Hsp70 is present at high levels specifically in the lens of the eye under non-stress conditions (Blechinger et al. 2002; Evans and Krone 2005). The high level of Hsp70 expression coincides with differentiation of fibre cells in the lens after which *hsp70* expression decreases (Krone et al.,

2003). Knockdown of *hsp70* expression using morpholino-modified antisense oligonucleotides resulted in fish with a small lens phenotype, demonstrating a requirement for Hsp70 in normal zebrafish eye development (Evans and Krone 2005).

The Hsp90 family consists of two genes that produce α and β forms of the protein in most vertebrates (Rupik et al. 2011). Like *hsp70*, they are inducible by heat stress but also have important roles in physiological processes under non-stress conditions. The role of Hsp90 in development was initially difficult to discern as complete loss of Hsp90 activity in knockout studies resulted in a lethal phenotype. Studies that used only a reduction in *hsp90* expression or that employed a specific Hsp90 inhibitor were instrumental in elucidating the roles Hsp90 played in development (Rupik et al. 2011).

Hsp90 proteins have roles in signal transduction and chromatin organization, functions that are unique to this family of hsps (Rupik et al. 2011). The involvement of Hsp90 in signal transduction has been demonstrated in a wide variety of organisms including yeast, *Drosophila*, several vertebrates and plants. For example, members of the Hsp90 family associate with steroid hormone receptors and are part of the mechanism that regulates receptor activity (Voellmy 2004). Hsp90 has been demonstrated to play a critical role in compound eye development as well as for spermatogenesis in *Drosophila* (Cutforth and Rubin 1994; van der Straten et al. 1997; Yue et al. 1999). In mammals, Hsp90β plays an essential role in placenta development (Voss et al. 2000). Inhibition of Hsp90 in geldanamycin treated zebrafish embryos resulted in a failure to form early somitic muscles cells called pioneer cells (Krone et al. 2003; Lele et al. 1999), and more recent studies have identified that the Hsp90α isoform is required for proper myofibril organization in skeletal muscles of zebrafish (Du et al. 2008; Etard et al. 2008; Hawkins et al. 2008). The protein is also important for neural cell differentiation, heart tissue morphogenesis and chondrocyte differentiation (Rupik et al. 2011). Another role for Hsp90 is regulating phenotypic variations in plants (Queitsch et al. 2002) and animals (Rutherford and Lindquist 1998; Sollars et al. 2003) as it has been demonstrated that inhibition of Hsp90 function allows previously concealed phenotypes to emerge. These phenotypes act as selectable traits that can be passed on to subsequent generations (Sollars et al. 2003).

Hsp47 is a heat inducible hsp that resides in the endoplasmic reticulum. Like other hsps it is also constitutively expressed during embryonic development, and is observed mainly in tissues derived from the mesoderm. Knockout of *hsp47* expression in mice results in a lethal phenotype

as offspring die *in utero*. The phenotype of knockout embryos includes a lack of mature α1(I)-collagen in mesenchymal tissues as well as disordered epithelia and ruptured blood vessels (Nagai et al. 2000). In developing chick and mouse embryos, *hsp47* is expressed in fibroblasts of connective tissue and in smooth muscle cells in the gastrointestinal tract, kidney epithelium and the basal lamina of the epithelium. Further, it has been demonstrated that Hsp47 has a critical role in the maturation of type I collagen and proper processing of type IV collagen in mice (Matsuoka et al. 2004; Nagai et al. 2000). Similarly, in early stages of zebrafish development, Hsp47 has been linked to the expression of type II collagen in the notochord. In later stages of zebrafish development, *hsp47* is also expressed in the otic capsule and fins (Lele et al. 1997; Krone et al. 2003).

The small heat shock proteins comprise a large family of proteins with varying ranges of molecular masses up to 45 kDa. One common characteristic of this diverse group of proteins is the fact that they all contain an α-crystallin domain similar to the crystallin proteins expressed in the lens of the eye (Mymrikov et al. 2011). While a large number of these proteins are known to have roles in preventing protein misfolding and aggregation during cell stress, the tightly regulated and time specific expression of many small hsps suggests an important role in development as well. The α-crystallins are specialized members of the small heat shock family and have an important role in the formation of the lens in the eye. Some members of the α-crystallin gene family have also been shown to be expressed developmentally in several non-ocular tissues including the muscle, heart and brain (Thonel et al. 2012). Transgenic mice that overexpressed the HspB8 (an ortholog to human Hsp22) protein showed significant hypertrophy of the heart (Mymrikov et al. 2011). The Hsp27 protein appears to be important for oogenesis and development of neural as well as the epidermal tissues (Thonel et al. 2012). The clarification of specific developmental roles for many small hsps is complicated by the large number and multiple isoforms that exist in vertebrates, and it has been suggested that the number of isoforms may be responsible for a certain amount of redundancy in small hsp function (Morrow and Tanguay 2012).

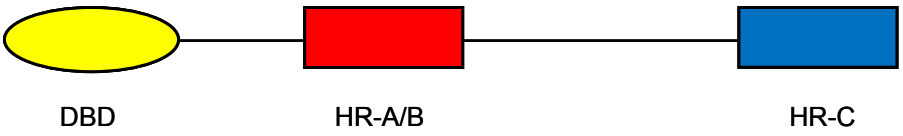
1.3 Heat shock transcriptions factors

Heat shock protein gene expression is regulated by a family of transcription factors called heat shock factors. These proteins regulate hsp gene expression in response to cellular stresses and also during different developmental stages in the absence of stress (Åkerfelt et al. 2007; Abane and Mezger 2010; Björk and Sistonen 2010; Fujimoto and Nakai 2010). Heat shock factors also regulate non-hsp genes such as those encoding fibroblast growth factors (*fgfs*) and *c-fos*, as well as having roles in chromatin modification and maintaining proteostasis (Morano and Thiele 1999; Pirkalla et al. 2001; Xing et al. 2005; Tu et al. 2006; Åkerfelt et al. 2007, 2010; Wilkerson et al. 2007; Fujimoto et al. 2008; Shi et al. 2009; Abane and Mezger 2010). Recently HSFs have been identified as potential targets for novel therapeutic agents targeting cancer and autoimmune diseases (Åkerfelt et al. 2007, 2010).

The amino acid sequences for HSFs are conserved across a range of organisms from single celled eukaryotes to mammals. There are several functional domains in HSF proteins that are also conserved between species (Fig. 1.1A) (Rabindran et al. 1993; Pirkalla et al. 2001). HSFs exist constitutively within cells in an inactive state as monomers and dimers. These proteins become active when cellular processes signal the monomers and dimers to trimerize, after which they localize to specific DNA sequences, HSEs, to activate target genes (Fig. 1.2) (Pirkalla et al. 2001; Björk and Sistonen 2010). This process of trimerization occurs through the action of several well characterized protein domains (Fig. 1.1). All HSFs contain a conserved amino-terminal winged helix-turn-helix DNA binding domain (DBD) capable of recognizing and binding to HSE DNA sequences consisting of a series of at least three inverted repeats of the sequence 5'-nGAAn-3' (Amin et al. 1988; Harrison et al. 1994; Morimoto 1998; Pirkkala et al. 2001).

Figure 1.1 Structure of HSF protein domains. (**A**) Three important protein domains have been identified and characterized in HSFs. The DNA binding domain (DBD) interacts with heat shock response element in the promoter of target genes. The trimerization domain (HR-A/B) is the region at which HSF monomers and dimers join to form an active HSF trimer. The inhibition of transactivation or trimerization domain (HR-C) is located in the carboxy portion of the protein. (**B**) The HR-C domain inhibits trimerization of the HR-A/B domain through interaction with a chaperone containing protein complex.

A



B

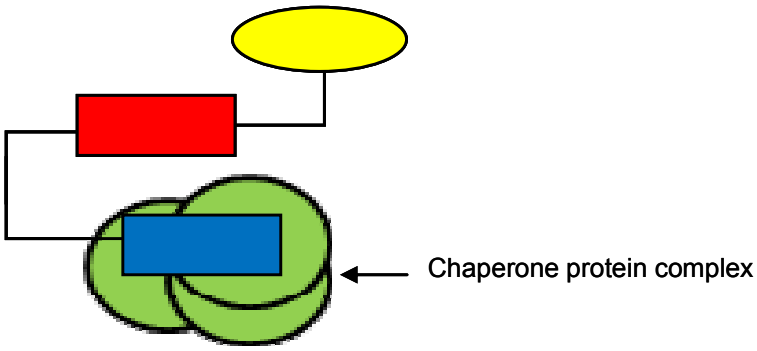
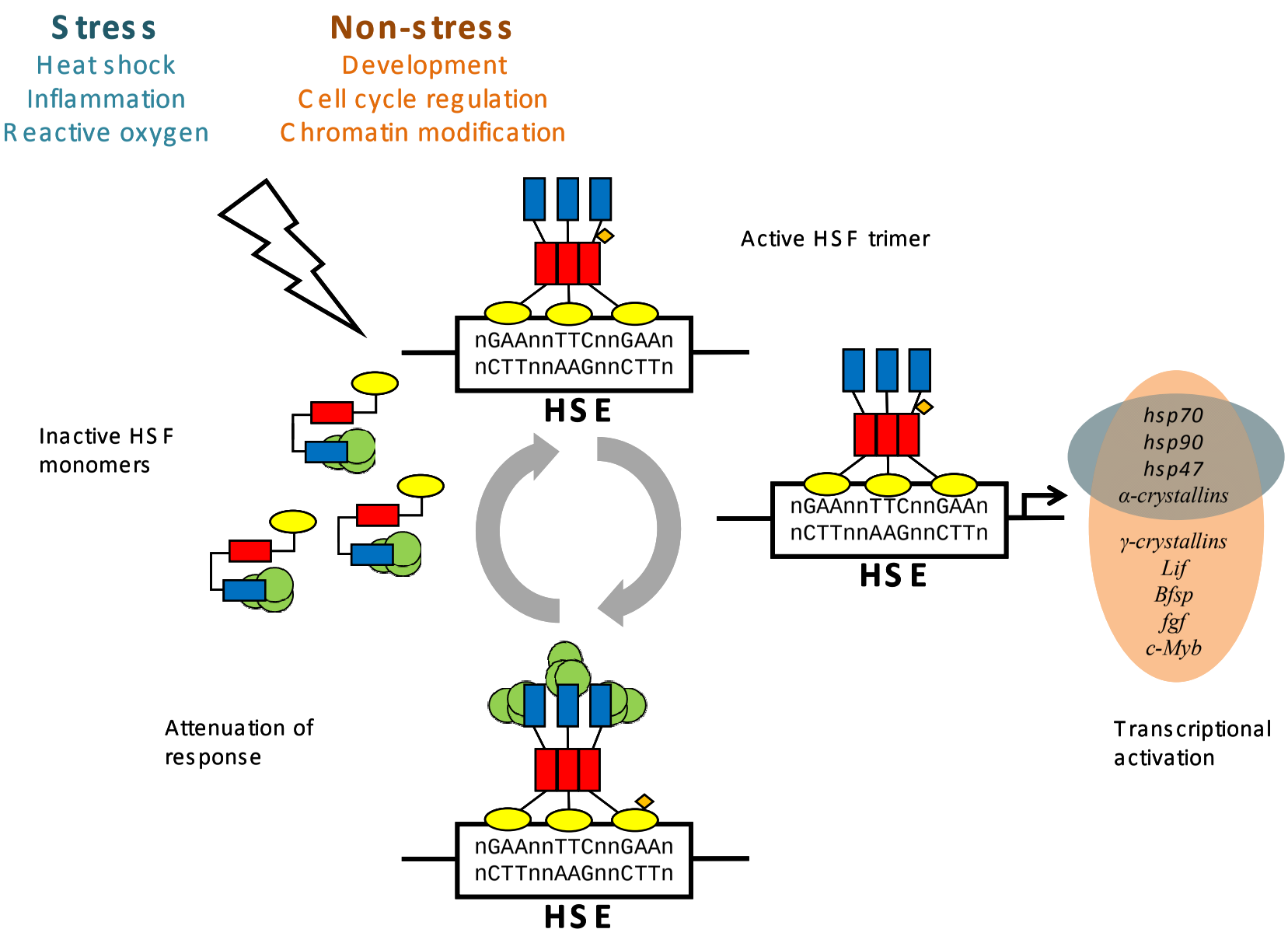


Figure 1.2 A schematic diagram of the process of HSF activation and attenuation of transcription. HSFs exist in the cell as inactive monomers or dimers (not shown). Protein complexes (green circles) which include the Hsp70 and Hsp90 chaperones, are bound to the inhibition of transactivation (HR-C) domain (blue rectangle) to keep the monomers and dimers inactive. Stress and non-stress cellular cues initiate the trimerization of HSFs at the trimerization (HR-A/B) domains (red rectangle) which involves the loss of the protein complex from the HR-C domain as well post-translational modifications (orange diamond). The trimer is recruited to the heat shock response element (HSE) DNA binding domain (DBD) where each DBD (yellow oval) associates with one of the GAA repeat sequences in the HSE. Transcriptional activation of target genes involved in either the stress response (shown in the blue oval) or non-stress responses (shown in the pink oval) is then initiated. The attenuation of HSF transcriptional activation requires the re-association with chaperone containing protein complexes as well as additional post-translational modifications in the DBD. HSFs then detach from the HSE and remain in the cell again as inactive monomers and dimers. Diagram is modified from Morimoto (1998) and Björk and Sistonen (2010).


Stress
Heat shock
Inflammation
Reactive oxygen
Non-stress
Development
Cell cycle regulation
Chromatin modification
Active HSF trimer
nGAAnnTTCnnGAAn
nCTTnnAAGnnCTTn
HSE
Inactive HSF
monomers
nGAAnnTTCnnGAAn
nCTTnnAAGnnCTTn
HSE
hsp70
hsp90
hsp47
α-crystallins
γ-crystallins
Lif
Bfsp
fgf
c-Myb
Attenuation of
response
nGAAnnTTCnnGAAn
nCTTnnAAGnnCTTn
HSE
Transcriptional
activation

The DBD consists of four antiparallel β sheets, a flexible 'wing' or loop of amino acids that extends outward between two of the β sheets and three helix bundles. The second and third helix bundles are what make up the helix turn helix motif (Sakurai and Enoki 2010). The nature of the DBD sequence results in a three dimensional structure where all of the GAA repeats end up on the same side of the helix motif. Typical HSEs have at least three 5'-nGAAn-3' repeats which allow each of the three DBD in an HSF trimer to each bind to a single GAA sequence (Fig. 1.2). The binding of the DBD to the GAA sequence occurs primarily as two hydrogen bonds between a single conserved arginine residue in the DBD and the guanine nucleotide of the GAA sequence (Sakurai and Enoki 2010). The 'wing' portion of the DBD does not appear to make contact with the DNA sequence (Sakurai and Enoki 2010).

The second conserved domain in the amino terminus of HSFs is the transactivation domain (HR-A/B) which is made up of two sub-domains, HR-A and HR-B, both consisting of hydrophobic heptad repeats characteristic of coiled coil structures. The HR-A/B domain is required for the trimerization of HSF monomers and dimers prior to DNA binding (Björk and Sistonen 2010; Sakurai and Enoki 2010). A third functional domain found in most HSFs also consists of a series of hydrophobic heptad repeats (HR-C), and is located at the carboxyl terminus of the protein. This region has been determined to interact with several other proteins and molecules, including Hsp70 and Hsp90 chaperones, for the purpose of inhibiting trimerization at the HR-A/B domain and thereby maintaining HSFs in their inactive form (Fig. 1.1 and 1.2) (Rabindran et al. 1993; Nakai and Morimoto 1993, 1998; Shi et al. 1998; Sakurai and Enoki 2010).

While only a single HSF gene has been identified in species such as yeast, *C. elegans* and *Drosophila*, four HSFs have been characterized in vertebrates and are named HSF1, HSF2, HSF3 and HSF4 in the order of their discovery. Two other members of the HSF family, HSFY and HSF5, have also been identified but little is known about the function of these proteins. Different isoforms of mammalian HSF1, HSF2 and HSF4 have been identified and are generated by alternative splicing of transcripts. Tissue and developmental specific expression patterns have been observed for some HSF although the regulation of their differential expression is poorly understood (Pirkkala et al. 2001). What is known to date about each of the HSFs and their regulation is summarized in the following sections.

1.3.1 Heat shock factor 1

In vertebrates, HSF1 was the first HSF identified and is known to be the primary regulator of heat shock genes following stress (Pirkalla et al. 2001; Björk and Sistonen 2010). For this reason it has been the most widely studied. Although HSF1 is the primary HSF that regulates the stress response in vertebrates, it also has several developmental roles including regulating genes involved in the formation of the brain and sensory placode (Åkerfelt et al. 2007, 2010; Abane and Mezger 2010). In zebrafish embryos, HSF1 has a direct role in development of a functional lens in the eye (Evans et al. 2007). A link between expression of *hsf1* and subsequent regulation of *hsp25*, *hsp90* and *αB-crystallin* expression was identified as being important for heart and kidney tissue development (Yan et al. 2002, 2005; Bierkamp et al. 2010).

Regulation of different physiological processes, including reproduction and the immune response, are also dependent on HSF1. For example, HSF1 binds to the promoter region of the interleukin-6 gene and a lack of HSF1 results in reduced proliferation of spleen cells in mice. This has an overall effect of impairing the immune response by limiting Ig production thus decreasing B cell response (Inouye et al. 2004). Several studies have identified HSF1 as an essential component for the development of a functional reproductive system in both male and female mice. In adult *hsf1* null mice, females are smaller, infertile, and display poor placental development while males have impaired spermatogenesis at the pachytene stage as well as an increased number of apoptotic spermatocytes (Xiao et al. 1999; Nakai et al. 2000; de Rooija and de Boerb 2003; Izu et al. 2004). As well, the viability of the oocytes appears to be linked to a reliance on HSF1 as a maternal factor during early cleavage in mammalian development (Christians et al. 2000). Recently it has also been determined that HSF1, along with HSF2, regulate X and Y linked multi-copy genes in mouse testis. These HSFs were found bound to these mutli-copy genes in the heterochromatin of post-meiotic cells during specific stages of spermatogenesis. This observation is intriguing as most multi-copy genes are silenced during meiosis at this stage of spermatogenesis. It has been suggested that the presence of HSFs may allow these specific multi-copy genes to continue to be expressed during meiosis as the expression of some of these genes are important for proper sperm differentiation and chromatin packing in male germ cells (Åkerfelt et al. 2010; Vihervaara et al. 2013).

The association of HSF1 with the sex chromosomes during a stage of meiosis when genes are normally silenced suggests that HSF1 could have roles in chromatin modulation.

Several recent studies have provided evidence for this possibility. HSF1 associates with histone deacetylase 1 and 2 both in the presence and absence of heat shock in a human cancer cell line (HeLa) to induce a large scale deacetylation of histone 4. This deacetylation was not observed in an HSF1 knockout mouse fibroblast cell line (Sakurai and Enoki 2010). Another study in a human erythroleukemia cell line (K562) showed that the number of DNA sites bound by HSF1 in non-cycling mitotic cells was severely reduced from those observed in cycling cells, suggesting there is a mechanism that limits the ability of HSF1 to interact with chromatin in non-cycling cells. However, HSF1 still was present on several DNA sites in non-cycling cells yet no active transcription from these sites was observed (Vihervaara et al. 2013). Furthermore, other studies have observed that the HSF1 transcription initiation complex includes the protein Brahma-related gene 1 (BRG1), a subunit of the chromatin remodeling complex that is required to allow RNA polymerase II activity at the transactivation site of HSF1 (Björk and Sistonen 2010).

HSF1 has also been investigated as a potential therapeutic target for cancer treatments. Cancer cells deprived of HSF1 became apoptotic while non-cancer cells survived (Åkerfelt et al. 2007, 2010). Additionally, mice carrying an oncogenic mutation that were also deficient in HSF1 developed fewer tumors that the same mouse strain with normal levels of HSF1 (Dai et al. 2007). The current explanation for these observations is that the dependency of cancer cells on HSF1 is due to the higher level of stress caused by increased metabolism, oxidative stress and spontaneous DNA damage that is observed in cancer cells which creates an increase in frequency of abnormal protein production and the accumulation of protein aggregates (Åkerfelt et al. 2010).

1.3.2 Heat shock factor 2

Unlike HSF1, HSF2 is maintained in an inactive state within the cells as both a dimer and a monomer, and it is these two forms that combine to form an active trimer (Åkerfelt et al. 2007). Although no studies have demonstrated that HSF2 is activated in response to heat shock, evidence suggests that when cells are exposed to stresses, HSF2 works cooperatively with HSF1 to regulate the expression of genes by forming herterotrimers with HSF1 (Sistonen et al. 1994; Morange 2006; Åkerfelt et al. 2007; Östling et al. 2007; Björk and Sistonen 2010). The occurrence of HSF1-HSF2 heterotrimers is more prominent during the acute phase of stress, and fewer heterotrimers are formed during prolonged exposure to heat stress (Åkerfelt et al. 2010).

While the activation of HSF2 has not been shown to have a direct link to cell stress, several studies have demonstrated a role for HSF2 in development. The absence of normal *hsf2* expression in mice has been linked to defects in development of the brain and the reproductive tract (Åkerfelt et al. 2007; Östling et al. 2007; Björk and Sistonen 2010). HSF2 levels are high early in mouse brain development (Åkerfelt et al. 2010). The discovery that HSF2 regulates p35 gene expression during the development of the cerebral cortex was the first evidence that HSFs could regulate non-classical heat shock genes during development (Chang et al. 2006). Mice lacking HSF2 have abnormal neuronal migration patterns, resulting in the formation of large ventricles and defects in lamination (Abane and Mezger, 2010; Åkerfelt et al. 2010). As well, two separate strains of *hsf2* knockout mice developed structural abnormalities in the hippocampus, vesicles and striatum in adult brains (Kallio et al. 2002; Wang et al. 2003). HSF2 is also expressed in proliferative neural progenitor cells during mouse development (Abane and Mezger 2010). The importance of HSF2 in brain development is further supported by observations that the expression of HSF2 in zebrafish is primarily observed in brain tissue of embryos between 36-84 hpf (Yeh et al. 2006).

Experiments studying the lack of HSF2 function in reproductive tissues in female mice have given conflicting results (Kallio et al. 2002; McMillan et al. 2002; Wang et al. 2003). Exactly what role HSF2 has in development of reproductive tissues has been the subject of some controversy in the past as two research groups using *hsf2* null mice demonstrated remarkably different phenotypic results. However, in two separate lines of *hsf2* knockout mice, females exhibited decreased fertility which was linked to several factors affecting normal oocyte development and ovulation (Kallio et al. 2002; Wang et al. 2003). These results could not be

confirmed in a third *hsf2* null mouse line (McMillan et al. 2002). However, the effects of the absence of *hsf2* on the development of male reproductive tissues have been more fully characterized. Male *hsf2* null mice developed smaller testes which contained abnormal seminiferous tubules with a resulting reduction in fertility. A double knockout of both *hsf1* and *hsf2* in male mice resulted in a more severe phenotype that included complete sterility, abnormal sperm shape and reduced sperm numbers (Wang et al. 2004; Åkerfelt et al. 2010). Gene expression studies demonstrated that *hsf2* is expressed specifically in pachytene spermatocytes and round spermatids during spermatogenesis (Åkerfelt et al. 2010). Further analysis revealed that *hsf2* expression in these cells is regulated in part by miR-18, a microRNA that binds to the 3' UTR of *hsf2* mRNA (Björk et al. 2010). It has also been suggested that the formation of HSF1-HSF2 heterotrimers is necessary during spermatogenesis to transform HSF1 into a transcriptionaly active state (Sandqvist et al. 2009). This suggestion is supported by evidence that both HSF2 and HSF1 occupy a series of multi copy genes on the X and Y chromosomes in post-meotic cells during the pachytene stage of spermatogenesis (Åkerfelt et al. 2010).

HSF2 also plays a role in chromatin modification as the sumoylated form of HSF2 prevents heterochromatin formation in the region of the chromosome to which it is bound (Hong and Sarge 1999; Xing et al. 2005). It has been suggested that this is a result of an apparent ability of HSF2 to interact with the enzymatic condensing complex and prevent it from inactivating genes during mitosis (Hong and Sarge 1999; Xing et al. 2005). This mechanism is referred to as gene bookmarking and results in less condensed regions of DNA in the promoter region of target genes including *hsp70* and *c-myc* (Abane and Mezger 2010). HSF2 occupies the promoter of *hsp70* during all stages of mitosis and may be important for allowing transcriptional activation of *hsp70* during the G1 phase of the cell cycle (Xing et al. 2005).

HSF2 may also be important for maintaining proteostasis during stress. In mouse MEF cell lines, lack of HSF2 caused cells to be more sensitive in response to long term mild heat shock resulting in increased ubiquitylated and misfolded proteins when compared to wild type cells (Shinkawa et al. 2011). The *hsf2* null MEF cells also had decreased αB-crystallin levels compared to wild type cells after heat shock. The knockout of *hsf2* in a Huntington's disease mouse model caused a similar increase in the accumulation of polyglutamine protein aggregates in the brain as well as shortening the lifespan of the mice (Shinkawa et al. 2011). Still other studies have shown that HSF2 constitutively binds to *hsp90*, *hsp27* and *c-fos* genes in Jurkat (T-

lymphocyte) cells suggesting that there are still other roles for HSF2 under non-stress conditions that need to be further characterized (Wilkerson et al. 2007).

1.3.3 Heat shock factor 3

Heat shock factor 3 was originally identified as an avian specific HSF (Nakai and Morimoto 1993) but an ortholog has recently been identified in mice (Fujimoto et al. 2010). The primary role of the HSF3 protein appears to be regulation of the cellular heat shock response during severe and long-term stresses (Tanabe et al. 1997, 1998; Kawazoe et al. 1999; Fujimoto et al. 2009). In avian species it is expressed at comparable levels to HSF2 in various tissues, and like HSF2 it exists as dimers and monomers when in its inactive state (Nakai and Ishikawa 2000; Pirkkala et al. 2001). However, unlike HSF2, the expression of HSF3 is induced by heat shock and therefore its function is more similar to HSF1 in this aspect. Although both HSF1 and HSF3 are expressed in response to heat stress, HSF3 is only expressed after severe heat shock and it appears that it becomes the dominant HSF regulating the cells response to severe and persistent heat stress (Kawazoe et al. 1999; Tanabe et al. 1997, 1998). Additionally, HSF3 was unable to induce *hsp70* expression in response to heat stress but did upregulate non-classical heat shock genes such as PROM2, a membrane glycoprotein, in mouse fibroblast cells (Fujimoto et al. 2010).

Some developmental roles for HSF3 have also been identified (Kanei-Ishii et al. 1997; Tanikawa et al. 2000; Fujimoto and Nakai 2010). HSF3 may have a developmental role in regulating hematopoiesis and cell proliferation as it has been demonstrated in unstressed cells that HSF3 binds to *c-Myb*, a transcription factor involved in these processes (Kanei-Ishii et al. 1997; Tanikawa et al. 2000). As well, the expression of Hsp90α in cultured DT40 chicken B lymphocytes was shown to be dependent on both HSF1 and HSF3 expression under non-stress conditions (Nakai and Ishikawa 2001).

1.3.4 Heat shock factor 4

HSF4 is the fourth HSF to be characterized in mammalian systems (Nakai et al. 1997; Tanabe et al. 1999). The structure of HSF4 has a pronounced difference from other HSFs in that, similar to the yeast HSF, it lacks the inhibition of trimerization domain (HR-C) (shown in Fig. 1.1), suggesting that HSF4 may be constitutively bound to HSE (Nakai et al. 1997). Two

isoforms of this protein have been identified in several mammalian species, both of which lack the HR-C domain. HSF4 in humans was initially reported to function as a repressor of *hsp70*, *hsp90* and *hsp25* in cultured HeLa cells (Nakai et al. 1997). Further studies revealed that it was only one isoform of the protein, HSF4a, that acted as a repressor and that the second isoform, HSF4b was an activator of hsps in a luciferase reporter assay (Tanabe et al. 1999; Frejtag et al. 2001; Zhang et al. 2001).

A role for HSF4 in development was first described by Bu et al. (2002) after studying three Chinese and one Danish family suffering from congenital cataracts. They identified a common missense mutation in *hsf4* of all the Chinese families that resulted in a conserved leucine residue being changed to a proline in the DBD of HSF4. A separate missense mutation was identified in the *hsf4* of the Danish family, which also resulted in an amino acid substitution in the DBD of the protein. An *hsf4*-null mouse line demonstrated that HSF4 had a tissue specific function as these mice had no major abnormalities in growth or organ development with the exception of cataract formation (Fujimoto et al. 2004). The development of the lens in these mice was normal early in embryogenesis but abnormalities such as inclusion-like bodies could be observed in lens fibre cells during late embryogenesis and postnatally (Fujimoto et al. 2004). Both classical and non-classical heat shock genes have been identified as targets for HSF4 regulation in mammalian lenses. Several studies have shown that a disruption of the regulation of many of these genes results in protein aggregates and inclusions in the lens due to improper lens fibre development and ultimately resulting in development of a cataract (Fujimoto et al. 2004; Min et al. 2004; Morange 2006; Shi et al. 2009; Abane and Mezger, 2010; Enoki et al. 2010; Zhou et al. 2011; Jing et al. 2013). In *hsf4* null mice, increased expression of heat shock genes *hsp60* and *hsp70* was observed in lens epithelial cells. As well, another heat shock gene, *hsp27*, had decreased expression in mouse lenses of *hsf4* knockout mice as well as in cultured HeLa cells in which *hsf4a* was overexpressed (Nakai et al. 1997; Fujimoto et al. 2004; Min et al. 2004; Fujimoto and Nakai 2010). However, when *hsf4b* was overexpressed in HeLa cells, expression of *hsp27* increased (Tanabe et al. 1999). Several non-hsp genes are also regulated by HSF4 in the epithelial cells of the mammalian lens. An increase in expression of several *fgf* genes in lens epithelial cells resulting in an increase in cell proliferation was observed in *hsf4* null mice (Fujimoto et al. 2004, 2008). HSF4 was also, shown to bind to the promoter region of *SKAP2*, a gene important for reorganization of actin filaments and linked to cell proliferation,

suggesting another possible target gene for HSF4 regulation in lens epithelia (Zhou et al. 2011). Results from Hu et al. (2013) also strongly suggest that HSF4b is required for lens epithelial cells to differentiate into lens fibre cells possibly through regulation of *RhoA* and *Rac1*, two genes known to be important in lens fibre cell differentiation.

The main target of HSF4 in lens fibre cells appears to be the γ-crystallin gene family but it also regulates other genes important for lens fibre differentiation and organization including *αB-crystallin*, *βB-crystallin* and structural proteins such as Bfsp (lens specific beaded filament), vimentin and connexins (Cheng et al. 2004; Fujimoto et al. 2004; Min et al. 2004; Somasundaram and Bhat 2004; Shi et al. 2009; Enoki et al. 2010; Mou et al. 2010; Jing et al. 2013). Recently it has been demonstrated in cell culture that activation of *αB-crystallin* expression by HSF4b is, in turn, regulated through the extracellular signal-regulated kinase (ERK) pathway in response to FGF2 stimulation (Li et al. 2013; Hu et al. 2013). However in the absence of FGF2, HSF4b was still able to regulate *hsp25* expression in the same cells suggesting multiple mechanisms for the regulation of HSF4b activation may exist in lens cells (Hu et al. 2013).

Regulation of transcription by HSF4 has also been identified in tissues other than the developing lens. In developing olfactory epithelium of post-natal mice, HSF4 increases expression of leukemia inhibitory factor (*Lif*) a gene which is necessary for normal sensory neuron development (Takaki et al. 2006). Most recently, microarray analysis of the expression of heat shock genes in neural lesions taken from multiple sclerosis patients identified a significant increase in expression of HSF4 but not of other HSFs (Mycko et al. 2013). These recent studies demonstrate that HSF4 also has a role in regulating developmental processes other than lens development.

Several studies have demonstrated that HSF4 works with other HSFs to regulate target genes. While an increase in mRNA expression was observed when HSF4 was bound to the promoter region of *fgf* genes in the lens epithelium and the *Lif* gene in the olfactory epithelium, a decrease in expression of both of these genes is seen when they are bound by HSF1 (Fujimoto et al. 2004; Takaki et al. 2006). These observations of the opposing effects of HSF4 and HSF1 were some of the earliest evidence that interplay between HSFs exists during development and also suggested that they play important roles in maintaining sensory organs (Åkerfelt et al. 2007; Morange 2006, Nakai 2010). More recently Kim et al. (2011) have

demonstrated that HSF4 interacts with HSF2 to modulate the expression of *hsp70* in hemin-treated cell lines as their study demonstrated that overexpression of *hsf4a* inhibited the oligimerization and DNA binding capacity of HSF2 in hemin-treated cells..

Recently, a role for HSF4 in regulating DNA repair or degradation has been identified. HSF4 was able to bind the *Rad51* recombinase gene both in human lens epithelial cells and in zebrafish lens. This enzyme has key roles in DNA repair mechanisms. Damage to DNA such as spontaneous double strand breaks that occur during cell division and migration, can disrupt the development of the lens and have been linked to cataract formation (Cui et al. 2012). Similar studies also suggest that HSF4 regulates DNAse 2β (DLAD) expression which is necessary for proper denucleation of lens fibre cells during lens development (Wride 2011; Cui et al. 2013). It has been suggested that impaired HSF4 activity would interfere with proper regulation of DNA repair or degradation. The inhibition of DNA degradation processes during lens formation in mice results in inclusions in the lens from damaged DNA remnants and ultimately results in cataract formation (Wride 2011).

There may also be a role for HSF4 in chromatin modulation. Initial studies conducted in HeLa cells demonstrated that both isoforms of HSF4 could bind to the BRG1 subunit in chromatin remodeling complexes during G1 phase of the cell cycle (Tu et al. 2006). The authors also suggest that this interaction may be dependent upon ERK 1/2 MAP kinases (Hu et al. 2006; Tu et al. 2006). He et al. (2010) demonstrated that the interaction of HSF4 with BRG1 was necessary for proper lens denucleation, possibly through the regulation of DNase IIβ gene. The interaction of HSF4 with genomic DNA from mouse lenses has been associated with decreased methylation of residues on histone 3, specifically H3K9 and H3L9 (Fujimoto et al. 2008; Sakurai and Enoki 2010). Methylated H3L9 has been frequently observed to be associated with non-transcriptionally active chromatin. The role of HSF4 in demethylation of this residue suggests that HSF4 could be allowing transcription factors including HSF1 to have access to these genes (Sakurai and Enoki 2010). A microarray analysis of HSF4 binding sites in mouse lens further support this role of HSF4. Results of the study demonstrated that only 5% of the sites bound by HSF4 were in the 5' promoter region of target genes while the remaining binding sites were mapped to introns, exons and DNA sites located more than 10 kb from target genes. This could suggest that HSF4 has other roles in the cell independent from transcriptional regulation (Fujimoto et al. 2008).

1.3.5 Other heat shock factors

1.3.5.1 Heat shock factor Y and heat shock factor X

Two genes encoding HSF family proteins have been identified in the sex chromosomes of humans. HSFY was identified as a member of the heat shock transcription factor family due to the high degree of homology it shares with the DBD of other HSFs (Skaletsky et al. 2003). The *HSFY* gene has been localized to the Y chromosome, is comprised of 7 exons, and has been identified in humans, mice, cattle, Rhesus monkey, opossum, kangaroo, zebu cattle and cats as shown in GenBank (Shinka et al. 2004; Tessari et al. 2004; Benson et al. 2005). An X chromosome homologue to *HSFY*, referred to as *HSFX* or *LW-1*, has been identified in human and opossum. The *HSFX* gene has only three exons, however the introns are present in the same location as those found in the DNA binding domain of *HSFY* which suggests that both arose from the same ancestral gene (Shinka et al. 2004). Phylogenic analysis of the DBD illustrated that HSFY and HSFX belong to a group distinct from other known HSFs (Shinka et al. 2004). Human HSFY and HSFX share 32% and 49% amino acid identity, respectively, with the ORF of human HSF1. As well, the DBD of both proteins share 53% amino acid identity to each other (Shinka et al. 2004).

The most striking difference between these proteins and other HSFs is that they appear to lack an HR-A/B or trimerization domain. An analysis of the amino acid sequence of HSFY and HSFX found few heptad repeats and the sequences were not homologous with HR-A/B or HR-C domains and not typical for a leucine zipper motif found in other HSF proteins (Shinka et al. 2004). The lack of an HR-A/B domain strongly suggests that these proteins would not be able to form active trimers and therefore questions whether these proteins can function as transcription factors in a way similar to other members of the HSF family. Furthermore, human HSFY proteins were not able to bind to a typical HSE in a gel mobility shift assay (Shinka et al. 2004). The DBDs of other HSFs have a conserved alpha helix 2 and 3 in the helix-turn-helix structure which is important for DNA recognition (Ahn et al. 2001). The alpha helix 2 and 3 structure is not conserved in the HSFY protein DBD even though the amino acid sequence is sufficiently similar to other HSFs to group it with this family of proteins (Shinka et al. 2004). However, a single nuclear localization sequence, similar to the ones found in human HSF1, HSF2 and HSF4, has been identified near the DBD of HSFY and HSFX, suggesting that the

protein may be capable of translocating to the nucleus (Sheldon and Kingston 1993; Shinka et al. 2004).

The Y chromosome consists of a short arm (Yp) and a longer arm (Yq). A large region, comprised of 95% of the total length of the human Y chromosome, has been identified as unique to the Y chromosome and has been termed the male specific region (MSY). The gene structure within the Y chromosome is complex with many of the protein coding genes existing in duplicate copies arranged within eleven large palindromic sequences, all of which are separated by regions of inverted repeat sequences and tandem arrays (Kichine et al. 2012). To date, eleven out of twenty seven distinct protein coding gene families have been determined to show testis specific expression and these testis specific genes families, including *HSFY*, are all located within the palindrome regions of the Y chromosome (Skaletsky et al. 2003; Kichine et al. 2012). The *HSFY* gene is present in two copies in the fourth palindrome of the Y chromosome, and are the only protein coding genes within this this region (Tessari et al. 2004). These genes encode RNA for three transcript variants of *HSFY* of which only transcript variant 1 contains a conserved DBD (Shinka et al. 2004). Two clusters of 2 genes each sharing 80% nucleotide identity to *HSFY*, arranged as inverted sequences, were also identified in palindrome 1 and palindrome 3 of the Y chromosome. While these genes maintain the seven exon structure of *HSFY*, it is assumed that they are non-functional pseudogenes due to the presence of a stop codon at the 16^{th} codon in the predicted amino acid sequence and the lack of an identified EST (expressed sequence tag) in GenBank (Tessari et al. 2004). Analysis of mRNA expression has determined that transcript variant one was detected in testis tissue, spermatozoa as well as in the brain and pancreas. Transcript splice variants 2 and 3 were only detected in testis tissue (Tessari et al. 2004). Within the testis, *HSFY* was determined to be present in the cytoplasm of Sertoli cells in all seminiferous tubules as determined by immunohistochemistry (Shinka et al. 2004). In similar experiments *HSFY* was also present transiently in the nucleus and perinuclear region of germ cells during specific phases of spermatogenesis in both humans and mouse, with the strongest expression being observed in round spermatids (Shinka et al. 2004; Tessari et al. 2004; Kinoshita et al. 2006). *HSFY* disappeared in fully mature sperm and could not be detected in other testicular tissues including Leydig cells or the lamina propria (Shinka et al. 2004).

Most of the research surrounding the function of HSFY has been focused on its potential role in fertility. Three separate regions of the Y chromosome in which microdeletions

have been associated with infertility were identified by deletion mapping and termed azoospermia factor intervals (AZFa, AZFb and AZFc) (Forresta et al. 2001; Skaletsky et al. 2003). Microdeletions occur in 10 to 15% of human males with idiopathic azoospermia and oligozoospermia, most of which are due to microdeletions in the AZFb region (Tessari et al. 2004; Kichine et al. 2012). It is this region where both copies of the *HSFY* gene reside (Ferlin et al. 2003; Skaletsky et al. 2003; Kichine et al. 2012). However most microdeletions in this area include a loss of multiple genes, in some instances up to as many as 18 genes were removed in a microdeletion.

The effect on fertility in human males due to the loss of *HSFY* through microdeletions within the Y chromosome has been analyzed in two studies. A study by Stahl et al. (2012) looked at the difference in expression of ten candidate genes in testicular tissue containing seminiferous tubules from 19 males with idiopathic azoospermia due to maturation arrest versus 8 males with obstructive azoospermia. They determined that only two genes, *CDY2* and *HSFY*, were consistently under-expressed by 12 and 16 fold respectively, in the patients with azoospermia. They concluded that both of these genes may be important for sperm maturation (Stahl et al. 2012). A second study was able to isolate 3 of 1186 infertile men who had a Y chromosome microdeletion in which both copies of *HSFY* were the only deleted protein coding genes and in which no similar microdeletion was found in 1179 control men (Kichine et al. 2012). The second study found that while the three men had reduced fertility, there was no evidence of azoospermia due to maturation arrest, contradicting the study by Stahl et al. (2012). The authors noted that while the *HSFY* deletion resulted in reduced fertility, it was not a severe enough phenotype to eliminate it from the population and continued to be passed on to future generations (Kichine et al. 2012). Interestingly, two of the three individuals with this specific microdeletion also observed a decline is sperm quality over time which may suggest that the phenotype of HSFY absence may only be of significance in individuals where there are multiple risk factors for infertility such as age (Kichine et al. 2012). This last point is especially interesting given recent investigation into the roles HSFs have in regulating protein homeostasis or proteostasis in cells. As has been previously described, the proteins involved in the HSR act as chaperones to prevent protein degradation and aggregation in response to cellular stresses. Aging organisms experience an increase in oxidative stress and corresponding increase in protein aggregation that can contribute to protein disease such as Parkinson's and Alzheimer's disease

(Åkerfelt et al. 2010; Morimoto 2011; Gidalevitz et al. 2011; Barna et al. 2012). It is interesting then to speculate that HSFY could have a role in regulating proteostasis in testicular cells, specifically Sertoli cells and spermatids, to maintain fertility. This would support the observation that while a deletion of the *HSFY* gene does not cause complete azoospermia, the loss of this gene may result in an earlier accumulation of stress/age-related protein aggregates compared to individuals that still maintain the *HSFY* gene on their chromosome. The ultimate result of an early accumulation of protein aggregates in the testis may be reduced fertility in males at an earlier age as suggested by Kichine et al. (2012). The observation that HSFY can not bind to HSEs makes it unlikely that this protein would protect cells from protein aggregates through transcriptional regulation of classical heat shock genes (Shinka et al. 2004). Therefore, the mechanism by which HSFY may protect cells from age related stress and prevent infertility is likely a novel one.

1.3.5.2 Heat shock factor family member 5

Heat shock factor family member 5 is the most recent HSF gene identified by sequence analysis of genomes deposited in GenBank (Benson et al. 2005). Homologues for this gene have been found in more than 30 species including human, mouse, zebrafish, *Xenopus* and chicken suggesting that the *hsf5* gene, like other HSFs, is conserved across genomes. Expression of *hsf5* in cell culture has been observed in two separate studies. Choi et al. (2010) demonstrated that the level of *hsf5* mRNA in mouse neural stem cells increased in response to ethanol treatment, similar to the pattern observed for *hsf2* mRNA levels in the same cells. A second study demonstrated that *hsf5* expression was also increased in 3T3-L1 preadipocyte cells exposed to high glucose levels (Gupta and Tikoo 2012). A third study analyzing gene expression in the zebrafish swimbladder determined *hsf5* to be one of the twenty most abundant transcription factors expressed in the adult tissue (Zheng et al. 2011).

To date, no characterization of the HSF5 protein sequence has been published. Alignment of the HSF5 predicted amino acid sequence with sequences from other HSFs using the BLAST sequence alignment tool confirms that the DBD for HSF5 is conserved with other HSFs. The predicted amino acid sequence of HSF5 DBD shares 36-37 % amino acid identities with the DBD of HSF1, HSF2, and HSF4 in humans. A similar pattern is observed in zebrafish where the DBD of HSF5 shares 37-38% amino acid identities with the DBD of zebrafish HSF1,

HSF2 and HSF4. However, no heptad repeat sequences were identified by the SMART analysis in HSF5 protein sequences from either human or zebrafish indicating that HSF5, like HSFY, does not contain an HR-A/B or HR-C domain and therefore would not be able to form active trimers.

The function of HSF5 is unknown. The predicted protein sequence, like the HSFY sequence, lacks all heptad repeat sequences typically associated with HSFs. Even though HSF5 does contain a conserved DBD, the lack of an HR-A/B domain suggests that the HSF5 protein will not be able to bind classical HSE sequences, similar to what was observed for HSFY. The discovery of the HSF5 sequence and its similarities to HSFY in the lack of HR domains suggests that these represent a new category of HSF. More studies are needed to determine how these HSFs lacking trimerization domains function in the cell. It will also be interesting to determine if these proteins do act as transcription factors and what regulatory roles, if any, they possess.

1.3.6 Heat shock factor transcriptional regulation

Heat shock factors regulate transcription of a number of genes that are involved in a large variety of physiological processes (Pirkkala et al. 2001; Åkerfelt et al. 2007; Fujimoto and Nakai 2010; Morimoto 2011). Gene targets include the classical heat shock genes as well as non-heat shock genes, both of which can be regulated by HSFs in response to stress. The diversity of genes regulated by HSFs is also highlighted in several studies that have demonstrated that the gene targets for HSFs can be tissue- or developmentally-specific. For example, a knockout of *hsf2* using RNAi techniques in HeLa cells resulted in decreased *hsp27* expression while the expression of *hsp27* in fibroblasts from an *hsf2* null mouse was not different than expression in wild type mice (Wilkerson et al. 2007). In developing mouse lens, HSF4 is able to repress the expression of several *fgf* genes in lens epithelial cells but this same inhibition could not be demonstrated in lens fibre cells (Fujimoto et al. 2004). Recently, Jing et al. (2013) were able to show that HSF4 bound to the *αB-crystallin* promoter in established epithelial cell lines but could not bind to the same gene in fibroblast cell lines. However the inverse was found to be true for HSF1 which did not bind to *αB-crystallin* in epithelial cells but did bind this gene in fibroblasts in response to a heat shock (Jing et al. 2013).

The high degree of similarity shared between protein sequences of all members of the HSF family suggests that there must be a complex network of processes that can regulate

HSF activation and differential recruitment of HSFs to specific gene targets as described above. There is therefore a keen interest in elucidating mechanisms by which HSF trimerization is regulated and how these HSFs differentially regulate target genes. Investigations into the regulation of HSFs have primarily been performed in mammalian species and what is currently understood about regulation of HSF activation is described below.

1.3.6.1 Activation of HSF trimers

The activation of HSFs is a multistep process that includes the trimerization of inactive HSF monomers and dimers, localization of the trimer to the HSE on target genes, and activation of gene expression through formation of the transcription initiation complex (Åkerfelt et al 2010; Björk and Sistonen 2010; de Thonel et al. 2012). Several hsps are involved in regulating the formation of HSF complexes. Hsp90 and to a lesser extent Hsp70 and Hsp40 have all been suggested to be involved in protein complexes that, when bound to HSF monomers or dimers prevent trimerization resulting in HSFs being maintained in an inactive state (Fig. 1.2) (Morano and Thiele 1999; Voellmy 2004; Björk and Sistonen 2010). Upon stress they are recruited away from HSF monomers to be used as chaperones for other proteins (Fig. 1.2). Depletion of the chaperones from HSF monomers and dimers results in the breakdown of the inactive heterocomplex allowing unbound HSFs to rapidly trimerize. These active HSF trimers then migrate to and bind to the HSE resulting in rapid synthesis of hsps. Once all of the damaged proteins are bound to hsps, excess hsps are once again free to interact with HSF, thereby dissociating active HSF trimers, and returning HSF to the inactive, heterocomplexed form creating a feedback loop (Fig. 1.2). A few proteins involved in opposing chaperone inhibition of the HSF complex have also been identified including the FAS death domain associated protein (DAXX) which associates with the HSF1 trimer. An Hsp70 co-chaperone, CHIP, also interacts with HSF1 to facilitate activation of transcription of target genes (Björk and Sistonen 2010). The mechanism by which the cell relieves chaperone inhibition in response to cues other than cellular stress, such as developmental processes, has not been studied to any large extent.

The cellular signaling events that initiate HSF activation in response to cellular stresses or developmental cues are not as well understood and the complete signaling pathways for these processes remain largely unknown. Studies using *Drosophila* HSF and mammalian

HSF1 purified from expression systems demonstrated that application of heat, a decrease in pH or simulation of oxidative stress using H_2O_2 could directly stimulate HSFs to trimerize, suggesting a stress sensing mechanism was inherent in the protein (Ahn and Thiele 2003; Björk and Sistonen 2010). There is also some evidence that a specific RNA molecule called heat shock RNA-1 acts as a thermosensor. The interaction of heat shock RNA-1 with elongation factor eEF1A is able to facilitate HSF1 trimerization in response to heat stress (Björk and Sistonen 2010). Still other studies have shown that inducing changes in membrane fluidity could increase HSF1 DNA binding although the cell signaling pathways involved in this process remain to be elucidated (Björk and Sistonen 2010).

One of the most widely studied aspects of HSF regulation has been to determine the role that post-translational modifications (PTMs) play in the regulation of HSF trimer formation. Most of the PTMs known to be associated with HSF regulation have been identified from studies where trimerization of HSFs was induced by stress. The nature of PTMs associated with trimerization of HSFs in response to developmental cues remain mostly unknown (Björk and Sistonen 2010). In yeast, different phosphorylation patterns were identified in response to heat stress versus oxidative stress which strongly suggests that PTMs are integral in the ability of HSFs to activate specific genes (Björk and Sistonen 2010; Sakurai and Enoki 2010).

There have been several inducible PTMs identified on mammalian HSFs including phosphorylation, acetylation, sumoylation, and ubiquitination (Åkerfelt et al. 2007; Björk and Sistonen 2010; Xu et al. 2012). The majority of the characterization of PTMs and their regulatory roles have been performed on HSF1. To date, twelve serine residues that undergo phosphorylation in response to heat shock have been identified on human HSF1, with the majority lying within a regulatory domain between the HR-A/B and HR-C domains (Björk and Sistonen 2010; Xu et al. 2012). This regulatory domain is responsible for inhibiting transactivation through the HR-C domain. It is possible that the large number of phophorylation events triggered in response to heat shock is linked to relieving the HR-C domains ability to inhibit trimerization (Björk and Sistonen 2010). Phosphorylation of serine 121 in the regulatory region also induces inhibition of transactivation by promoting the association of HSF1 with Hsp90 (Xu et al. 2012). However, phosphorylation modifications can also promote other PTMs on the HSF1 complex which allow the HR-C to inhibit transactivation. For example, phosphorylation of serine 303 is necessary for a nearby lysine residue to be modified by

sumoylation, the result of which is the inhibition of HSF1 transactivation (Björk and Sistonen 2010; Xu et al. 2012). Phosphorylation of still other residues has additional regulatory roles. For example the phosphorylation of serine 419 is required for nuclear translocation of the active trimer while phosphorylation of serine 216 is associated with HSF1 ubiquitination (Xu et al. 2012).

Acetylation of several lysine residues in the DBD is another PTM that has been identified in HSF1. This modification allows dissociation of HSF from the HSE. It has also been demonstrated that deacetylation of HSF1, which occurs via longevity factor sirtuin 1 (SIRT1), promotes DNA binding (Björk and Sistonen 2010). This acetylation modification is critical in the displacement of HSF1 from DNA given that the binding of hsps to the HR-C domain of an active HSF1 trimer can not inactivate the complex on its own (Björk and Sistonen 2010).

The existence and roles of PTMs on other mammalian HSFs is less well characterized. HSF2 is a short lived protein and therefore the identification of several ubiquitination sites within HSF2 trimers is not surprising (Björk and Sistonen 2010; Xu et al. 2012). HSF2 also interacts with subunits of the anaphase promoting complex which could promote movement of HSF2 to the ubiquitin/proteosome degradation pathway (Björk and Sistonen 2010). The abundant sites for ubiquitination in HSF2 trimers suggest that transcriptional activation of target genes by this protein is under tight regulation. Additionally two lysine residues with sumoylated modifications have been identified on HSF2. These sites are associated with the DNA binding capacity of HSF2 and therefore may be involved with its bookmarking function. However, conflicting results have been found for whether or not the sumoylation promotes or inhibits DNA binding (Björk and Sistonen 2010). No acetylation or phosphorylation sites have been identified in HSF2 trimers to date (Xu et al. 2012).

Only one phosphorylation site and one sumoylation site have been identified on mammalian HSF4, specifically on the HSF4b isoform. The sumoylation of lysine 293 is directly preceded by the phosphorylation of the nearby serine 298, the result being a repression of HSF4b activation. This process is very similar to what was observed for HSF1 with the exception that sumoylation occurs in HSF1 as a response to stress while HSF4b appears to by constitutively sumoylated, consistent with the HSF4 complex being constitutively trimerized (Tanabe et al. 1999; Björk and Sistonen 2010; Xu et al. 2012)

Only limited investigations have been performed to identify PTM in HSFY and HSFX proteins. Mass spectroscopy based proteomics analysis has identified two phosphorylation sites on HSFY isoform 1 and a sumoylation site has been identified on isoform 1 of HSFX (Xu et al. 2012). The function of these PTMs is unknown.

As discussed above, the processes that regulate HSF trimerization and their ability to transactivate genes are important mechanisms by which the few types of known HSFs differentially regulate target genes in response to a variety of cellular stimuli including stress and environmental cues. Another mechanism by which HSFs differentially regulate gene expression is by working together to co-regulate target genes. It has been shown that several processes such as brain and lens development are more severely hampered in mice when multiple HSFs are knocked out as opposed to only a single HSF (Björk and Sistonen 2010). This may be due in part to the presence of multiple HSEs in the promoter of some HSF target genes, allowing different HSFs to bind to their promoters as the same time. There have been several studies demonstrating that the promoters of HSF target genes are occupied by more than one type of HSF (Björk and Sistonen 2010; Åkerfelt et al. 2010). These include regulation of genes during spermatogenesis by both HSF1 and HSF2 and also several genes in the *fgf* gene family that are occupied by both HSF4 and either HSF1 or HSF2 during lens development (Fujimoto et al. 2008; Abane and Mezger 2010; Björk and Sistonen 2010).

Another mechanism by which HSFs can differentially regulate their targets is the composition of the HSF trimer itself. Initial studies indicated that all HSFs formed as homotrimers but several studies now suggest that HSFs may also form heterotrimers (Sandvqist et al 2009; Åkerfelt et al. 2010; Björk and Sistonen 2010). An initial observation was made by Loison et al. (2006) that the promoter region of the *clusterin* gene, which contains only a single HSE, was occupied by both HSF1 and HSF2 after stress. Subsequent studies in K562 erythroleukemia cells demonstrated that monomers from different HSFs can bind to each other through the HR-A/B domain to form heterotrimers (Östling et al. 2007; Sandvqist et al. 2009). Additionally HSF1 and HSF2 were localized to the same binding sites in nuclear stress bodies (Sandvqist et al. 2009). More recently it has been demonstrated that HSF1 and HSF2 synergistically regulate expression of multi-copy genes on the Y chromosome giving further evidence for HSF heterotrimer formation (Åkerfelt et al. 2010; Vihervaara et al. 2013). The ability of HSFs to work together to co-regulate gene expression as well as form both hetero and

homotrimers illustrates the complex number of mechanisms by which so few HSF proteins can differentially regulate target genes (Björk and Sistonen 2010).

1.3.6.2 Architecture of the heat shock response element

The architecture of the DNA sequence targeted by HSFs has a role in regulating gene activation. The HSE sequence that HSFs bind to consists of at least three inverted repeats of the sequence 5'-nGAAn-3' in the 5' UTR of target genes. Typically, each DBD of an HSF trimer associates with one of the GAA subunits in HSEs with three GAA repeats (Fig. 1.2). However many genes have HSEs that contain more than three HSE repeats. In yeast, an HSE containing four repeats was found to be occupied by two HSF proteins with one subunit from each trimer not making contact with the DNA (Sakurai and Enoki 2010). As well, multiple copies of HSEs have been found in the promoter regions of hsp genes which allow simultaneous binding of different HSFs. The number of 5'-nGAAn-3' repeats or length of the HSE may also play a role in specificity of HSF binding. Kroeger and Morimoto (1994) demonstrated that HSF1 proteins preferentially bound to long arrays of GAA repeats while HSF2 preferred shorter arrays. The organization of HSEs in this manner allows for the binding of multiple HSFs to work cooperatively to regulate gene expression (Åkerfelt et al. 2007; Sakurai and Enoki 2010).

Different HSF proteins exhibit different affinities for perfect 5'-nGAAn-3' versus discontinuous, or non-perfect, HSE target sequences. Discontinuous HSE sequences are prevalent in many species and may consist of single nucleotide changes to one or more of the inverted GAA subunits in an HSE sequence pattern. For example, a study in *Xenopus laevis* identified an HSE in the promoter region of *hsp70* that contains three inverted GAA repeats with a single nucleotide change in the second repeat (Ovsenek and Heikkila 1990). Other discontinuous HSE sequences, referred to as a gap HSE, are characterized by missing a complete GAA unit in the middle of a series of inverted repeats. Step HSE sequences containing a series of only GAA subunits separated by 5 nucleotides instead of the expected inverted repeats have also been identified in yeast (Yamamoto et al. 2009).

Several studies have been performed to determine if discontinuous HSE sequences affected the way in which different HSFs bound to these sites. Yeast heat shock factor binds both perfect and discontinuous HSE sequences (Hashikawa et al. 2007). In contrast, human HSF1 preferentially bound to continuous HSEs, while HSF2 had a higher binding affinity for

discontinuous HSEs than HSF1, and HSF4 exhibited the highest binding affinity for discontinuous HSEs (Yamamoto et al. 2009). Studies of HSF binding in mouse lens tissue also demonstrated that mouse HSF4 was able to bind to discontinuous and more ambiguous HSE sequences more strongly than HSF1 and HSF2 (Fujimoto et al. 2008; Sakurai and Enoki 2010). A region of amino acids in the carboxy terminus of HSF proteins near the HR-C domain has been recently identified as influencing HSFs ability to bind to discontinuous HSEs (Enoki and Sakurai 2011). Expression of alternatively spliced forms of *Drosophila* HSF identified two regions of amino acids flanking the HR-C domain that influenced the proteins ability to form tight trimers, and also influenced the binding affinity of the protein for different discontinuous HSEs (Enoki and Sakurai 2011). Similar experiments expressing two isoforms of zebrafish HSF1 (a and c) also identified a region near the HR-C domain, present in HSF1a and not in HSF1c, that allowed HSF1a to exhibit a higher binding affinity to discontinuous HSEs (Wang et al. 2001; Enoki and Sakurai 2011). Similarly, a substitution in the human HSF1 HR-C domain allowed HSF1 to bind discontinuous HSE in yeast cells (Takamori et al. 2009). It has been suggested that this observed preferential binding of HSFs to continuous and discontinuous HSEs may be an important part of the mechanism by which HSFs can differentially regulate gene expression in different tissues, or to regulate different biological functions (Yamamoto et al. 2009; Sakurai and Enoki 2010; Enoki and Sakurai 2011). However, a more recent study using ChIP sequencing to identify HSF1 and HSF2 binding sites in a human leukemia cell line demonstrated that the HSE sequences that each HSF bound were very similar across the genome. They also noted that even though the HSE sequences were very similar, HSF1 and HSF2 still could preferentially bind certain sites suggesting additional mechanisms are required for determining target binding sites (Vihervaara et al. 2013). The number of 5'-nGAAn-3' repeats or length of the HSE may also play a role in determining target specificity. Kroeger and Morimoto (1994) demonstrated that HSF1 proteins preferentially bound to long arrays of GAA repeats while HSF2 preferred shorter arrays.

1.4 Zebrafish as a developmental model system

A number of organisms have been used as models for developmental research, each having specific characteristics that are beneficial for experimental use. The zebrafish, *Danio rerio*, is a small, tropical, freshwater fish native to the Ganges River in India and other areas of southern Asia. There are several advantageous features of zebrafish that have made it an excellent model system to use for developmental research. Zebrafish eggs are externally fertilized and the embryos are transparent which allows easy detection of mutations, facilitates manipulative studies, such as transplanting tissues or injecting dyes, as well as allows for easy visualization of staining from *in situ* hybridization procedures. The transparency of zebrafish embryos allows for live imaging of cell migration, adhesion and differentiation in several tissues which is a technique that can not be easily done in mammals (Greiling and Clark 2009, 2012). Zebrafish developmental processes are also similar to those in higher vertebrates. Recently a comparative study was able to demonstrate that the zebrafish genome contains orthologues for 70% of the genes in humans (Howe et al. 2013). One obvious difference however, is that zebrafish do not contain X and Y sex chromosomes (Traut and Winking 2001; Anderson et al. 2012). Technically, zebrafish are easy to care for and maintain because individuals are inexpensive and can be housed in large numbers. Zebrafish also produce large numbers of eggs (20-200 per mating), and develop rapidly allowing experiments to be conducted and repeated quickly. The availability of a comprehensive zebrafish techniques manual containing valuable references on embryo care, developmental staging, and common molecular techniques is also an attractive advantage when using this system (Westerfield 1995). Collectively, these features make the zebrafish an excellent model system to study vertebrate development.

1.5 Development of the zebrafish eye

The development of functional vision in zebrafish is necessary for prey capture and avoidance behaviours (Easter and Nicola 1996). The zebrafish display startled avoidance behaviour in response to light and dark shadows by 69-72 hpf, eye movements indicative of tracking prey develop between 73-80 hpf and they begin to capture prey at about 4 dpf (Easter and Nicola 1996; Gestri et al. 2012). This suggests that rapid eye development is occurring prior to 72 hpf or 3 dpf.

The two primary components of the eye that facilitate vision are the lens and the retina. In zebrafish, development of a spherical lens can be observed as early as 20 hpf while the retina only begins to develop at 28 hpf (Greiling and Clark 2009). Zebrafish larvae have a lens and retina that are both about 100 μM in thickness and the inner retinal membrane is juxtapositioned to the epithelial surface of the lens (Easter and Nicola 1996). As the zebrafish develops into an adult, the lens increases in size to about 10x the thickness of the retina and both of these tissues become separated by a vitreous space. The three layers of neural and glial cells are distinguishable in the retina by 3 dpf (Gestri et al. 2012). At the same time, the zebrafish lens has become spherical and the center of the lens begins to develop a transparent organelle free zone (OFZ) by 3 dpf (Greiling and Clark 2009). Retinal image formation and the maturation of the extraocular muscles complete their development simultaneously at about 72 hpf, which coincides with the time of hatching and onset of visual prey tracking behaviour (Easter and Nicola 1996). However, many aspects of functional vision are aquired in later stages of zebrafish development. For example, cone photoreceptor cells in the retina, required for vision in bright light, do not become fully functional until 5 dpf while the rod photoreceptor, required for vision in low light, is not functional until 15 dpf (Easter and Nicola 1996; Gestri et al. 2012). As well, some proteins in the lens, such as γS-crystallins, do not appear until six weeks of age (Greiling et al. 2009). Both retinal and lens cells proliferate throughout the life of the fish as the fish and fish eye continue to grow in size (Gestri et al. 2012; Greiling and Clark 2012).

Different developmental mechanisms exist in various species to adapt the eye structure to adjust the focal point on the retina. As zebrafish develop from larvae into adult fish, the lens becomes progressively thicker, much more so than the retina. The difference in the focal point on the retina from a thinner lens to the thicker lens is compensated by a continuing increase in the vitreous space between the lens and retina and not by any change in the shape of the eye (Gestri et al. 2012). In mammals, the adjustment for focal point between the juvenile and adult eye is due to changes in the dimension and shape of the whole eye. In amphibians the change in lens between the water dwelling juvenile and air dwelling animal occurs rapidly over a day and is due to changes in lens position and shape without a change in the shape or size of the whole eye (Easter and Nicola 1996).

1.5.1 Lens development

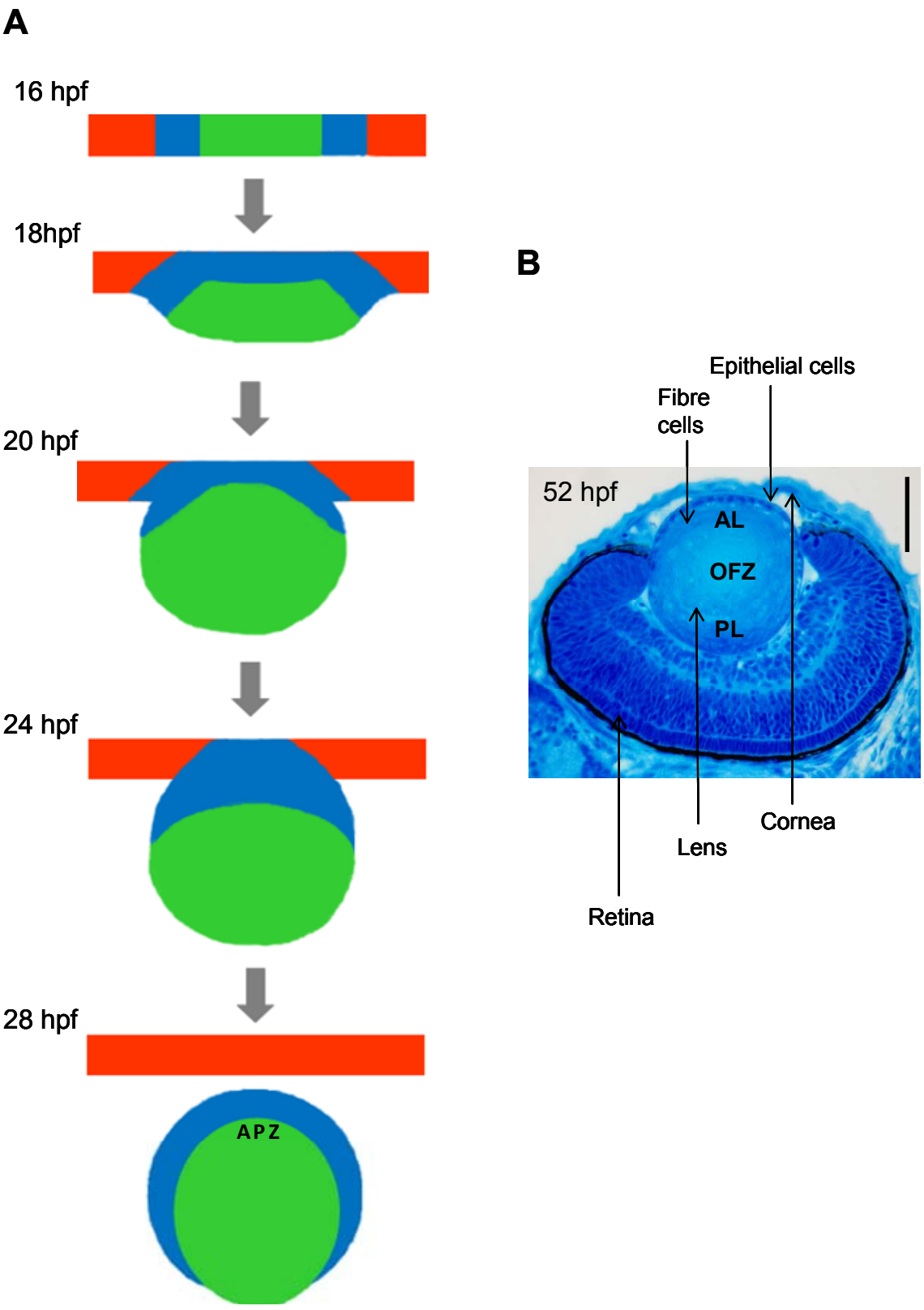
The vertebrate lens contains layers of organized, symmetrical, transparent, refractile cells that are responsible for transmission of light and focusing of images on the retina (Greiling and Clark 2008; Greiling et al. 2009). The lens develops as concentric layers which ultimately lose nuclei and organelles to become transparent. Lenses also go through a dehydration process to achieve a high concentration of proteins with an organized structure and eventually forms a tissue with the most densely packed protein composition in the body (Greiling and Clark 2009; Abane and Mezger 2010). For normal vision, the lens cells need to become and remain transparent for the entire life of the animal. When lens cells lose transparency, the result is the formation of a cataract which reduces or prevents functional vison due to the presence of occlusions (Greiling and Clark 2009). Cataracts can develop in aging individuals due to protein degradation and aggregation in the lens. Age related changes in protein composition due to oxidation, proteolysis, or disruption of protein-protein interactions can all contribute to the degradation and aggregation of proteins in the lens (Wride 2011). Cataracts can also be present in congenital (at birth) or juvenile (post-birth) forms and are responsible for approximately one third of infant blindness. These cataracts form during lens development as a result of either improper regulation of gene expression or incomplete organelle degradation and removal (Bu et al. 2002; Smaoui et al. 2004; Mellersh et al. 2006; Engelhardt et al. 2007; Wride 2011).

The zebrafish lens is morphologically similar to that of other vertebrates, containing a single layer of epithelial cells surrounding layers of elongated fibre cells throughout the interior of the lens. The zebrafish lens is more spherical than typical vertebrate lenses, a characteristic that may be attributed to the fact that the zebrafish eye continues to grow in size throughout life by adding layers of fibre cells to the lens (Gestri et al. 2012). The initial development of the lens and cornea in zebrafish is very similar to lens development in other species, with the lens placode forming from ectoderm. However, unlike mammalian and avian lens development, the zebrafish lens does not develop a fluid filled vesicle that pinches off from the ectoderm but rather forms a mass of cells that delaminate inward to form a sphere that eventually detaches from the ectoderm (Greiling and Clark 2008). This lack of cavity formation during organogenesis is a common developmental characteristic of zebrafish as the neural tube, otic vesicle and optic vesicle all initially form from a solid mass of cells (Greiling and Clark 2009, 2012). While some of the mechanisms driving lens development are different between fish and other species, the end

result for all is the formation of a symmetrical, transparent, refractile organ with similar organization and structure (Greiling and Clark 2009).

Development of the zebrafish lens can be divided into three stages, namely development of the lens placode, delamination of cells forming the lens mass and final delamination, detachment and maturation into a fully functional lens (Greiling and Clark 2009, 2012). The cranial placodes, which include the lens, olfactory, and otic placodes as well as the lateral line in aquatic species, develop from the ectoderm that is positioned at the interface where epithelial cells and neural cells develop, termed the preplacodal region. The lens placode in zebrafish is evident at 16 hpf as a single layer of elongated cells which overlays the area of the optic primordium destined to become the retina. The end of the first stage occurs between 16-18 hpf at which point the layer of cells in the lens placode rapidly double in size and take on a columnar shape. The cells then divide and form a cell mass two to three layers thick with an appearance resembling a bulging disc by 20 hpf (Fig. 1.3).

Figure 1.3 Development of the zebrafish lens. (**A**) shows a schematic representation of lens development starting from the lens placode at 16 hpf through to the formation of a spherical lens which has detached from the ectoderm at 28 hpf. At 28 hpf, secondary fibres begin to differentiate from the anterior proliferative zone (APZ). Cell fates are depicted by different colours. Cells fated to become cornea are shown in orange, anterior epithelium is shown as blue, and cells fated to become primary fibres are shown in green. (**B**) A cross section of a zebrafish eye at 52 hpf stained with methylene blue-azure II. The single layer of lens epithelial cells surrounding the lens and the layers of elongated fibre cells in the interior of the lens are indicated. The organelle free zone (OFZ) has also started to develop in the lens nucleus at this stage. The cornea is juxtaposed to the anterior lens (AL) while the retina is juxtaposed to the posterior lens (PL). Histological section is a 5.5 μM thick methacrylate cross section. Scale bar = 50 μm. This figure is modified from Evans (2006) and Greiling and Clark (2012).


A
16 hpf
18hpf
20 hpf
24 hpf
28 hpf
APZ
B
Epithelial cells
Fibre cells
52 hpf
AL
OFZ
PL
Cornea
Lens
Retina

The second stage of lens development begins with a reorganization of the cells into three distinct regions (Greiling and Clark 2009, 2012). A central mass of round cells delaminates from the center of the lens placode and forms the embryonic lens core or organizing center. Cells that will define the posterior and lateral borders of the lens become elongated and are fated to become primary fibre cells surrounding the lens core. Finally, still undifferentiated cells in the anterior of the lens organize themselves parallel to the surface ectoderm. The second stage ends with a series of cell delaminations, resulting in a spherical mass of cells that can be observed by 20 hpf (Fig. 1.3). This mass of cells still remains attached to the layer of ectoderm that will form the cornea.

The cell mass detaches from the cornea to form a distinct lens by 24 hpf. At this point the lens has become round with clearly defined borders, and is separated from the cornea by only a narrow aqueous chamber (Greiling and Clark 2008, 2012). The cells in the posterior-middle of the mass continue to become round and form an organizing centre. Primary fibre cells migrate to the posterior of the lens and elongate in a symmetrical manner to surround the nuclear mass. At 24 hpf, the cluster of cells originating from undifferentiated cells at the point of detachment in the anterior of the lens begins to divide and by 36 hpf a single layer of epithelial cells surrounds the spherical mass of primary fibre cells (Greiling and Clark 2009). All lens epithelial cells are competent to undergo differentiation into secondary lens fibre cells (Wride 2011). By 28 hpf the first secondary fibre cells can be observed (Fig. 1.3). Epithelial cells in the anterior portion of the lens, termed the anterior proliferative zone, divide inwardly to form secondary lens fibre cells (Fig. 1.3). These cells migrate towards the lens posterior while elongating to form semi-circular layers around the primary fibres (Greiling and Clark 2009, 2012). At 48 hpf, the lens has increased in size, more secondary fibre cells are present and existing elongated fibre cells have become extremely narrow. At 3 dpf (larva stage), the lens is spherical in all dimensions and individual cells in the lens nucleus can not be distinguished. Continuing from this point to 4 dpf, the lens diameter increases in size through additional layers of secondary fibre cells. As the entire zebrafish grows in size, the lens continues to develop and increase in size by synchronized proliferation, migration and elongation of the epithelial cells and their subsequent differentiation into fibre cells (Greiling and Clark 2009, 2012).

The final stage of lens fibre differentiation involves the process of degradation and removal of organelles and DNA from primary and secondary fibre cells, creating the organelle

free zone (Wride 2011; Greiling and Clark 2012). Organelles such as mitochondria, Golgi apparatus and nuclei scatter visible light when it passes through the fibre cells and therefore the process of organelle and DNA degradation is integral to the formation of a functional transparent lens. This is evidenced by the zebrafish mutant line, *bumper*, in which fish develop cataracts during lens development due to retention of nuclei in secondary fibres (Wride 2011). This process of organelle degradation and denucleation begins at 50 hpf in zebrafish and occurs simultaneously with the synthesis and organized dense packing of multiple crystallin proteins. The organelle loss begins initially in the primary fibres at the center of the lens. The loss of organelles continues progressively from the center outward towards the lens border until the organelle free zone occupies all but the newest secondary fibres at 72 hpf (Wride 2011; Greiling and Clark 2012).

Currently, very little is known about the molecular process involved in achieving lens transparency yet retaining cell viability. This process is even more intriguing considering the clearance of organelles occurs without vasculature or lymphatics (Greiling et al. 2009). Two different pathways have been implicated in the unique process of organelle breakdown in the lens, one being an attenuated apoptosis pathway and the other being through the ubiquitin proteosome pathway (Wride 2011; Greiling and Clark 2012). The first hypothesis designed to explain the clearance of organelles suggested that it was the result of an attenuated apoptosis pathway. The process of clearing the lens of organelles involves similar processes to those observed for apoptosis including the condensation of chromatin and disintegration of the nucleus with a similar DNA fragmentation pattern. Several apoptosis related genes are also implicated in lens organelle degradation. Expression of members of the B-cell lymphoma family, cytokines such as tumour necrosis factors and members of the p53 family of tumour suppressor proteins have been detected in developing mammalian lenses. As well, mice with knockouts for either caspase-3 or caspase-6, while showing no difference in organelle breakdown, did develop cataracts (Wride 2011). The knockdown of *hsp70* in zebrafish embryos resulted in a small lens phenotype as well as an increase in the duration of apoptosis during lens fibre differentiation (Blechinger et al. 2002; Krone et al. 2003; Evans et al. 2005, 2007; Evans 2006). Chaperones such as Hsp70 have been shown to protect against stresses that lead to apoptosis in many cell types and therefore it has been proposed that Hsp70 may play a role in the attenuated apoptosis process during lens fibre maturation (Evans et al. 2005; Evans 2006).

Even though some of the mechanisms involved in degradation of organelles in the lens resemble classic apoptosis, several other key aspects of classical apoptosis do not occur during lens formation. Lens cells still retain their structure after organelle degradation as they do not undergo cytoskeleton breakdown or membrane blebbing. While numerous apoptosis related genes are expressed in the lens it still remains unclear whether the mechanisms facilitating organelle loss in the lens are truly an attenuated form of apoptosis or another signaling process that overlaps or shares redundancy with classical cellular apoptosis (Wride 2011). For these reasons the process has been referred to as attenuated apoptosis.

More recent studies have identified that the ubiquitin-protesome pathway also has an essential role in lens fibre differentiation and transparency (Greiling and Clark 2012). Ubiquitin was identified as the fourth most abundant protein in the mouse lens fibre proteome. Inhibition of the proteosome pathway in chick lenses resulted in retention of mitochondrial protein and occlusion of the lens. As well, a fully functional ubiquitin-proteosome pathway has been characterized in bovine lenses (Greiling and Clark 2012). In zebrafish, the *volvox* mutant, which is deficient in a subunit of the 26S proteosome, has a small lens phenotype with disorganized cell structure, improper nuclei placement and insufficient elongation of secondary lens fibres (Imai et al. 2010). *Volvox* mutants had higher levels of ubiquitinated proteins than did wild type fish and it is presumed that in these cells the mechanisms required to tag proteins with ubiquitin are still present but the defective proteosome could not degrade the tagged proteins (Imai et al. 2010). In kidney fibroblasts Grune et al. (2011) demonstrated that Hsp70 was necessary for disassociation and re-association of the 19S regulators and 20S proteosome to form a functional 26S proteosome. They were also able to demonstrate that inhibitors of Hsp70 resulted in an inactive 26S proteosome. These observation suggest that the small lens phenotype observed by Evans et al. (2005) in *hsp70* knockdown zebrafish could also be due in part to a disrupted ubiquitin-proteosome pathway instead of or as well as having a role in an attenuated apoptosis pathway. While further investigation is required to fully elucidate the roles of both attenuated apoptosis and the ubiquitin-proteosome pathways in lens development, current evidence strongly suggests both processes, and possibly others, are important for the development of a functionally transparent lens (Wride 2011; Greiling and Clark 2012).

Lens function depends on a smooth gradient of refractive index that facilitates the passage of light from outside of the fish through the transparent lens to focus on the retina

(Greiling et al. 2009). The lens has the most densely packed protein composition of any tissue in the body. The densely packed, organized protein structure allows for the high index of light refraction. The lenses of aquatic animals adjust for light refraction differently than terrestrial animals because in terrestrial animals the majority of light refraction occurs between the air and the cornea. However, in zebrafish and other aquatic species, the index of refraction is very similar between water and the cornea and little light refraction occurs here, leaving most of the refraction to occur while passing through the lens (Greiling and Clark 2012). The most abundant group of proteins found in the lens is the family of proteins called crystallins. There are three subclasses of crystallins, α, β and γ, with each group consisting of several different protein isoforms identified alphabetically based on structural similarities. It is the high level of structural organization in which these proteins are packed into the fibre cells that is primarily responsible for the refraction of light through the lens. Maintaining tightly controlled ratios in the crystallin complement of the lens is necessary to ensure the formation of a transparent lens. The importance of this is highlighted by the cataract phenotype observed in the zebrafish *cloche* mutant. The development of cataracts in the lenses of *cloche* mutants between 2-4 dpf has been attributed to reduced solubility of γ-crystallin proteins resulting from reduced levels of αA-crystallin chaperones (Goishi et al. 2006). Mapping of the zebrafish lens proteome demonstrated that the crystallin complement of the adult zebrafish lens is similar to that observed in the mammal but the amounts and subclasses of crystallins present in the lens differ between these species (Greiling and Clark 2008; Posner et al. 2008; Greiling et al. 2009). The zebrafish has an extra αB-crystallin isoform and a series of γ M and γ N-crystallins which have been identified only in lenses of aquatic species (Posner et al. 2008). The levels of α-crystallins are lower in zebrafish but the ratio of αA to α B-crystallins are the same compared to humans (Greiling and Clark 2008). β-crystallins are the earliest detected crystallins in zebrafish lenses and remain the most abundant through the larval stage of zebrafish. As the fish ages both α and γ-crystallins increase in abundance while the levels of β-crystallins remain the same (Greiling et al. 2009). This is different from mammals where β-crystallins are not detected in the mouse embryonic lens until after birth. Zebrafish lenses, like those of many other fish species, contain higher levels and more diverse γ-crystallins proteins than mammals. The high level of γ-crystallins, which are smaller than α-crystallins, allows for tighter packing of the proteins, creating a harder lens and providing a high refractive index of light required for underwater vison (Posner et al. 2008). The

unique γM-crystallins, which are only found in aquatic lens, are the most abundant γ-crystallin and it has been suggested their role is to facilitate cold tolerance in the lens (Posner et al. 2008).

Cytoskeleton proteins in the lens are important for maintaining cellular processes such as cell division, migration and elongation. Disruption in the expression of cytoskeleton proteins, such as Bfsp, vimentin and connexins, results in reduced lens opacity and cataract formation (Fujimoto et al. 1999; Cheng et al. 2004; Shi et al. 2009; Mou et al. 2010; Greiling and Clark 2012). Intermediate filament proteins, Bfsp, are the most abundant non-crystallin protein in the lens (Greiling and Clark 2012). These proteins function to connect crystallins as well as form lens specific beaded filament networks that connect lens fibre cells to each other (Shi et al. 2009; Greiling and Clark 2012). While bfsps have been identified in zebrafish lenses, little work has been done looking at the function of this protein in zebrafish (Greiling and Clark 2012). In mouse lenses, expression of *Bfsp* is regulated by HSF4. Lens fibre cells in an *hsf4* knockout mouse line were disordered, and irregularly shaped with protein inclusions all of which were attributed to decreased expression of *Bfsp* (Shi et al. 2009). Vimentin is another protein that forms intermediate filament networks between the nucleus and plasma membrane in lens cells. Vimentin filament assembly was characterized in zebrafish lenses but most of what is known about the function of this protein in the lens has been determined in mammalian models (Greiling and Clark 2012). While vimentin can be seen in the epithelial cells of mature lenses, it is only present in fibre cells prior to denucleation (Mou et al. 2010). Vimentin is upregulated in the lenses of *hsf4* knockout mice, resulting in disruption of the denucleation process. Overexpression of vimentin in mouse lens fibre cells mimicked the phenotype observed in *hsf4* knockout mice. Disruption of lens fibre denucleation resulted in retained DNA fragments, contributing to cataract formation in these *hsf4* knockout mice (Mou et al. 2010). Connexin proteins are vital for the formation of intercellular channels and gap junctions, which are important for intracellular communication and fluid transport in the lens (Cheng et al. 2004; Greiling and Clark 2012). The four isoforms of connexins identified in zebrafish are expressed in a developmentally and cell specific manner with all but one isoform being exclusively expressed in the epithelial cells of mature lenses. Connexins are also important for normal development of lenses as knockdown of only a single isoform of this protein in zebrafish and mice resulted in the formation of a cataract (Cheng et al. 2004).

2.0 HYPOTHESES AND OBJECTVES

Recent studies by our group have shown that regulation of *hsp70* gene expression by HSF1 is important for development of the zebrafish eye (Blechinger et al. 2002; Evans et al. 2005, 2007; Evans 2006). The knockdown of *hsf1* and *hsp70* by antisense morpholino-modified oligonucleotides resulted in a consistent and reproducible small eye phenotype in microinjected embryos due primarily to a reduction in lens size. The number of animals exhibiting the small lens phenotype following *hsf1* knockdown was similar to that seen in *hsp70* knockdowns but the small lens phenotype was less severe in *hsf1* knockdowns (Evans et al. 2007). The reduced severity of the small lens phenotype in the *hsf1* knockdowns suggested that there are likely other factors involved in regulating *hsp70* expression in the zebrafish lens. However, similar experiments using morpholino knockdowns for zebrafish *hsf2* showed no effect on eye development (Evans et al. 2007).

The zebrafish genome contains genes encoding at least three heat shock factors; two isoforms of HSF1 (Råbergh et al. 2000), HSF2 (Yeh et al. 2006) and a novel, uncharacterized orthologue of HSF4. Several studies have demonstrated that multiple HSFs work together to stimulate gene expression. Many genes that are targets for HSFs contain multiple HSE in their promoters and it appears that different combinations of HSFs may simultaneously bind to the promoters of target genes to produce specific responses (Åkerfelt et al. 2007; Östling et al. 2007). Both HSF1 and HSF4 are required for mammalian lens development and have been shown to both co-regulate and independently regulate several genes involved in this process (Fujimoto et al. 2004, 2008; Min et al. 2004; Smaoui et al. 2004; Mellersh et al. 2006; Shi et al. 2009; Abane and Mezger 2010). This suggests that the novel zebrafish HSF4 could play a similar role and work with HSF1 to regulate the process of lens development in zebrafish. Therefore, the overall focus of the work presented in this thesis was to characterize the novel zebrafish HSF4 and compare its expression to that of other HSFs with a focus on identifying the roles of these proteins in zebrafish lens development. Three hypotheses were developed for the purpose of this investigation.

Hypothesis 1: Zebrafish HSF4 represents a novel subfamily of vertebrate HSFs.

The zebrafish *hsf4* identified in an expressed sequence tag (EST) library codes for a severely truncated form of HSF4 and is much shorter in length than any other known vertebrate HSF. The first objective was to determine how similar this truncated protein was to known members of the HSF family both in structure and function and to confirm this putative protein is an ortholog of mammalian HSF4. An extensive *in silico* analysis of both the gene and protein sequences of zebrafish HSF4 was performed to determine the homology of this protein to other HSF proteins. As a second objective, the ability of HSF4 to bind to HSE sequences was determined by EMSA.

Hypothesis 2: Zebrafish HSFs are expressed in distinct developmental patterns in zebrafish embryos.

Different expression patterns of HSFs have been identified in mammals. A basal level of HSF1 expression is observed in most tissues during mammalian development while HSF2 and HSF4 show tissue and developmentally specific patterns of expression (Abane and Mezger 2010; Björk and Sistonen 2010). Studies investigating the expression of HSFs in zebrafish are limited and to date, a comprehensive analysis of zebrafish HSF expression in developing embryos and adult tissues has not been performed (Råbergh et al. 2000; Yeh et al. 2006). Therefore, to test this hypothesis, the first objective was to characterize the expression of all three zebrafish HSF genes across several developmental stages as well as in adult tissues using quantitative PCR and *in situ* hybridization techniques. A second objective was to characterize the abundance of HSF proteins in developing zebrafish by Western blot analysis using custom designed antibodies for zebrafish HSFs.

Hypothesis 3. Zebrafish HSF4 is involved in lens development and also works cooperatively with other HSFs to regulate lens-specific gene expression.

The transcriptional regulation of target genes by HSF4 is crucial to lens development in mammals (Fujimoto et al. 2004, 2008; Min et al. 2004; Smaoui et al. 2004; Mellersh et al. 2006). As well, HSF4 co-regulates the expression of certain genes in the lens with other HSFs (Shi et al. 2009; Abane and Mezger 2010). In zebrafish, no HSF4 target genes have been identified for lens development. Therefore, to test this hypothesis, the effect of a knockdown of *hsf4* expression using MO antisense technology on lens development in zebrafish embryos was investigated. The first objective was to determine if knockdown of *hsf4* produced a small lens phenotype, similar to that observed for *hsf1* and *hsp70* knockdowns. A second objective was to investigate the effect of a knockdown of either *hsf4* or *hsf1* on the expression of *hsp70* and *fgf1*, two potential gene targets required for lens development using qPCR.

3.0 MATERIALS AND METHODS

3.1 Zebrafish maintenance and embryo collection

Procedures and protocols followed for maintaining and breeding zebrafish (*Danio rerio*) are established and detailed in "*The Zebrafish Book*" (Westerfield 1995). Adult wild-type zebrafish stocks were obtained through local pet stores and housed in an Aquatic Habitat™, system (Aquatic Eco-Systems, Apopka, FL) at a density of 15-25 fish per five gallon tank with an approximately equal female to male ratio. Tanks were filled with charcoal filtered potable water (fishwater) and maintained at a temperature between 26-28°C, with a constant photoperiod of 14 hours of light followed by 10 hours of darkness. Adult fish were fed once per day with a mixture of frozen bloodworms (San Francisco Bay Brand, Newark, CA), dried flake food (Nutrifin, Rolf C. Hagen (USA) Corp., Mansfield, MA; Tetra United Pet Group, Blacksburg, VA), brine shrimp (Hikari, Hayward, CA) or floating pellets (Aquatic Eco-Systems, Apopka, FL).

One gallon breeding tanks maintained at 26-28^{o}C were used for embryo collection. The bottoms of the tanks were lined with marbles and contained a single artificial plant. The constant photoperiod, described above, kept the fish in a breeding state and the fish began spawning in the morning when the light cycle began. A total of six to eight male and female adult zebrafish were moved from the stock tanks into each of the breeding tanks the afternoon prior to breeding. Eggs were siphoned from between the marbles after spawning the following morning. The number of eggs collected from each tank was variable ranging from 50-1000 eggs.

Embryos were cleaned by removing debris and then transferring into 25 mm Petri dishes containing fishwater supplemented with 0.05% methylene blue (used as a fungicide). Methylene blue was not used when embryos were maintained for periods longer than 72 hours. Embryos were maintained in an incubator (Revco Scientific, Asheville, NC) set to 28^{o}C under the same photoperiod described above for adult zebrafish. The water was changed daily and embryos were monitored under a dissecting microscope to remove any unfertilized eggs or dead embryos. All surviving embryos were staged using the published staging guide by Kimmel et al. (1995). Developmental stages are described as hours post fertilization (hpf) at 28^{o}C.

Embryos were dechorionated either manually under a low powered dissecting microscope using fine forceps or by enzymatic reaction with pronase. The procedure for dechorionating embryos with pronase was performed by replacing the fish water in a petri dish with a solution of

0.5 mg/mL pronase in fishwater and incubating for 12 minutes at 28°C. The pronase was removed and embryos were washed five times in fishwater to remove the chorions and residual pronase (Westerfield 1995; Evans 2006).

3.2 Embryonic heat shock

Embryos were heat shock treated by transferring them to a 25 mm Petri dish in fishwater and incubating at 36°C for one hour (heat shock was started one hour prior to the developmental stage desired for analysis). Samples were either snap frozen immediately on dry ice for RNA extraction or returned to the 28°C incubator for 5 hours prior to collection for protein extraction. Samples requiring a recovery period after heat shock prior to analysis were returned to the 28°C incubator for 15 hours prior to collection for RNA analysis as described above.

3.3 Imaging zebrafish embryos

Whole embryos were placed in a depression slide for microscopy. Images of whole embryos were captured using a Nikon Coolpix digital camera mounted on a Nikon Eclipse E600 photomicroscope (Nikon Instruments Inc., Melville, NY). Embryos and larvae from whole mount *in situ* hybridization treatments were screened and imaged using either bright field or fluorescent microscopy on a Nikon Y-FL epifluorescence attachment. All images were processed using Adobe Photoshop Elements (version 6) and arranged in Powerpoint (Microsoft Office 2003, Microsoft Corporation, Redmond, WA).

3.4 Analysis of *hsf4* sequences *in silico*

Heat shock factor sequences for genomic DNA, cDNA and proteins were identified and obtained from both the GenBank and Ensembl sequence databases. Alignments of heat shock factor nucleic acid and amino acid sequences were performed using the ClustalW alignment program (Larkin et al. 2007). Boxshade software was used to outline similarities between the aligned sequences (Hofmann and Baron 2008; http://www.ch.embnet.org/software/BOX_form.html). Structural domains within the zebrafish HSF2 and HSF4 protein sequences were identified using NCBI's Conserved Domain Database (Marchler-Bauer et al. 2009) as well as the SMART protein analysis software (Ponting et al. 1999). The location of the protein domains identified using the above software was confirmed

by alignment of zebrafish protein sequences with HSF proteins from other species deposited in Genbank. The amino acid identity of protein domains was determined by comparing the amino acid sequence of only the DBD, HR-A/B and HR-C from each of the three zebrafish HSFs to HSF protein sequences from 9 other species found in GenBank using BLAST software available on the NCBI website (Altschul et al. 1990). The Coils software version 2.2 (ch.EMBnet.org) was used to predict the probability of amino acid sequences within HSF proteins to adopt a coiled-coil conformation. The software compares the target sequence to a database of known sequences that produce coiled-coil structures by scanning amino acid sequence in a 14, 21 and 28 amino acid windows and generating a comparative score. Sequences with a score closer to 1 have a higher probability of adopting a coiled-coil conformation. Phylogenetic analysis was performed for the three zebrafish HSF sequences and 24 other HSF sequences that were identified in GenBank and Ensembl using the Phylogeny.fr software (Dereeper et al. 2008, 2010). Species and gene accession numbers used for sequence alignments and phylogenetic analysis are shown in Table A1 of the Appendix.

3.5 Molecular biology techniques

3.5.1 Plasmid preparation

The procedure outlined in this section for isolating and maintaining plasmid stocks was performed following standard molecular biology protocols from Sambrook and Russell (2001). Single bacterial colonies were isolated by streaking glycerol stocks or newly transformed *E. coli* on Luria-Bertaini (LB) plates (10 g tryptone, 5 g yeast extract, 10 g NaCl, pH 7.0 solution plus 15 g of agar per litre of deionized water, sterilized by autoclaving, and stored at 4°C) containing appropriate antibiotic selection at 37°C overnight. Single colonies isolated from the plates were grown overnight at 37°C with agitation in LB media (same as above without the agar) under appropriate antibiotic selection. Glycerol stocks (400 µL of bacterial culture plus 100 µL of sterile 80% glycerol) were prepared from cultures and stored at -80°C.

Plasmids were isolated from bacterial cultures using the alkaline lysis miniprep procedure (Sambrook and Russell 2001). Briefly, inoculated cultures of *E. coli* containing the desired plasmid were incubated in LB media under appropriate antibiotic selection with vigorous shaking overnight at 37°C. Following incubation, cultures were placed on ice for 10 minutes to

stop further growth. Three to five mL of culture were transferred to a 15 mL centrifuge tube and centrifuged for 5 minutes at 3000 rpm in a swinging bucket centrifuge (Damon IEC - ThermoFisher Scientific, Waltham, MA) after which the supernatant from each sample was discarded. The cell pellet was resuspended in 200 μL of Solution I (50 mM glucose, 25 mM Tris pH 8.0, 10 mM EDTA pH 8.0), mixed gently by pipetting and incubated at room temperature for 5 minutes. Cells were then lysed in 400 μL of freshly made Solution II (1% SDS w/v, 0.2 M NaOH) by vigorous mixing then placing the tube on ice for 5 minutes. The lysis solution was neutralized using 300 μL of ice cold Solution III (3M potassium acetate pH 6.0), mixed by inversion and placed on ice for 5 minutes. The resulting mixture was then centrifuged at 4°C in a Sorvall microfuge (model Mc12) for 5 minutes at maximum speed (Sorvall, ThermoFisher Scientific, Waltham, MA). The supernatant was transferred to a sterile 1.5 mL microfuge tube and 20 μg of RNase A (Fermentas; ThermoFisher Scientific, Waltham, MA) was added followed by incubation for 1 hour at room temperature. Proteins were removed from the retained supernatant with a 200 μL phenol extraction (Tris buffered phenol, pH 8.0) followed by the addition of 200 μL of chloroform. DNA was precipitated by adding 1 mL of 95% EtOH (v/v) to the supernatant and incubating at -20°C for 20 minutes or overnight. Samples were centrifuged in a microfuge (Sorvall) at 4°C for 10 minutes at maximum speed, and ethanol was removed by aspiration. The pellet was then washed with 500 μL of ice cold 70% EtOH (v/v) and centrifuged at 4°C again for 5 minutes at maximum speed. The 70% EtOH (v/v) was removed by aspiration and the DNA pellet was allowed to air dry for 20 minutes at room temperature before dissolving it in 50-100 μL of triple distilled water. Isolated plasmids were stored at -20°C until use or at -80°C for long term storage.

3.5.2 Restriction enzyme digestion of plasmids

Purified plasmids (section 3.5.1) were digested with appropriate restriction enzymes following manufacturer's instructions for the specific enzymes to obtain DNA fragments required for cloning. Restriction enzyme analysis was also used to confirm the successful ligation of DNA into a vector. Analysis of restriction enzyme digestion was performed by agarose gel electrophoresis (section 3.5.3).

3.5.3 DNA gel electrophoresis

Concentration and DNA integrity was analysed by agarose gel electrophoresis as described in Sambrook and Russell (2001). Briefly, an agarose gel was prepared by melting agarose (Fermentas) in 1× TAE buffer (40 mM Tris-acetate, 1 mM EDTA) and adding 1 µL of ethidium bromide (10 mg/mL). An agarose gel with a concentration of 1% (w/v) was used to resolve DNA fragments with an expected size larger than 500 base pairs (bp). For smaller DNA fragments, a 1.5% (w/v) agarose gel was used. A commercially prepared DNA ladder (GeneRuler™ 1kb DNA ladder Plus, Fermentas) was loaded onto each gel as a standard. DNA standards and samples in 5× loading buffer (0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF, 30% (v/v) glycerol) were loaded directly into the gel. Gels electrophoresis was performed in 1× TAE buffer run at approximately 125 volts using an Owl electrophoresis apparatus (ThermoFisher Scientific, Waltham, MA). The size and concentration of electrophoresed DNA samples was estimated by visual comparison of the sample DNA band to the bands on the DNA ladder.

3.5.4 Purification of DNA from agarose gel

DNA fragments to be purified were identified using an ultraviolet transilluminator and specific bands were excised from the agarose gel using a razor blade. The DNA was purified from the agarose gel using a GeneJET™ gel extraction kit or DNA extraction kit as per manufacturer's instructions (both Fermentas). DNA was eluted by adding 20-30 µL nuclease free water to the purification column, centrifuged for 30 seconds at maximum speed in an microfuge (model 5418; Eppendorf Canada, Mississauga, ON) and eluant was stored at -20°C until use.

3.5.5 DNA sequencing

The identity of both purified DNA fragments and DNA inserts within purified plasmid were confirmed by both forward and reverse sequencing using T7, SP6 or M13 primers at the NRC Plant Biotechnology Institute (Saskatoon, SK).

3.5.6 RNA extraction

Embryos/larvae were collected in 1.5 mL microfuge tubes, fishwater was aspirated and discarded, and the sample was snap frozen in powdered dry ice and stored at -80^{o}C. RNA was extracted from these samples using the RNeasy RNA extraction kit (Qiagen Inc., Toronto, ON). Briefly, embryos were resuspended in 600 µL of RLT buffer (containing guanidine thiocyanate) and homogenized using a pestle. Samples were processed following the instructions from the RNeasy kit including the on-column DNase treatment. RNA was eluted from the column in 30 µL of RNase free water and stored at -80°C until use.

Adult zebrafish tissues were collected as described by Gupta and Mullens (2010). Tissues were immediately placed in RNAlater (Qiagen) solution, incubated at 4°C overnight and then stored at -20°C until RNA extraction. Immediately prior to RNA extraction, the adult tissues were removed from the RNAlater solution, frozen in liquid nitrogen and ground to a powder using a pestle. Tissue was then immediately resuspended in 600 µL RLT buffer from an RNeasy RNA extraction kit (Qiagen), disrupted using a 2 mL Potter-Elvehjem mortar and pestle homogenizer (WHEATON, Millville, NJ, USA) and collected into a 1.5 mL microfuge tube. Isolation of RNA from these samples was then performed following the instructions from the RNeasy RNA extraction kit including the on-column DNase treatment (Qiagen). RNA was eluted from the column in 30 µL of RNase free water and stored at -80°C until use.

3.5.7 RNA quantification and gel electrophoresis

The concentration of RNA was quantified using a SmartSpec Plus spectrophotometer (Bio-Rad Laboratories, Inc., Hercules, CA). A sample of RNA was diluted into 10 mM Tris-HCl (pH 7.0) and dispensed into trUView cuvettes (Bio-Rad Laboratories) and absorbance was measured at 260 nm wavelength. The concentration of RNA in each sample was determined using the following equation: 40 µg/mL x A_{260} x dilution factor = µg/mL RNA. The purity of RNA samples was determined by calculating the ratio of the absorbance at 260 nm to the absorbance at 280 nm. A ratio between 1.6 and 2.1 was accepted as sufficiently pure RNA for use in subsequent experiments.

The quality of RNA was analyzed by formaldehyde gel electrophoresis. Solutions and buffers were prepared according to the RNeasy RNA extraction kit (Qiagen). A 10× MOPS buffer (200 mM MOPS, 50 mM sodium acetate, 10 mM EDTA, in nuclease free water and pH

adjusted to 7.0) was prepared for use in RNA analysis. Four microlitres of RNA was denatured in RNA loading buffer (4 mM EDTA pH 8.0, 2.7% (v/v) of 12.3 M formaldehyde, 20% (v/v) glycerol, 30.8% formamide, 0.16% saturated bromophenol blue, 40% (v/v) 10× MOPS buffer), at 70°C for 10 minutes and placed on ice. Two microlitres of 0.5 mg/mL ethidium bromide was added to each sample, mixed by pipette, and loaded into a 1.2% (w/v) formaldehyde denaturing agarose gel (agarose dissolved in 1× MOPS buffer containing 2.7% of 12.3 M formaldehyde). Samples were electrophoresed at 120 volts for 1 hour in 1× MOPS buffer and RNA bands were visualized under ultraviolet light. The RNA was accepted as adequate quality for use in subsequent analysis if the 28S and 18S bands appeared to be at least a 1:1 ratio by visual inspection.

3.5.8 cDNA synthesis

Synthesis of cDNA was performed using the RevertAid H minus first strand cDNA synthesis kit (Fermentas). Each reaction was performed using 2 µg of RNA extracted from an embryo/larvae sample or 0.5-2 µg of RNA from adult tissue samples. Reactions were performed as per kit instructions using only oligo(dT) primers and RevertAid H minus M-MuLV reverse transcriptase. Samples were stored at -20°C until use.

3.6 Generating heat shock proteins in a bacterial expression system

3.6.1 Plasmid stocks

Bacterial glycerol stocks of plasmids containing full length cDNA clones for *hsf1*, *hsf2*, and *hsf4* were purchased from the I.M.A.G.E. consortium. The I.M.A.G.E. clone numbers and vectors are summarized in Table 3.1 below. Plasmids were isolated from bacterial stocks as described in section 3.5.1.

Table 3.1. Summary of I.M.A.G.E. cDNA clones.

Name	GenBank Accession #	I.M.A.G.E. clone ID	Vector	Selectable Marker
hsf1	NM 131600	8754667	pExpress1	Ampicillin
hsf2	NM 131867	7002403	pExpress1	Ampicillin
hsf4	NM 001013317	7143689	pDNR-LIB	Chloramphenicol

3.6.2 Preparation of the pRSET expression vector

The pRSET bacterial expression system (provided by Dr. Ovsenek, Department of Anatomy and Cell Biology, University of Saskatchewan) was used to generate heat shock transcription factor proteins *in vitro*. Expression of DNA fragments inserted in the multiple cloning site of pRSET is regulated by a T7 promoter site upstream of an ATG translation initiation codon (pRSET user manual, Invitrogen, Life Technologies, Carlsbad, CA). The pRSET vector is designed to generate the protein of interest as a fusion protein attached to a 6 histidine (His) tag at the amino terminus of the protein. The pRSET vector is used in conjuction with the BL21(DE3)pLysS cell line which contain a "helper" DE3 bacteriophage lambda lysogen (*E. coli* genotype F-, *omp*T *hsd*SB (rB-mB-) *gal dcm* (DE3) pLysS (CamR)). The lambda lysogen contains the gene for T7 RNA polymerase under control of the *lac*UV5 promoter as well as a small portion of the *lac*Z gene. The *lac* repression of the T7 RNA polymerase gene is lifted with the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to the growth media thereby allowing the DNA insert from the pRSET plasmid to be transcribed (Anonymous 2010a).

Polymerase chain reaction (PCR) was used to amplify the open reading frame from each of the full length cDNA clones identified in Table 3.1. Primers were designed to incorporate an *Xho*I restriction enzyme site on the 5' end and a *BstB*I restriction enzyme site on the 3' end of the PCR product to facilitate directional cloning of the PCR product into the pRSET-A expression vector. An additional four nucleotides were added to the 5' ends of each primer to act as an anchor and facilitate restriction enzyme binding to the PCR product. Primer sequences used to amplify hsf sequences were as follows:

hsf1 forward 5'-CGGGCTCGAGATGGAGTATCACAG-3' and
reverse 5'-GCCCTTCGAACTATGATAGTTTGG-3';
hsf2 forward 5'-CGGGCTCGAGATGAAACACAGCTCG-3' and
reverse 5'-GCCCTTCGAATCAGATATCAAGCGG-3';
hsf4 forward 5'-CGGGCTCGAGATGCAGGAGAACCC-3' and
reverse 5'-GCCCTTCGAATCACCTTGATTCC-3'. All reactions contained PCR buffer (Invitrogen), 0.2 mM dNTPs (Invitrogen), 3 mM $MgCl_2$ (Invitrogen), 9-15 ng of plasmid template, 0.5 μM primers (Sigma-Aldrich Corporation, St. Louis, MO), and 1 μL of *Taq* polymerase (in house). Polymerase chain reactions were all performed using an MJ Research Inc. PTC-100™ programmable thermocycler (MJ Research, Waltham, MA) with the following

reaction conditions: initial denaturing at 95°C for 2 minutes, 40 cycles of (95°C for 30 seconds, annealing at 55°C for *hsf2* and *hsf4*, 50°C for *hsf1* for 30 seconds, 72°C elongation for 2 minutes), followed by a final elongation cycle of 15 minutes at 72°C. PCR products were purified as per section 3.5.4 and stored at -80^{o}C until use.

Products from the PCR reactions were initially cloned into the pCR®4-TOPO vector using the TOPO TA Cloning® Kit for Sequencing following the manufacturer's instructions (TOPO TA cloning manual, Invitrogen). The pCR®4-TOPO vector is supplied as a linearized plasmid with single deoxythymidine residues on the 3' ends the vector. These deoxythymidine residues can then ligate to the 3' deoxyadenosine residues present on PCR products generated with *Taq* polymerase. The Topoisomerase I enzyme, which is covalently bound to the vector, then facilitates the final incorporation of the PCR product into the vector.

Briefly, the cloning procedure was performed by combining commercially prepared pCR®4-TOPO vector with 8-15 ng of purified PCR product and incubating for one hour at room temperature then placed on ice. Chemically competent DH5α™ (*E. coli* genotype F-φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(rk-, mk+) *pho*A *sup*E44*thi*-1 *gyr*A96 *rel*A1 λ^{-}) (Invitrogen) were then transformed with 2 µL of the TOPO/HSF vector using a heat shock at 42°C as per manufacturer's instructions. Transformed cells were streaked on LB plates containing 50 µg/mL kanamyacin (section 3.5.1). Plasmids were isolated from single colonies (section 3.5.1) and were then analysed by restriction digestion (section 3.5.2) and sequencing (section 3.5.5) to confirm the identity and orientation of the DNA insert. Glycerol stocks of bacterial colonies that contained the insert were prepared (section 3.5.1) and stored at -80°C.

The HSF DNA inserts were subcloned from the pCR®4-TOPO vector into the pRSET expression vector. Both the isolated TOPO/HSF plasmids and empty pRSET vector were digested with *Xho*I and *BstB*I restriction enzymes, and then gel purified as per sections 3.5.2 and 3.5.4. The purified DNA insert was subsequently ligated into the pRSET-A vector (5' to 3') using T4 DNA ligase (Invitrogen) as per manufacturer's instructions. Briefly, the purified inserts and pRSET plasmid were combined (6:1 insert to vector ratio, both heated to 55°C before combining) with 0.1 units of T4 DNA ligase in a 20 µL reaction containing ligase buffer and incubated overnight at 4°C. Chemically competent DH5α™- *E. coli* (see genotype above) were transformed using a 42°C heat shock with 4 µL of the ligated products as per manufacturer's

instructions. Bacterial were then streaked on LB plates containing 50 µg/mL ampicillin. Plasmids were isolated from single colonies (section 3.5.1) and analysed by restriction enzyme digestion (section 3.5.2) and sequencing (section 3.5.5) to confirm the identity and orientation of the DNA insert. Glycerol stocks of bacterial colonies that contained the insert were prepared (section 3.5.1) and stored at -80°C.

3.6.3 Purifying HSF proteins from a bacterial expression system

Expression of proteins was performed as per manufacturer's instructions (pRSET user manual, Invitrogen) using Super Optimal Broth (SOB; 20 g tryptone, 5 g yeast extract, 0.5 g NaCl, 0.186 g KCl per litre of deionized water, pH to 7.0 and sterilized by autoclaving). After autoclaving, 10 mL of 1 M sterile $MgCl_2$ solution was added and the final solution was stored at 4°C. Plates of SOB were prepared by adding 15 g of agar per litre of media and autoclaving.

Chemically competent BL21(DE3)pLysS cells were transformed with purified pRSET plasmid containing an HSF DNA insert (section 3.6.2) and the resulting transformations were plated on SOB agar plates containing 50 µg/mL ampicillin and 35 µg/mL chloramphenicol. One colony was selected to inoculate a starter culture of 5 mL of SOB media containing the above antibiotics and then incubated at 37°C. Approximately eighteen hours later, 25 mL of SOB media containing no antibiotics was inoculated with 2-3 mL of the starter culture. The newly inoculated culture was incubated at 28°C until the culture had an OD_{600} of 0.4-0.6 (approximately 4 hours) at which time protein expression was induced with a final concentration of 1 mM IPTG. Bacteria were pelleted by centrifugation and the supernatant removed. Pellets were disrupted in 4 mL of phosphate buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH to 7.4) by vortexing, a 30 second sonication and a single freeze/thaw event at -80°C after which the cell debris was pelleted by centrifugation. Both the supernatant and pellet were retained and stored at -80°C.

Proteins were purified by adding 200-250 µL of Profinity™ IMAC Ni-Charged Resin (Bio-Rad Laboratories) suspended in PBS plus 3 M NaCl and10 mM imidazole to 2-4 mL of the retained supernatant. The mixture was incubated overnight at 4°C with gentle rocking to maintain the resin in suspension. The resin was then allowed to precipitate by gravity and washed a total of three times in fresh buffer (above). The purified protein was eluted from the

resin with 250 µL of PBS plus 3 M NaCl and 250 mM imidazole. Glycerol was added to the purified proteins at a final concentration of 5% and the total solution was stored at -80°C.

The protein concentration for each sample was determined using a Pierce 660 nm Protein Assay kit with the associated Ionic Detergent Compatibility Reagent (ThermoFisher Scientific, Waltham, MA) as per manufacturer's instructions. Briefly, protein was assayed using a dye based colorimetric analysis to determine unknown protein concentrations by extrapolating the absorbance of the sample from a standard curve of known concentrations of bovine serum album (BSA).

3.7 Western blot analysis of proteins

3.7.1 Antibodies

Zebrafish specific polyclonal antibodies were custom designed and synthesized commercially by ABR Affinity BioReagents™ (Colorado, U.S.A.). Briefly, rabbits were immunized against protein-specific peptides consisting of 13-15 amino acids (shown in brackets) designed specifically for zebrafish proteins HSF1 (SPDPVKETESGVD), HSF2 (HQEAEPKTTPDVS) and HSF4 (SSSMALQMQAEESR). Animal sera were purified by affinity chromatography to produce antibody stocks. Upon receipt, purified antibodies were thawed on ice, dispensed into microcentrifuge tubes in 5 µL aliquots and stored at -80^{o}C until use.

The appropriate dilution to use for each antibody was determined experimentally using protein extracted from zebrafish embryos and the results are summarized in Table 3.2. The specificity of the antibodies was confirmed experimentally. Briefly, this was achieved by using HSF proteins purified from a bacterial expression vector (section 3.6.3) to determine specificity of antibodies to each of these proteins using the Western blot technique (section 3.7.3). As well, a commercially prepared HIS-tag antibody was used at a dilution of 1:10,000, to confirm detection of proteins purified from the pRSET expression vector (Invitrogen, supplied by Dr. Troy Harkness, Department of Anatonmy and Cell Biology, University of Saskatchewan).

Table 3.2. Expected molecular masses of proteins targeted by custom designed zebrafish specific antibodies. Peptide sequences used to generate the antibodies were common to multiple isoforms of the HSF1 and HSF4 proteins and therefore detection of more than one band is expected.

Protein Designation	Accession Number	Amino Acid Number	Molecular Mass (kDa)	Primary Antibody Dilution
HSF1a	AAF72750	512	57	1:3,000
HSF1b	NP 571675.1	538	59	
HSF2	NP 571942.1	489	55	1:10,000
HSF4a	NP 001013335.1	286	32	1:5,000
HSF4b	N/A	441	50	

3.7.2 Isolation and quantification of protein from embryos

Zebrafish embryos were first dechorionated with pronase as described in section 3.1. Next, embryos younger than 80 hpf were deyolked prior to protein extraction. Deyolking was performed by first collecting 35-70 dechorinoated embryos in a 1.5 mL centrifuge tube. Water was removed and the embryos were resuspended in approximately 100 µL of ice cold PBS containing 300 µM PMSF, 2 µM leupeptin, 0.15 µM aprotinin, 1 mM EDTA and 1 mM EGTA. The embryos were pipetted gently up and down through a p200 pipet tip (Eppendorf) to break the yolk membrane. The volume of buffer in each sample was made up to 1 mL with the PBS/protease inhibitor solution (above) and the embryos were washed by gently inverting the tube 2-4 times. The embryos were pelleted at 800 ×g in an microfuge (Eppendorf) at room temperature, the wash buffer was removed and embryos were homogenized manually with a pestle in 100 µL 2× Laemmli buffer (0.125 Tris-HCl pH 6.8, 4% w/v SDS, 20% (v/v) glycerol, 10% 2-mercaptoethanol, 0.02% bromophenol blue; Laemmli et al. 1970). Protein extracts were stored at -80°C until use.

The protein concentration for each sample was determined using a Pierce 660 nm Protein Assay kit with the associated Ionic Detergent Compatibility Reagent (ThermoFisher Scientific) as described in section 3.6.3.

3.7.3 SDS-PAGE and blotting technique

Analysis of protein by SDS-PAGE and Western blotting techniques was performed following standard molecular biology protocols (Sambrook and Russell 2001). Proteins were separated on a 10% SDS-polyacrylamide denaturing resolving gel (29:1 acrylamide:bis-acrylamide in 0.375 M Tris-HCl pH 8.8, 0.1 5 (w/v) SDS with 0.025% (w/v) of each of TEMED and ammonium persulfate) overlayed with a 4% SDS-polyacrylamide stacking gel (29:1 acrylamide:bis-acrylamide in 0.125 M Tris-HCl pH 6.8, 0.1 5 w/v SDS with 0.025% w/v each of TEMED and ammonium persulfate) using a Mini-Protean® III electrophoresis apparatus (Bio-Rad Laboratories) as per manufacturer's instructions.

Protein samples were prepared for analysis by heating the samples at 95^{o}C for 5 minutes then cooling on ice for two minutes immediately prior to loading samples onto the gel on an equal protein basis. A commercially prepared protein ladder (Fermentas) was also loaded on every gel to be used for molecular mass comparison. The gel was then electrophoresed in 1× running buffer (0.025 M Tris-HCl, 0.192 M glycine and 0.1% SDS, buffer pH 8.3) at 150 volts for approximately 2 hours. After electrophoresis the resolving gel was placed in cold Towbin transfer buffer (0.025 M Tris-HCl, 0.192 M glycine at pH 8.3 plus 0.0375% SDS and 20% MeOH; Towbin et al. 1979) for 25 minutes. Gels were either stained with Coomassie blue (0.25% (w/v) Coomassie brilliant blue R-250, 50% (v/v) MeOH, 10% (v/v) glacial acetic acid) to detect protein bands or transferred to a PVDF membrane for Western blotting. Transfer of the protein bands from the acrylamide gel to the PVDF membrane was performed with Towbin buffer using the Trans-Blot® Semi-Dry transfer apparatus as per manufacturer's instructions (Bio-Rad Laboratories). Transfer was performed at 25 volts for 60 minutes.

Immunodetection was performed by first incubating the membrane in a blocking solution of 5% (w/v) blotting grade blocker non-fat dry milk (Bio-Rad Laboratories) in TTBS buffer (20 mM Tris-HCl, 137 mM NaCl, 0.1% v/v Tween 20, pH 7.4) for at least 1 hour to prevent non-specific binding. The blot was removed from the blocking solution and incubated with primary antibody diluted in fresh blocking solution (dilutions shown in Table 3.2) overnight at 4^{o}C with mild agitation. The antibody was then removed and the blot was washed 4 × 5 minutes in TTBS solution with mild agitation. A goat anit-rabbit IgG secondary antibody conjugated to horseradish peroxidase (Bio-Rad Laboratories) diluted 1:20,000 in TTBS was added to the blot and incubated at room temperature for 1 hour. After the second incubation, the

blot was washed 4 × 5 minutes with TTBS. Western Lightning™ Chemiluminescence Reagent Plus (PerkinElmer Inc., Waltham, MA) was added to the membrane to detect the secondary antibody as per manufacturer's instructions. Finally the chemiluminescence was detected using HyBlot CL™ autoradiography film (Denville Scientific Inc., South Plainfield, NJ).

3.8 Electrophoretic mobility shift assays

3.8.1 Preparation of ^{32}P-labelled oligonucleotide probes

Sequences of HSE oligos used in experiments are listed in Table 3.3. An oligonucleotide sequence containing a CCAAT box (CBF) was used for competition assays to determine HSF specificity. Single stranded DNA oligos were purchased from Sigma-Aldrich Corporation (St. Louis, MO) and were annealed in house to create double stranded DNA with a five base pair overhang at both the 5' and 3' ends. The annealing reaction was performed by combining 1 nmol each of the sense and antisense oligos with annealing buffer (10 mM Tris-HCl pH 7.5, 0.1 M NaCl, 1 mM EDTA) in a 100 μL reaction. The annealing mixture was allowed to incubate at 92°C for 10 minutes and then allowed to slowly anneal at room temperature. The resulting double stranded (ds) DNA oligos were stored at -20°C.

Twenty-five pmol of the dsDNA sequence (above) was radio-labelled using a reverse transcriptase fill-in reaction. Five units of mouse leukemia virus (MmluV) reverse transcriptase (Fermentas), 2 μg BSA, 2 mM each dATP, dGTP, and dTTP and 20 μCi α^{32}P-dCTP (PerkinElmer Inc.) in a final volume of 20 μL were incubated at 37°C for 30 minutes. Reactions were then diluted to 100 μL with distilled water, extracted once with phenol/chloroform (1:1) and precipitated with 1/3 volume of 7.5 M ammonium acetate and 2.5× volume of 100% ethanol at -20°C overnight. The labeled dsDNA oligo was pelleted by centrifugation at maximum speed for 10 minutes at 4°C. The pellet was then allowed to air dry and finally resuspended in 20 μL of DEPC-water.

Table 3.3. Sequence of HSE oligonucleotides used for EMSA. Nucleotides altered from the perfect inverted 5'-nGAAn-3' sequence are underlined. Over-hanging nucleotides used for end-fill labeling are indicated by lower case letters.

Oligonucleotide Name	Nucleotide Sequence
HSE	5' ctaggGAA ATG GAA GC TTC GG GAA AC TTC GG GTCGG 3'
HSEsn	5' ctaggGAA ATG GAA GC CTC GG GAA AC TTC GG GTCGG 3'
HSEmod1	5' ctaggGAA ATG GAG GC GTC GG GAG AC GTC GG GTCGG 3'
HSEmod2	5' ctaggGAA ATG GCA GC TCC GG GCA AC TCC GG GTCGG 3'
HSEmod3	5' ctaggGAA ATG GTC GC AGC GG GTC AC AGC GG GTCGG 3'
HSEmod4	5' ctaggGAA ATG CAA GC TTA GG CAA AC TTA GG GTCGG 3'
CBF	5' ctaggGAG CCA ATC ACC GAG CTCG 3'

3.8.2 Sample preparation and electrophoresis

DNA binding reactions contained 750 ng of purified HSF in PBS/250 mM imidazole combined with 100 cps labeled HSE or CBF dsDNA probe (approximately 1 pmol of labeled DNA), 0.5 μg poly deoxy-inosine-deoxy-cytosine (dI-dC), 25 mM Tris-HCl pH 7.8, 1 mM EDTA, 1 mM DTT, and 5% (v/v) glycerol in a final volume of 25 μL. Reactions were incubated for 20 minutes at room temperature, and then mixed with 2 μL of 5× loading dye (0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol, 50% (v/v) glycerol). The reactions and dye were immediately loaded onto a 4% (w/v) non-denaturing polyacrylamide gel (44:0.8 acrylamide to bis-acrymamide) and electrophoresed in 1× Tris/glycine buffer (25 mM Tris-HCl, 192 mM glycine) at 4°C for 2 hours at 175 volts. The gel was dried on filter paper at 80^{o}C for 50 minutes and then exposed to HyBlot CL™ autoradiography film (Denville Scientific Inc., South Plainfield, NJ) at -80°C overnight with an intensifying screen.

3.9 Whole mount *in situ* hybridization

3.9.1 Generating DIG-labeled probes

3.9.1.1 *In vitro* transcription

In situ hybridization experiments were performed using antisense RNA probes that hybridize to specific mRNA sequences. Probes were prepared by *in vitro* transcription and were labeled by incorporating digoxigenin (DIG) labeled UTP nucleotides (Roche Diagnostics Corporation, Indianapolis, IN) into the antisense RNA probe during the *in vitro* transcription procedure.

Linearized plasmids containing the desired DNA sequences were used as templates for the *in vitro* transcription procedure. Plasmids were isolated as per section 3.5.1 under ampicillin selection (20 µg/mL). Zebrafish *alpha-tropomyosin* in Bluescript KS+ plasmid (Ohara et al. 1989; accession # NM_131105) and *hsf4* in pCR®4-TOPO plasmid (refer to section 3.6.2) were linearized with *Eco*RI and *Not*I respectively as per the enzyme manufacturer's instructions (Fermentas). Linearized plasmids were extracted with phenol/chloroform (1:1) followed by a second chloroform extraction. DNA was precipitated in 100% EtOH containing 0.15 M NaAc overnight at -80°C. Samples were centrifuged at maximum speed in a microfuge (Sorvall) at 4°C for 10 minutes after which the ethanol was removed from the pellet by aspiration. The pellet was washed with ice cold 70% (v/v) EtOH, centrifuged again as above, the EtOH was again removed and the DNA pellet was allowed to air dry before dissolving it in 100 µL of DEPC treated triple distilled water. A sample of linearized plasmid was analysed by gel electrophoresis to determine the concentration of the plasmid sample (refer to section 3.5.3) and to confirm that the reaction resulted in complete digestion of the plasmid. Linearized plasmids were stored at -20^{o}C until use.

Alpha-tropomyosin and *hsf4* antisense RNA probes, 860 nt and 866 nt in length respectively, were synthesized using T3 RNA polymerase according to manufacturer's instructions (Fermentas). Each 20 µL reaction was prepared by adding 1 µg DNA template, T3 RNA polymerase buffer, 40 units of T3 RNA polymerase, DIG labeled nucleotide mix (1mM ATP, CTP, GTP, plus 0.65 mM UTP and 0.35 mM DIG-UTP) to DEPC treated triple distilled H_2O. The reaction was thoroughly mixed and incubated for 2 hours at 37°C. Following the incubation, 2 units of RNase-free DNase (Fermentas) were added to the reaction and incubated at 37°C for 2 hours after which the reaction was stopped by adding 0.02 M EDTA (pH 8.0). The *in*

vitro transcribed RNA was precipitated by adding 2.5 µL of 4 M LiCl, and 75 µL of 100% EtOH to the reaction and incubating overnight at -80°C. RNA was then pelleted by centrifuging at 12,000 ×g for 25 minutes at 4°C in a microfuge (Sorvall), followed by a wash in ice cold 70% (v/v) EtOH and finally centrifuged again at 4°C. Ethanol was removed and the RNA pellet was allowed to air dry at room temperature. Finally, RNA was dissolved in 50 µL triple distilled DEPC water and stored at -80^{o}C until use.

Each sample was analysed using a denaturing agarose gel electrophoresis to determine if the *in vitro* transcription procedure was successful (section 3.5.7). The amount of RNA probe was quantified as described in section 3.5.7.

3.9.1.2 One-sided oligonucleotide PCR

A one-sided PCR technique was developed in collaboration with Dr. Bill Kulyk (Department of Anatomy and Cell Biology, University of Saskatchewan) to generate antisense single stranded (ss) DNA probes for use for *in situ* hybridization experiments as a more stable alternative to antisense RNA probes. This method was used to generate ssDNA probes to detect zebrafish *alpha-tropomyosin* and *hsf4* mRNA (section 3.9.1.1 and Table 3.1 for accession numbers).

Existing methods (Emanuel 1991; Konat 1996; Condon 1999) were modified to use a commercially synthesized oligonucleotide as a PCR template for the purpose of generating DIG labeled, ssDNA probes. The oligonucleotides templates were designed to a unique sequence of their respective target mRNA as follows:
alpha-tropomyosin (100 bp) 5'-TTGGACAGAGCAGAGCAGGCGAGACCGACAAGAAA GCAGCTGAGGAGAGGAGCAAGCAGCTCGAGGATGACCTGGTCGCCCTGCAGAAGA AGCTGAA-3' and *hsf4* (90 bp) 5'-AGACGGAGAGAGGAACAGCTTCAGCAGCGCACAT CCTAAGAGAACATCAATTTGGAAATGGACTAAAATACCACAGGATTAAGGAGACTC-3'. Primer sequences used to amplify the above templates were for *alpha-tropomyosin*, forward 5'-TTGGACAGAG CAGAGCAGGC-3' and reverse 5'-TTCAGCTTCT TCTGCAGGGC-3'; and for *hsf4* were forward 5'-AGACGGAGAGAGGAACAGCT-3' and
reverse 5'-CCACAGGATTAAGGAGACTC-3'.

One-sided PCR reactions were prepared by adding 500 ng of the appropriate oligonucleotide template, reverse primer at a final concentration of 1 µM, 5 µL of 10× DIG

labeled nucleotide mix (10 mM dATP, dCTP, dGTP, plus 5 mM dTTP and 5 mM DIG-dTTP (Roche Diagnostics Corporation)) and 5 units of *Taq* polymerase (Fementas) to 1× PCR buffer containing 2 mM $MgCl_2$. This mix was designed to incorporate the DIG label at a rate of 50% of the dTTP nucleotides. Control reactions using a 10× nucleotide mix with only unlabeled dNTPs; a reaction without a DNA template; one without *Taq* polymerase; and a positive control using a symmetric reaction that included both forward and reverse primers were also performed. Reactions were run under the following conditions: initial denaturation at 94°C for 2 minutes; followed by 40 cycles of 94°C for 1 minute, 55°C for 1 minute, 72°C for 2 minutes, followed by a final elongation of 72°C for 10 minutes. The labeled ssDNA probes were stored at -20°C until use.

A sample of each reaction was analysed by gel electrophoresis using a commercially prepared DNA ladder (as per section 3.5.5). The expected size of the PCR product was confirmed by comparing the relative mobility of the unlabelled control reaction to a commercial ladder. The DIG labeled reaction migrated slower on the gel than the unlabelled product due to the incorporation of the DIG residues (Roche Diagnostic GmbH 2008).

3.9.2 Embryo preparation and assay

Embryos were manually dechorionated as described in section 3.1. Between five and twenty embryos or larvae were collected in 1.5 mL microfuge tubes and fixed in 4% PFA (paraformaldehyde dissolved in PBS) for 2 hours at room temperature or at 4°C overnight. Embryos were dehydrated for storage prior to *in situ* hybridization through a series of graded (25%, 50%, 75%) methanol (MeOH) in PBST (PBS containing 0.01% Tween-20) solutions, and stored at -20°C in a final solution of 100% MeOH. Immediately prior to *in situ* hybridization, these samples were rehydrated with MeOH/PBST using concentrations in the reverse order of that described above.

The *in situ* hybridization protocol was performed as follows. Staged embryos were permeabilized using 10 µg/mL proteinase K (Boehringer-Manheim, Roche Diagnostics Corporation) in PBST at room temperature for the following lengths of time: 24 hpf or less are not exposed to proteinase K, 30-50 hpf incubated for 10 minutes; 60-72 hpf incubated for 15 minutes. Samples were then washed 2× 5 minutes in PBST, and post-fixed in 4% PFA in PBS at room temperature for 30 minutes. Following the fixing, samples were incubated in pre-

hybridization buffer (50% formamide in 5× SSC (750 mM NaCl, 75 mM trisodium citrate, pH 7.0), 5 mM EDTA, 0.1% v/v Tween-20, 50 µg/mL heparin, and 100 µg/mL tRNA in triple distilled water) at 65°C for RNA probes and at 50°C for ssDNA probes for 4 to 6 hours. Following pre-hybridization, heat denatured probe (90 sec at 85°C followed by ice for 2 minutes) was added to fresh pre-hybridization buffer. RNA probes were added at a concentration of 100 ng probe per 1 mL of pre-hybridization buffer. Single stranded DNA probes were added at a volume of 5 µL (from 50 µL reaction) to 1 mL of pre-hybridization buffer.

Following an overnight hybridization at 65°C (RNA)/ 50°C (ssDNA), embryos were washed several times to prevent non-specific binding of the probe. Washes were performed at 65°C (RNA)/50°C (ssDNA) for 10 minutes each in the following order: 75% hybridization solution /25% 2× SSC; 50% hybridization solution /50% 2× SSC; 25% hybridization solution /75% 2× SSC; 100% 2× SSC (all v/v) respectively. This series of washes was followed by two, 30 minutes washes in 0.2× SSC. Samples were then washed at room temperature for 5 minutes in each of the following: 75% 0.2× SSC/25% PBST; 50% 0.2× SSC/50% PBST; 25% 0.2× SSC/75% PBST; 100% PBST (all v/v) respectively.

Embryos were incubated in a blocking solution (2% (w/v) fetal calf serum and 2 mg/mL BSA in PBST) with mild agitation for 4 hours. Immunodetection of either RNA or ssDNA DIG-labeled probes was performed by incubating the samples with an anti-DIG antibody (Roche Diagnostics Corporation) diluted 1:5,000 in blocking solution for 3-4 hours at room temperature and mild agitation. Excess antibody was removed by washing the embryos 3× 15 minutes in PBST at 4°C, followed by an overnight wash in PBST at 4°C all with mild agitation. Embryos were then washed once in PBST for 15 minutes followed by one 5 minute wash in triple distilled water and 2× 5 minute washes in staining buffer (100 mM Tris-HCl pH 9.5, 0.1 M NaCl, 0.1% (v/v) Tween-20) all at room temperature. Staining was performed by incubating the embryos in staining solution (250 µg/mL BCIP, 250 µg/mL NBT and 4 µM levamisol diluted in staining buffer) at room temperature with mild agitation until colour was visualized within the embryos (usually 1-2 hours). This was followed by two 5 minute washes in PBST and a post-fix in 4% PFA in PBS (v/v), all at room temperature. Following post-fixation, embryos were washed two times in PBST at room temperature and stored in PBS with 0.02% sodium azide at 4°C.

The *in situ* stained whole mount embryos/larvae were prepared for visualization under the microscope using the following procedure. Whole mount samples were dehydrated by a series of graded washes in PBST:MeOH as described above (25% MeOH to 100% v/v MeOH). The dehydrated samples were placed in a clearing solution containing 2:1 benzyl alcohol: benzyl benzoate immediately prior to imaging. Samples were placed in a depression slide and viewed using a Nikon Eclipse E600 photomicroscope. Results were observed and recorded using light microscopy described in section 3.3.

3.10 Real-time quantitative reverse transcription PCR

The level of mRNA transcripts for *hsf1b*, *hsf1a+b*, *hsf2* and both *hsf4a* and *hsf4b* isoforms were analysed by quantitative real time reverse transcription polymerase chain reaction (qPCR) using the primers (Sigma-Aldrich) listed in Table 3.4. *Elongation factor alpha* (*Elfα*) and *beta-actin* (*βactin*) were used as reference genes to normalize data. Amplification and quantification of reference genes was always performed in conjunction with the analysis of target genes. The extraction and quantification of RNA was performed as per sections 3.5.6 and 3.5.7, cDNA prepared as per section 3.5.8 and primer sets were optimized as per section 3.10.1 to comply with qPCR quality control measures outlined in the MIQE guidelines (Bustin 2000; Bustin et al. 2009).

Table 3.4. Summary of primers used for qPCR. Reaction efficiencies (E) for each primer set used in qPCR reactions described in section 3.10, were calculated as described in section 3.10.1. Reaction efficiencies (E) were considered to be within acceptable parameters if the percent reaction efficiency was between 80-120% which corresponds to E between 1.8 and 2.2. Therefore primers with calculated E values between 1.8 and 2.2 were considered acceptable for use in qPCR reactions. Note: a reaction efficiency for the *hsf4* splice-MO was not determined (ND).

Gene Name (Accession #)	**Forward Primer**	**Reverse Primer**	**E**
Elfα (NM_131263.1)	5' CAAGGAAGTCAGCGCATACA 3'	5' TCTTCCATCCCTTGAACCAG 3'	2.2 Ssofast 2.0 SYBR
βactin (NM_131031)	5' CGAGCAGGAGATGGGAACC 3'	5' CAACGGAAACGCTCATTGC 3'	2.1 Ssofast 2.0 SYBR
hsp70 (NM_131397.2)	5' ATCAACGAGCCCACGGCTGC 3'	5' ACATGCGGTTCGCAGCCTCC 3'	2.0
hsf1a+b (NM_131600.1)	5' TGTGGACACGCCCCTTTCGC 3'	5' GCAGGCGACGCTCAAGCACT 3'	2.2
hsf1b (NM_131600.1)	5' GGCTTCTCCACCTCATCTCT 3'	5' TGTCGATGGACTCCAGATG 3'	2.0
hsf2 (NM_131867)	5' GGATGAGTCCCTGGAGATGA 3'	5' GATAGAAGAGCGTGGCTTCG 3'	2.1
hsf4 (both transcripts)	5' GGTCCAGGTGTTGCGGAGCC 3'	5' AGGGGTGGAGCAGCCGTCAT 3'	2.0
hsf4a (NM_001013317)	5' GCCAAGGAAGTTCTGCCCAA 3'	5' GCTAAAAGTGGTCTCGCCCC 3'	2.1
hsf4b (Ensemble HSF4-201)	5' GGTCCAGGTGTTGCGGAGCC 3'	5' AGCCAGCGGTGGGGTACAGC 3'	1.9
hsf4 splice-MO (both transcripts)	5' GGTGTTGCGGAGCCAACAGG 3'	5' GTGACACCACTTCCCTCCACA 3'	ND
fgf1 (NM_200760.1)	5' GCCGTAGGTAAGCAGCTCGCT 3'	5' GTCGCGCGGGCTGGACTTTA 3'	2.2

For individual PCR reactions, cDNA from each sample was diluted five fold in water and 2 µL of the diluted cDNA was added to each reaction. Reactions to detect *hsf4a* and *hsf4b* mRNA were performed by adding cDNA to a 20 µL reaction mixture containing 10 µL of 2× Sso-Fast™ EvaGreen® Supermix (Bio-Rad Laboratories) and 300 nM final concentration of each primer. Reactions for all other genes of interest were performed by adding cDNA to a 25 µL reaction mixture containing 12.5 µL of 2× Maxima SYBR green/ROX qPCR Master Mix (Fermentas) and 300 nM of each primer. All PCR reactions were performed using triplicate technical samples and a single non-template control (NTC) for each biological sample to be tested and were performed in a MiniOpticon thermocycler (Bio-Rad Laboratories). Reactions were performed using the following conditions: denaturing reaction at 95^{o}C for 10 minutes followed by 50 cycles of (95°C for 15 seconds, 62°C for 30 seconds, 72°C for 45 seconds). A melting curve of the amplified DNA product was also generated for all samples after the completion of the PCR reactions. Critical threshold values (C_T) and melting curve analysis was performed using the CFX Manager software with default settings (Bio-Rad Laboratories). Samples with C_T values that were within 1 C_T of the value of the corresponding non-template control (NTC) were determined to be below the limit of detection and assigned a value of 0 for the purposes of graphing results. Reactions that resulted in products with multiple melting peaks or melting peaks that did not cross the programs default threshold were also classified as below the limit of detection.

3.10.1 Primer optimization

All primers designed for qPCR analysis (Table 3.4 and A2) were optimized for reaction conditions (described below) using cDNA generated from zebrafish embryos at 80 hpf. An optimal range of annealing temperatures was determined for each primer set by running PCR reactions using a temperature gradient from 55-70°C. Optimal primer concentrations were determined from a series of experiments using 100 µM-1 mM primer concentrations in order to minimize the occurrence of primer dimers. The reaction efficiency (E) of each primer set was calculated from a standard curve using the equation $E=10^{[-1/slope]}$ and the percentage efficiency was calculated as %E = (E-1)*100 (Rasmussen 2001). Standard curves for each primer set were generated using a three fold dilution series of cDNA with one exception being the primers for *hsf4a* which used a sequential dilution series from 1 to 7 fold dilutions of cDNA. It is important

to note that while calculating efficiency from serial dilutions is highly reproducible, this method of calculation frequently results in efficiencies greater than two which is theoretically impossible suggesting a limitation in the accuracy of this calculation method (Pfaffl 2006). Therefore, for the purposes of this study, primers that had a calculated efficiency value between 80-120%, corresponding to an E between 1.8 and 2.2, were considered to have acceptable reaction efficiency and an optimal efficiency of 2 was assumed in subsequent calculations. Calculated efficiencies are shown in table 3.4.

3.10.2 Analysis of qPCR data

All qPCR data obtained from the CFX Manager software were analysed using the Relative Expresstion Software Tool (REST) 2009 (Pfaffl et al. 2002). The C_T values for the βactin reference gene were used to normalize C_T values obtained for the genes of interest. The normalized values were then compared to a control treatment or tissue and expressed as a ratio or fold increase relative to that control. The REST software calculated the resulting relative expression ratios using the following equation (Livak and Schmittgen 2001; Pfaffl et al. 2002):

$$\text{Ratio} = [(E_{target})^{\Delta Ct \text{ target (mean control – mean sample)}}] \div [(E_{reference})^{\Delta Ct \text{ reference (mean control – mean sample)}}]$$

The calculated relative expression ratios were further analysed for statistical significance using the REST software. This software algorithm uses a Pair-wise Fixed Reallocation Randomization Test (Pfaffl et al. 2002) which is a non-parametric test designed to analyse data in a robust way without making assumptions about the distribution of the data. The randomization calculations are subjected to a hypothesis test to determine the probability that the samples are significantly different at $p \leq 0.05$ (Pfaffl et al. 2002). For the purpose of this study, the REST software was set to perform 2000 randomizations of each data set. The REST program also employs bootstrapping techniques to determine 95% confidence intervals for data sets. For the purposes of this study, calculated fold difference over control values with a $p \leq 0.05$ were determined to have a significant difference in mRNA levels and fold difference values with a $p \geq 0.05$ but ≤ 0.15 were considered to be showing a trend to an increase or decrease in mRNA levels.

3.11 Sequence identification of *hsf4* transcripts from zebrafish embryos

Polymerase chain reaction was used to amplify each of *hsf4* transcripts a and b from zebrafish embryos at 80 hpf for the purpose of identifying the sequence of the amplified product. Ribonucleic acid was extracted from zebrafish embryos as per section 3.5.6 and 3.5.7 and cDNA was generated from each of these samples as per section 3.5.8. Amplification of the *hsf4a* transcript was performed with Sso-Fast™ EvaGreen® Supermix and the *hsf4b* transcript was amplified using with 2x Maxima SYBR green/ROX as described in section 3.10 using the primers listed in Table 3.4. An additional reaction was performed to amplify *hsf4b* without the use of the SYBR green dye and instead used an in house PCR mix consisting of 1× PCR reaction buffer, 3 mM $MgCl_2$, 300 nM of each primer, 200 µM dNTPs and 5 U *Taq* DNA polymerase (primers from Sigma-Aldrich; remaining reagents from Invitrogen). Polymerase chain reactions were performed as follows: denaturing reaction at 95^{o}C for 10 minutes followed by 50 cycles of (95°C for 15 seconds, 60°C for 30 seconds, 72°C for 45 seconds) and a final extension time of 5 minutes at 72°C was added to ensure adequate adenosine overhang residues were present to facilitate TOPO cloning.

All PCR products were sequenced as per section 3.5.5 and identity of the PCR products was confirmed by comparing the obtained sequences to those deposited in GenBank and Ensembl using BLAST (section 3.5.5).

3.12 Morpholino knockdown of *hsf1* and *hsf4* expression

Morpholino modified antisense oligonucleotides (MO) are synthetic DNA sequences modified with a morpholino ring and a neutral phosphate backbone (Ekker and Larson, 2001). They have a high affinity for RNA and low cellular toxicity which makes them ideal for use as an *in vivo* gene targeting tool that cause the knockdown of expression of a particular gene (Summerton 1999). Morpholinos were purchased from Gene Tools, LLC (Philomath, OR) as a 25 base oligo complementary to a target mRNA. Knockdown of zebrafish *hsf* expression was performed using MO tagged with either 3'-carboxyfluorescein or 3'-lissamine. The 3'-carboxyfluorescein is a green emitting fluorescent tag (excitation peak at 501 nm; emission peak at 524 nm) and adds a mass of 490 Da to the MO while the lissamine tag, also known as sulforhodamine B, is a red-emitting fluorescent tag (excitation peak at 575 nm; emission peak at 593 nm) and adds a mass of 785 Da to the MO. Morpholino sequences used to knockdown *hsf1*

gene expression contained a lissamine tag and were designed to do so by binding to a sequence spanning the start codon of the mRNA, thereby preventing the binding of the ribosomal subunit and inhibiting translation (Summerton 1999). Initial experiments designed to look at the phenotypic effect of *hsf4* knockdown used MO sequences that contained a carboxyfluorescein tag and were also designed to bind DNA sequences spanning the start codon. Morpholino sequences used to knockdown *hsf4* gene expression for analysis of effect on potential downstream gene targets, were designed to an *hsf4* sequence that spanned the intron/exon junction at the 5' end of the 5th exon in *hsf4* pre-mRNA sequences. This type of MO is referred to as a splice-MO and results in the exon targeted by the MO being co-spliced out with the adjoining intron during pre-mRNA processing (Draper et al. 2001). The 5th exon that was targeted by the *hsf4* splice-MO is the coding region for approximately one third of the transactivation domain of the HSF4 protein. Embryos injected with a splice-MO that targets and removes a portion of the coding region for this essential protein domain should result in the synthesis of proteins that are not active unlike embryos injected with MO sequences that block translation which results in a decrease in synthesis of the target protein.

Morpholinos were designed to be complementary to a specific *hsf1* or *hsf4* mRNA sequence based on mRNA sequence obtained from Genbank. The following sequences were used as MO targets and the overlap with the ATG start codon is underlined where applicable: *hsf4* MO (5'-TATAGAGCCTGGGTTCTCCTGCATG-3') and *hsf4* splice-MO (5'-TCTGCCTGCAAGATAGATAGATAGA-3') - Accession # NM001013317; *hsf1* MO (5'-CACGGAGAGTTTAGTGATGATTTCT-3') - Accession # NM 131600. For both genes, the morpholinos sequences chosen were able to bind both known transcripts of *hsf1* and *hsf4*. Morpholinos with a five base pair mismatch were injected as a control for non-specific phenotypic effects (mismatched base pairs are indicated by lower case letters);

hsf4 5pbmm MO (5'-TATtGAGgCTGGcTTCaCCTGgATG-3');

hsf4 5bpmm splice-MO (5'-TCTcCCTGgAAcATAcATAcATAGA-3');

hsf1 5bpmm-MO (5'-CAgGGAcAGTaTAGaGATcATTTCT-3'). Morpholino mismatch sequences were designed to have no significant sequence similarity to any part of the zebrafish genome while still being as similar as possible to the original MO sequence (differing from the original MO by only five to six base pairs).

Morpholino stock solutions for each of the *hsf1* MO, *hsf1* 5pbmm-MO, *hsf4* MO and *hsf4* 5bpmm-MO were prepared in triple distilled water to a final concentration of 8.33 µg/µL, aliquoted and stored at -20^{o}C until use. Stock MO were further diluted in Danieau solution (58 mM NaCl, 0.7 mM KCl, 0.4 mM $MgSO_4$, 0.6 mM $Ca(NO_3).H_2O$, 5 mM HEPES (pH 7.6); Nasevicius and Ekker 2000) to a concentration of 3.46 µg/µL immediately prior to embryo injection. Needles used for microinjection (glass tubes 100 mm in length and 0.02 mm in diameter sourced from Drummond Scientific Co., (Broomall, PA) were made in house using a model 700C vertical pipette puller (David Kopf Instruments, Tujunga, CA). Needles were calibrated to deliver a constant volume of 3.14 nL per injection using a Narashige IM-300 microinjector (Drummond Science Co). This was achieved by marking off 1 mm increments on each needle and adjusting the injection psi so that 10 injections dropped the meniscus of the injection fluid by 1 mm. The microinjector used human grade compressed nitrogen (Praxair, Saskatoon, SK) to pressurize the injection system and needle pressure was maintained at 80-85 psi for the duration of the injections.

Newly fertilized embryos were lined up along a glass slide placed at the bottom of a petri dish. An absorbent tissue was used to remove most of the water surrounding the embryos to assure that the embryos did not move during the injection procedure. The petri dish was placed under a dissecting microscope (Leica Microsystems Inc., Concord, ON) and illuminated with an Intralux 5000-1 fiber optic light source (Volpi USA, Auburn, NY). A needle was inserted into the microinjector and filled with MO solution. Embryos were injected by passing the needle through the chorion and into the yolk of the embryos. Needles were washed a minimum of five times with Danieau solution between injections of different solutions. All embryos were injected prior to the 32 cell stage. Following injection, embryos were maintained in 25 mL petri dishes as described in section 3.1. Embryo mortality was monitored daily and non-viable embryos were removed. Treatment groups consisted of embryos injected with MO, 5bpmm-MO or Danieau solution both of which were compared to an uninjected treatment group for the analysis described below ($n \geq 80$ embryos for all treatments).

Embryos analysed for phenotypic effects of *hsf4* knockdown were screened for the presence and distribution of fluorescein at 48 hpf using a Nikon 41002C HQ:G filter as described in section 3.3. The diameter of the eye as well as the length of the body was measured in embryos with intense fluorescein distribution throughout the embryo at 48 hours post

fertilization (hpf). The effect of the MO on lens phenotype was determined by calculating the ratio of eye and lens diameter and eye surface area to total body length similar to that previously described (Evans et al. 2007).

Embryos to be analysed by qPCR were screened for the presence and distribution of lissamine at 48 hpf using using a Nikon 41002C HQ:R filter as described in section 3.3. Only embryos observed to have intense lissamine distribution throughout the embryo were considered to be successful injections and these embryos were retained for further analysis. Embryos were screened again prior to 60 hpf to remove embryos that were dead or had severe developmental abnormalities (example: stubby tail, crooked spine or missing fins or eyes). The remaining embryos were screened for a small eye phenotype using bright field microscopy (section 3.3). Finally embryos were collected at 60 hpf and frozen for RNA extraction (sections 3.5.6-3.5.7) and subsequent qPCR analysis ($n \geq 30$ embryos for all treatments in each biological replicate).

Quantitative PCR was performed on the MO-injected and uninjected embryos to analyse mRNA levels for *fgf1, hsp70*, *hsf4* splice-MO and *hsf1a+b* using the primers outlined in Table 3.4. Reactions were prepared using Sso-Fast™ EvaGreen® Supermix (Bio-Rad Laboratories) and PCR analysis was performed exactly as described for *hsf4a* and *hsf4b* analysis in section 3.10.

4.0 RESULTS

4.1 *In silico* analysis of zebrafish heat shock factors

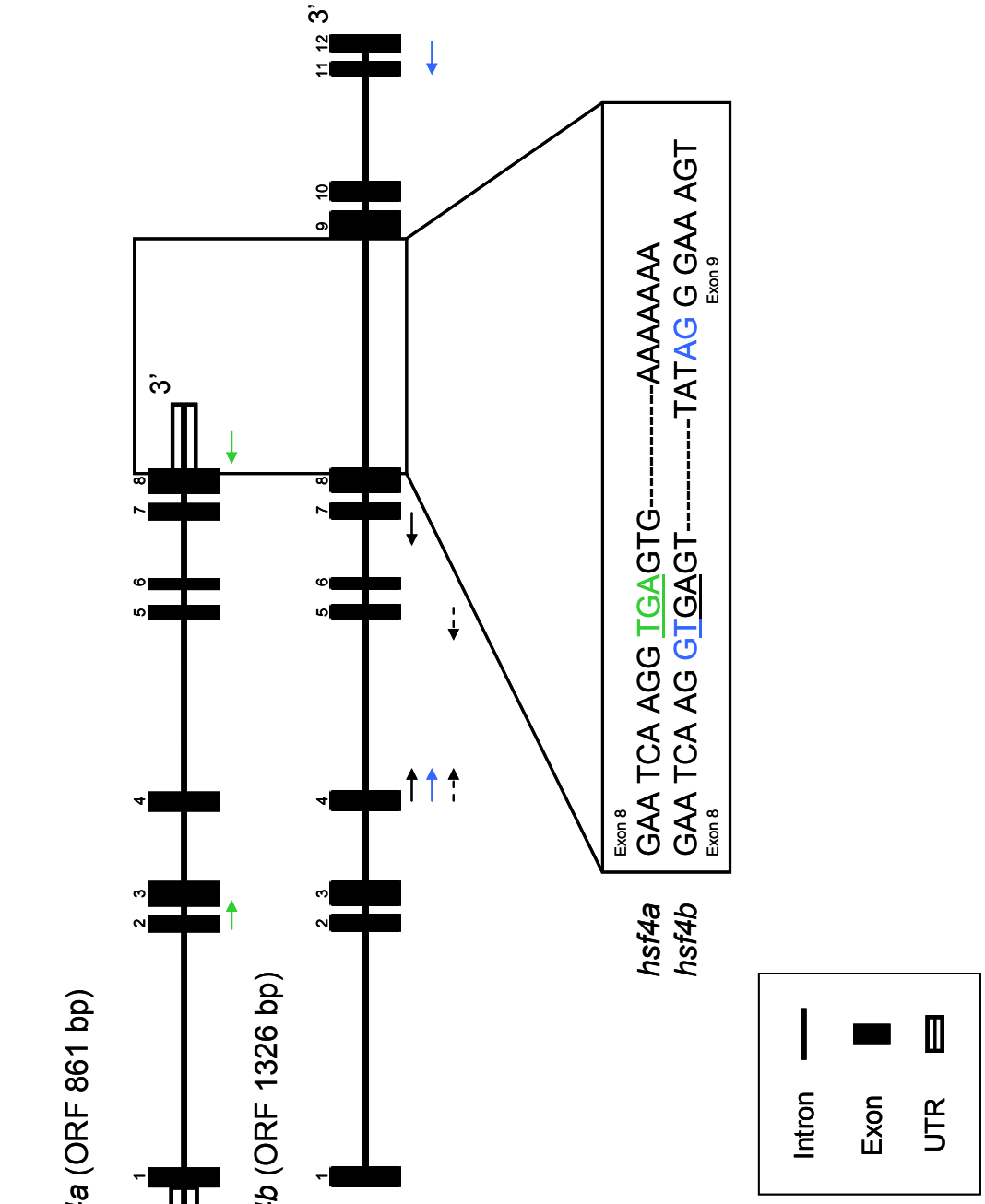
4.1.1 Nucleic acid sequence analysis of two zebrafish *hsf4* transcripts

A novel zebrafish HSF (*hsf4*) was previously identified in an EST screen of an embryonic cDNA library derived from zebrafish embryos spanning stages from fertilization to 72 hpf (Strausberg et al. 2002). The identified clone consists of a 1650 bp full length cDNA sequence, (Genbank accession # NM_001013317), corresponds to Mammalian Genome Collection (MGC)/zgc:113344, and is referenced in the ZFIN library as ZDB-GENE-050306-18 (Sprague et al. 2006). Sequence analysis of the cloned cDNA reveals a high degree of identity to previously characterized zebrafish *hsf1* and *hsf2* nucleotide sequences (Råbergh et al. 2000; Yeh et al. 2006), and to mammalian *hsf4*. The identified mRNA sequence is encoded by a single *hsf4* gene that has been mapped to zebrafish chromosome 18, with the genomic sequence spanning 7103 bp (Genbank accession # BX088710.16) and consisting of 8 exons and 7 introns (Fig. 4.1A). My analysis of the genomic sequence identified a TATA box in the 5' end of the gene, 214 bp upstream of the putative ATG translation initiation site. There were also three additional in frame ATG sites identified in the 5' region of the gene located at 57 bp, 165 bp and 205 bp upstream of this site. My analysis of the 3' untranslated region identified two ATTAAA sequences that may act as a poly adenylation signals. The sequences are located 154 bp and 456 bp downstream of the TGA stop codon. The second ATTAAA sequence is also 163 bp upstream of the previously identified poly A tail.

The open reading frame of a second, longer putative *hsf4* transcript has also been predicted in the Ensembl database, and originates from the same gene on chromosome 18 that encodes for the shorter transcript (Ensembl Zv9, release 71; transcript ID ENSDART00000130961). My analysis determined that nucleic acid sequence in the first 8 exons of the longer transcript is identical to the shorter transcript but also includes an additional 4 exons derived from genomic sequence that spans an additional 3266 bp downstream from the end of the shorter transcript (Fig. 4.1A). The open reading frame of this transcript is predicted to contain 1326 bp from the total 12 exons. Since the longer sequence is predicted, no 5' or 3' UTR was identified in Ensembl. However, my analysis of the genomic sequence in the 3' region downstream of the 12th exon identified an ATTTAA poly adenlyation signal sequence 482 bp downstream of the TAA stop codon.

Figure 4.1 Sequence comparison of two splice variants of zebrafish *hsf4*.

(**A**) The genomic sequence of the zebrafish *hsf4* gene contains twelve exons and eleven introns. Exons are numbered sequentially in the figure from the 5' end to the 3' end. Two splice variants that are transcribed from this gene have been identified and are named *hsf4a* and *hsf4b*. The sequence of *hsf4a* (identified from an EST screen of zebrafish embryos) contains the entire sequence from exons 1 through 8 terminating at a TGA stop codon located at 859 bp from the ATG start site (depicted with green letters and underlined). Additionally, a full 5' and 3' UTR has been sequenced for this transcript. The sequence predicted for *hsf4b* also contains exons 1 through 8 plus additional sequence from exons 9 through 12. The *hsf4b* transcript contains a splice site which splits the AGG codon (depicted with blue letters) 858 bp from the ATG start codon allowing the spliceosome to bypass the TGA stop codon (underlined) and read through the entire 12 exons of the gene. No UTR regions have been sequenced for the *hsf4b* transcript. The arrows shown above represent primer binding sites (Table 3.4) which flank specific mRNA sequence used for qPCR analysis of the two *hsf4* transcripts. Arrows in green flank the region specific for *hsf4a*, blue arrows flank the region specific for *hsf4b*, in black flanks a region common in both transcripts and the dashed line flanks the region amplified for analysis of *hsf4* splice-MO knockdown experiments. (**B**) A sequence alignment of *hsf4a* transcript from GenBank, the predicted *hsf4b* from Ensembl and the portion of *hsf4b* sequence from an in house PCR product. Residues that are identical between the three sequences are shaded in black. The red square highlights the position of the TGA stop codon that occurs in the *hsf4a* transcript which is absent from the *hsf4b* transcript.


A
hsf4a (ORF 861 bp)
hsf4b (ORF 1326 bp)
5'
3'
1
2
3
4
5
6
7
8
9
10
11
12
hsf4a
hsf4b
Exon 8
GAA TCA AGG TGAGTG----------AAAAAAA
GAA TCA AG GTGAGT----------TATAG G GAA AGT
Exon 8
Exon 9
Intron
Exon
UTR

B

```
hsf4a_ORF          1 ATGCAGGAGAACCCAGGCTCTATAGGTGTGGATGGGGGATATGCTAGCAATGTGCCCGCT
hsf4b_predicted    1 ATGCAGGAGAACCCAGGCTCTATAGGTGTGGATGGGGGATATGCTAGCAATGTGCCCGCT
hsf4b_sequence     1 ------------------------------------------------------------

hsf4a_ORF         61 TTCCTCACTAAACTCTGGACTCTGGTCGAGGACCCGGAGACTAATCATCTGATCTGCTGG
hsf4b_predicted   61 TTCCTCACTAAACTCTGGACTCTGGTCGAGGACCCGGAGACTAATCATCTGATCTGCTGG
hsf4b_sequence     1 ------------------------------------------------------------

hsf4a_ORF        121 AGTGCTACCGGGACCAGCTTTCATGTTTTCGACCAAGGCAGATTTGCCAAGGAAGTTCTG
hsf4b_predicted  121 AGTGCTACCGGGACCAGCTTTCATGTTTTCGACCAAGGCAGATTTGCCAAGGAAGTTCTG
hsf4b_sequence     1 ------------------------------------------------------------

hsf4a_ORF        181 CCCAAATACTTCAAACACAACAACATGGCGAGTTTCGTACGCCAGCTCAACATGTACGGC
hsf4b_predicted  181 CCCAAATACTTCAAACACAACAACATGGCGAGTTTCGTACGCCAGCTCAACATGTACGGC
hsf4b_sequence     1 ------------------------------------------------------------

hsf4a_ORF        241 TTCAGAAAAGTGGTGAATATTGAGCAGAGTGGGCTGGTGAAGCCAGAGCGAGACGACACT
hsf4b_predicted  241 TTCAGAAAAGTGGTGAATATTGAGCAGAGTGGGCTGGTGAAGCCAGAGCGAGACGACACT
hsf4b_sequence     1 ------------------------------------------------------------

hsf4a_ORF        301 GAGTTCCAGCACCTCTATTTTCTCCAGGGCCACGAGCATCTGCTGGAGCACATTAAACGC
hsf4b_predicted  301 GAGTTCCAGCACCTCTATTTTCTCCAGGGCCACGAGCATCTGCTGGAGCACATTAAACGC
hsf4b_sequence     1 ------------------------------------------------------------

hsf4a_ORF        361 AAAGTGTCAATAGTAAAAAGCGAGGAGACTAAAGTGCGACAGGAGGACTTGAGTAAGCTG
hsf4b_predicted  361 AAAGTGTCAATAGTAAAAAGCGAGGAGACTAAAGTGCGACAGGAGGACTTGAGTAAGCTG
hsf4b_sequence     1 ------------------------------------------------------------

hsf4a_ORF        421 TTGTATGAGGTCCAGGTGTTGCGGAGCCAACAGGAAAACATGGAGATGCAGATGCAGGAT
hsf4b_predicted  421 TTGTATGAGGTCCAGGTGTTGCGGAGCCAACAGGAAAACATGGAGATGCAGATGCAGGAT
hsf4b_sequence     1 --------GGTCCAGGTGTTGCGGAGCCAACAGGAAAACATGGAGATGCAGATGCAGGAT

hsf4a_ORF        481 ATGAAACAGCAGAATGATGTTCTGTGGAGGGAAGTGGTGTCACTGAGACAGAATCACACA
hsf4b_predicted  481 ATGAAACAGCAGAATGATGTTCTGTGGAGGGAAGTGGTGTCACTGAGACAGAATCACACA
hsf4b_sequence    53 ATGAAACAGCAGAATGATGTTCTGTGGAGGGAAGTGGTGTCACTGAGACAGAATCACACA

hsf4a_ORF        541 CAGCAGCAGAAAGTCATGAACAAGCTGATTCAGTTCCTGTTCAGCCAGATGCAATCCAAC
hsf4b_predicted  541 CAGCAGCAGAAAGTCATGAACAAGCTGATTCAGTTCCTGTTCAGCCAGATGCAATCCAAC
hsf4b_sequence   113 CAGCAGCAGAAAGTCATGAACAAGCTGATTCAGTTCCTGTTCAGCCAGATGCAATCCAAC

hsf4a_ORF        601 TCTCCAAGCACTGTGGGAATGAAGAGAAAGCTTCCTCTGATGCTGGATGACGGCTGCTCC
hsf4b_predicted  601 TCTCCAAGCACTGTGGGAATGAAGAGAAAGCTTCCTCTGATGCTGGATGACGGCTGCTCC
hsf4b_sequence   173 TCTCCAAGCACTGTGGGAATGAAGAGAAAGCTTCCTCTGATGCTGGATGACGGCTGCTCC

hsf4a_ORF        661 ACCCCTCCTGCCTCCAAATTCAGCCATAACCACTCCATGGAGTCTTTACAGGAATCATTC
hsf4b_predicted  661 ACCCCTCCTGCCTCCAAATTCAGCCATAACCACTCCATGGAGTCTTTACAGGAATCATTC
hsf4b_sequence   233 ACCCCTCCTGCCTCCAAATTCAGCCATAACCACTCCATGGAGTCTTTACAGGAATCATTC

hsf4a_ORF        721 TATATCCAATCGCCATCCACAGAAAGTGCTTCCTGCTCGACTAGCAGCGTGATGACAGGA
hsf4b_predicted  721 TATATCCAATCGCCATCCACAGAAAGTGCTTCCTGCTCGACTAGCAGCGTGATGACAGGA
hsf4b_sequence   293 TATATCCAATCGCCATCCACAGAAAGTGCTTCCTGCTCGACTAGCAGCGTGATGACAGGA
```

B

```
hsf4a_ORF          781 GGGCCCATCATCTCAGATGTCACTGAAATCCCACAGTCCAGCTCCATGGCTTTACAAATG
hsf4b_predicted    781 GGGCCCATCATCTCAGATGTCACTGAAATCCCACAGTCCAGCTCCATGGCTTTACAAATG
hsf4b_sequence     353 GGGCCCATCATCTCAGATGTCACTGAAATCCCACAGTCCAGCTCCATGGCTTTACAAATG

hsf4a_ORF          841 CAGGCAGAGGAATCAAGGTGA-------------------------------------
hsf4b_predicted    841 CAGGCAGAGGAATCAAGG---GAAAAGTGTCTGATGCTAATTAAGGAGGAGCCGGTGAGT
hsf4b_sequence     413 CAGGCAGAGGAATCAAGG---GAAAAGTGTCTGATGCTGATTAAGGAGGAGCCGGTGAGT

hsf4a_ORF              ------------------------------------------------------------
hsf4b_predicted    898 CCTGGGGTCCAGGGACGAGCAGAGGGTGTTCCACTCGGGTCTTGTGAAGTGTGTGCTGAA
hsf4b_sequence     470 CCTGGGGTCCAGGGACGTGCAGAGGGTGTTCCACTCGGGTCTTGTGAAGTGTGTGCTGAA

hsf4a_ORF              ------------------------------------------------------------
hsf4b_predicted    958 CCCCCTGTCCTCCCTGTTGCCATGGTGCAGTCCGTCCTGGAGGGCCGAGGATCTAATTTG
hsf4b_sequence     530 CCCCCTGTCCTCCCTGTTGCCATGGTGCAGTCCGTCCTGGAGGGCCGAGGATCTAATTTG

hsf4a_ORF              ------------------------------------------------------------
hsf4b_predicted   1018 GGAGAGAGAAGGGCCAAAAGACCCATGCTGGAGAGACCAGAGATTCCTGATGGCGTGGAG
hsf4b_sequence     590 GGAGAGAGAAGGGCCAAAAGACCCATGCTGGAGAGACCAGAGATTCCTGATGGCGTGGAG

hsf4a_ORF              ------------------------------------------------------------
hsf4b_predicted   1078 AATGTTGACATGAGTCTGGAGGATTTACAGCTGCTTCTGAGGAGTCACCAGCAGAGCATG
hsf4b_sequence     650 AATGTTGACATGAGTCTGGAGGATTTACAGCTGCTTCTGAGGAGTCACCAGCAGAGCATG

hsf4a_ORF              ------------------------------------------------------------
hsf4b_predicted   1138 GAAAATAATGCTGCAGCAATGGATCAATTTACATTTAGTCTGCCTTTGAATGAGTGGAAT
hsf4b_sequence     710 GAAAATAATGCTGCAGCAATGGATCCCTTTACATTTAGTCTGCCTTTGAATGAGTGGAAT

hsf4a_ORF              ------------------------------------------------------------
hsf4b_predicted   1198 TTCGCAGAAATGGACCCGAACCTGAAATCGGAACTGGCTAATGCCCTCATCCCTGCTGCT
hsf4b_sequence     770 TTCGCAGAAATGGACCCGAACCTGAAATCGGAACTGGCTAATGCCCTCATCCCTGCTGCT

hsf4a_ORF              ------------------------------------------------------------
hsf4b_predicted   1258 GTGTCCCAATACATGTTTCAGGGTCAGGAGGGAGAGCTGTACCCCACCGCTGGCTATGAA
hsf4b_sequence     830 GTGTCCCAATACATGTTTCAGGGTCAGGAGGGAGAGCTGTACCCCACCGCTGGCT-----

hsf4a_ORF              ---------
hsf4b_predicted   1318 GAGCAGTAA
hsf4b_sequence         ---------
```

I named these two transcripts in accordance with the ZFIN zebrafish gene nomenclature guidelines in which new genes are named sequentially and using similar identifiers as those used for human and mouse orthologs but without the use of capital letters (Mullins 1995). The shorter transcript was designated *hsf4a* and the longer transcript *hsf4b*.

I confirmed the presence of both the *hsf4a* and *hsf4b* transcripts in zebrafish embryos by sequencing the products from two or more independent reverse transcription PCR amplifications from zebrafish embryos at 80 hpf as described in section 3.11. Two sets of primers specific to *hs4fa* were designed. The first set was designed to amplify a 680 bp portion of the mRNA sequence which spanned the region from within the 3rd exon through the stop codon and into the 3' UTR. The second set was designed to amplify a 968 bp portion which spanned from within the 2nd exon through the stop codon and into the 3' UTR. A primer set specific to *hsf4b* was designed to amplify an 866 bp portion of the ORF which spanned from within the 4th exon to within the 12th exon. I performed a sequence alignment of the *hsf4b* sequencing results compared to the known *hsf4a* ORF and the predicted *hsf4b* ORF are shown in Fig. 4.1B. The sequence alignment indicates that the PCR products for both transcripts matched the previously identified and predicted sequences for their respective ORFs. The amplified sequence of *hsf4a* was identical to the sequence deposited in Genbank (NM_001013317) and sequence analysis confirmed the existence of a TGA stop codon. The sequence of *hsf4b* was identical to both that of *hsf4a* and the predicted sequence up to 858 bp position where it appears to lack the TGA stop codon (Fig. 4.1B). The remainder of the *hs4fb* sequence is identical to its predicted sequence with the exception of a thymine base sequenced at position 487 of the PCR product instead of an adenine in the predicted sequence as well as two cytosine bases sequenced at positions 735 and 736 of the PCR product instead of two adenine bases (Fig. 4.1B).

4.1.2 Amino acid sequence analysis of two isoforms of zebrafish HSF4

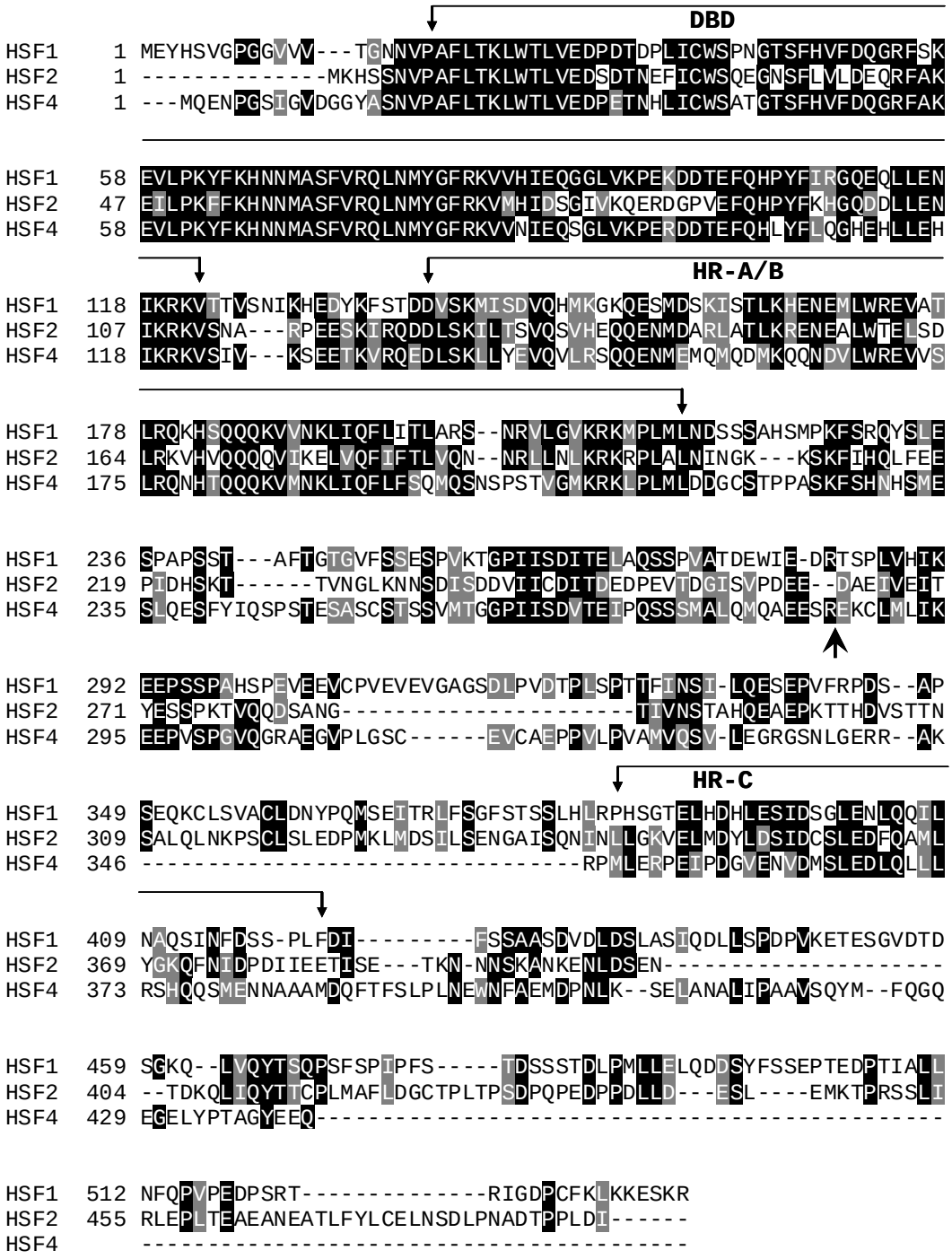
The zebrafish *hsf4* sequence is similar to other zebrafish and mammalian HSF sequences in that exons 1-3 encode for the DBD and exons 4-6 encode for the HR-A/B domain of the predicted protein (Pirkkala et al. 2001; Sprague et al. 2006). Both *hsf4* transcripts are identical in sequence up to the end of the 8th exon of the gene while the *hsf4b* transcript contains sequence from four additional downstream exons. Therefore, the DBD and HR-A/B domains predicted for both HSF4 isoforms are identical in amino acid sequence composition.

Sequence alignment (Fig. 4.2A and B) of the predicted amino acid sequences for both HSF4s demonstrates the high degree of amino acid sequence conservation between the zebrafish HSF4 proteins and zebrafish HSF1 and HSF2. My analysis of the predicted amino acid sequence of both HSF4 isoforms using SMART and Pfam protein analysis tools (UniProtKB) indicates that amino acids 19-123 in both HSF4 sequences comprise a region with a high degree of identity to the highly conserved DBD of proteins in the HSF superfamily. I also identified an HR-A/B trimerization domain, similar to the domain present in other zebrafish HSF proteins, in both HSF4 isoforms spanning amino acid sequence 133-220 (Fig. 4.2A and B) by sequence alignment. The HR-A/B domain corresponds to a predicted coiled coil structure, characteristic of a leucine zipper domain, identified in the HSF4 sequence by the UniProtKB and Coils version 2.2 software shown in Fig. 4.3C (Parry 1982; Lupas et al. 1991).

I was not able to identify a second coiled-coil structure or leucine zipper characteristic of an HR-C domain in the carboxy terminus of the HSF4a isoform using software analysis or with a sequence alignment. The lack of an HR-C trimerization inhibition domain in the carboxy terminus of the protein is consistent with what has been reported for mammalian HSF4 protein sequences (Nakai et al. 1997; Pirkkala 2001). However, the predicted protein sequence for this novel HSF4a is only 286 amino acid residues in length. This is effectively 177-250 amino acids shorter than any other HSF identified in a vertebrate species and is unique within the HSF family among the animal kingdom. Thus, the lack of an HR-C domain in the carboxy terminus is not due to the lack of a heptad repeat sequence in the amino acids spanning that region of the protein as is the case for mammalian HSF4s, but instead to the fact that this region is completely absent in the zebrafish HSF4a protein.

Figure 4.2 Comparison of zebrafish heat shock factors protein sequences.
(**A**) Zebrafish HSF1, HSF2, and HSF4 proteins are compared with each other by protein sequence alignment. Residues that are identical between the three sequences are shaded in black while positions with conserved amino acid substitutions are shaded in grey. The DNA binding domain (DBD), trimerization domain (HR-A/B) and inhibition of trimerization domain (HR-C) are indicated above the HSF1 sequence. The arrow indicates the final amino acid in the HSF4a isoform sequence. (**B**) A diagrammatic representation of zebrafish HSF proteins and the corresponding human ortholog comparing the location of specific protein domains within each sequence. The range of amino acids comprising each domain is indicated above each domain and the total number of amino acids in each protein is shown at the end of each line.

A


DBD
HR-A/B
HR-C
HSF1 1 MEYHSVGPGGVVV---TGNNVPAFLTKLWTLVEDPDTDPLICWSPNGTSFHVFDQGRFSK
HSF2 1 -------------MKHSSNVPAFLTKLWTLVEDSDTNEFICWSQEGNSFLVLDEQRFAK
HSF4 1 ---MQENPGSIGVDGGYASNVPAFLTKLWTLVEDPETNHLICWSATGTSFHVFDQGRFAK
HSF1 58 EVLPKYFKHNNMASFVRQLNMYGFRKVVHIEQGGLVKPEKDDTEFQHPYFIRGQEQLLEN
HSF2 47 EILPKFFKHNNMASFVRQLNMYGFRKVMHIDSGIVKQERDGPVEFQHPYFKHGQDDLLEN
HSF4 58 EVLPKYFKHNNMASFVRQLNMYGFRKVVNIEQSGLVKPERDDTEFQHLYFLQGHEHLLEH
HSF1 118 IKRKVTTVSNIKHEDYKFSTDDVSKMISDVQHMKGKQESMDSKISTLKHENEMLWREVAT
HSF2 107 IKRKVSNA---RPEESKIRQDDLSKILTSVQSVHEQQENMDARLATLKRENEALWTELSD
HSF4 118 IKRKVSIV---KSEETKVRQEDLSKLLYEVQVLRSQQENMEMQMQDMKQQNDVLWREVVS
HSF1 178 LRQKHSQQQKVVNKLIQFLITLARS--NRVLGVKRKMPLMLNDSSSAHSMPKFSRQYSLE
HSF2 164 LRKVHVQQQQVIKELVQFIFTLVQN--NRLLNLKRKRPLALNINGK---KSKFIHQLFEE
HSF4 175 LRQNHTQQQKVMNKLIQFLFSQMQSNSPSTVGMKRKLPLMLDDGCSTPPASKFSHNHSME
HSF1 236 SPAPSST---AFTGTGVFSSESPVKTGPIISDITELAQSSPVATDEWIE-DRTSPLVHIK
HSF2 219 PIDHSKT------TVNGLKNNSDISDDVIICDITDEDPEVTDGISVPDEE--DAEIVEIT
HSF4 235 SLQESFYIQSPSTESASCSTSSVMTGGPIISDVTEIPQSSSMALQMQAEESREKCLMLIK
HSF1 292 EEPSSPAHSPEVEEVCPVEVEVGAGSDLPVDTPLSPTTFINSI-LQESEPVFRPDS--AP
HSF2 271 YESSPKTVQQDSANG----------TIVNSTAHQEAEPKTTHDVSTTN
HSF4 295 EEPVSPGVQGRAEGVPLGSC------EVCAEPPVLPVAMVQSV-LEGRGSNLGERR--AK
HSF1 349 SEQKCLSVACLDNYPQMSEITRLFSGFSTSSLHLRPHSGTELHDHLESIDSGLENLQQIL
HSF2 309 SALQLNKPSCLSLEDPMKLMDSILSENGAISQNINLLGKVELMDYLDSIDCSLEDFQAML
HSF4 346 ----------RPMLERPEIPDGVENVDMSLEDLQLLL
HSF1 409 NAQSINFDSS-PLFDI---------FSSAASDVDLDSLASIQDLLSPDPVKETESGVDTD
HSF2 369 YGKQFNIDPDIIEETISE---TKN-NNSKANKENLDSEN----------
HSF4 373 RSHQQSMENNAAAMDQFTFSLPLNEWNFAEMDPNLK--SELANALIPAAVSQYM--FQGQ
HSF1 459 SGKQ--LVQYTSQPSFSPIPFS-----TDSSSTDLPMLLELQDDSYFSSEPTEDPTIALL
HSF2 404 --TDKQLIQYTTCPLMAFLDGCTPLTPSDPQPEDPPDLLD---ESL----EMKTPRSSLI
HSF4 429 EGELYPTAGYEEQ----------
HSF1 512 NFQPVPEDPSRT----------RIGDPCFKLKKESKR
HSF2 455 RLEPLTEAEANEATLFYLCELNSDLPNADTPPLDI------
HSF4 ----------

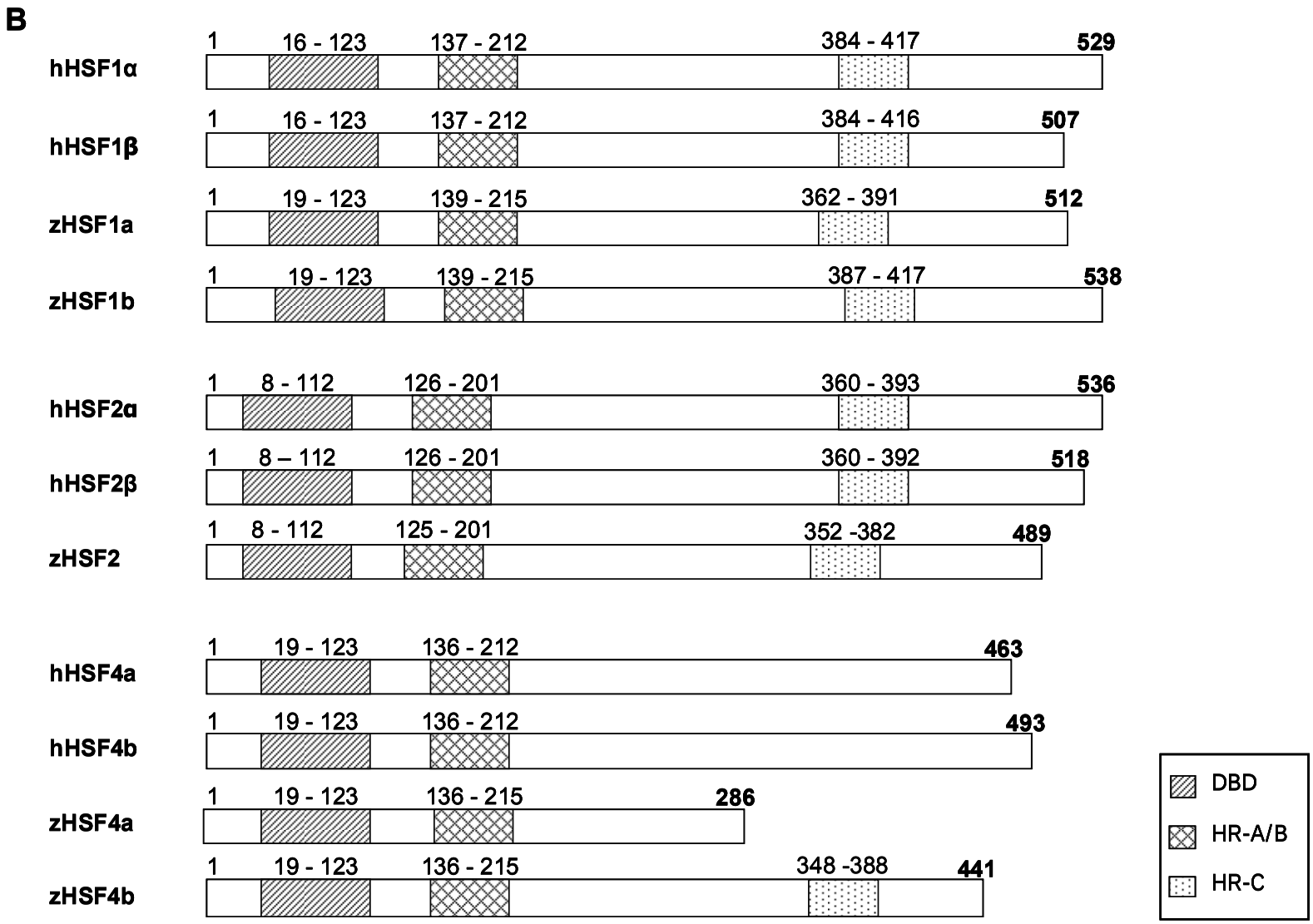
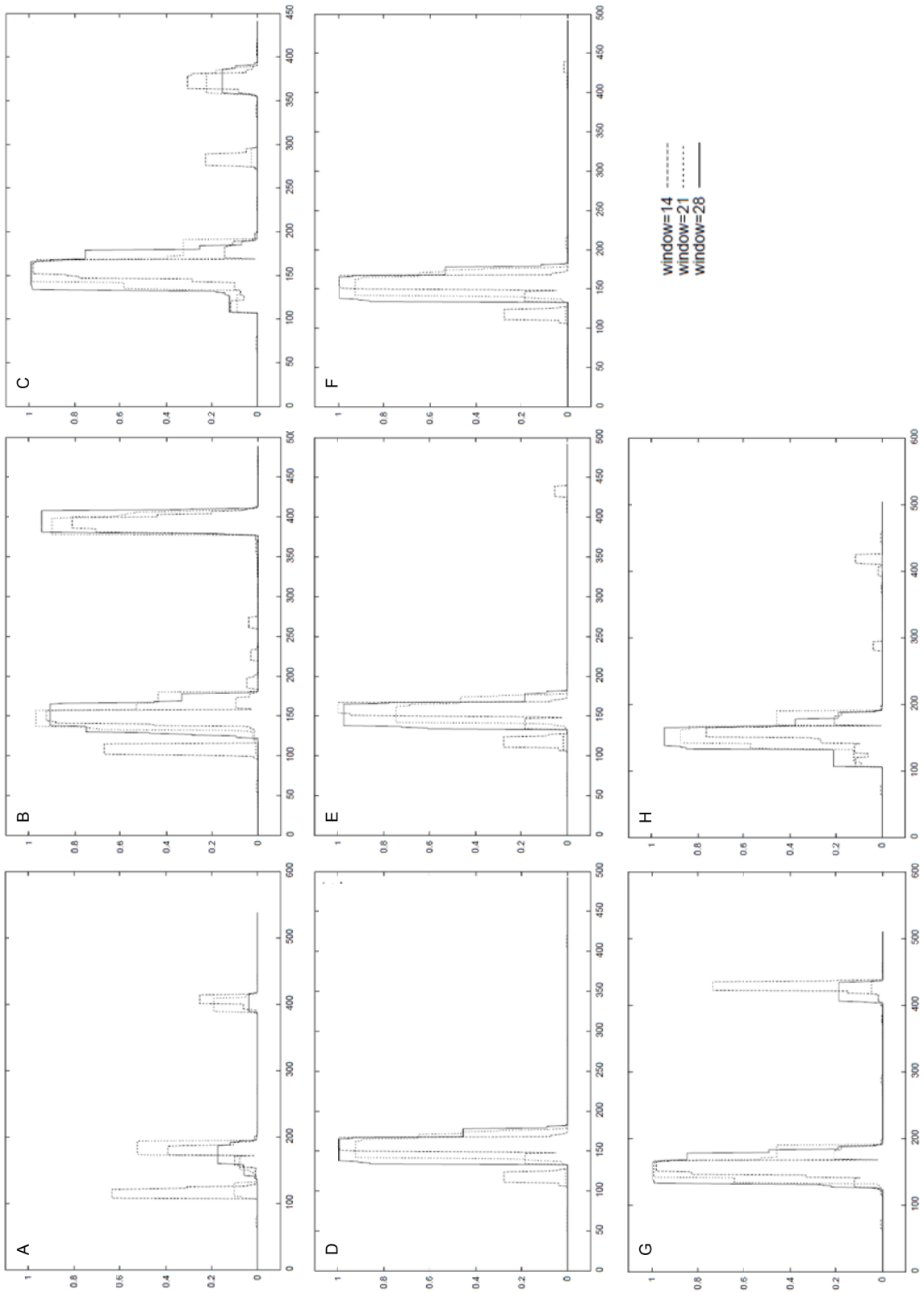

B
hHSF1α
1
16 - 123
137 - 212
384 - 417
529
hHSF1β
1
16 - 123
137 - 212
384 - 416
507
zHSF1a
1
19 - 123
139 - 215
362 - 391
512
zHSF1b
1
19 - 123
139 - 215
387 - 417
538
hHSF2α
1
8 - 112
126 - 201
360 - 393
536
hHSF2β
1
8 – 112
126 - 201
360 - 392
518
zHSF2
1
8 - 112
125 - 201
352 -382
489
hHSF4a
1
19 - 123
136 - 212
463
hHSF4b
1
19 - 123
136 - 212
493
zHSF4a
1
19 - 123
136 - 215
286
zHSF4b
1
19 - 123
136 - 215
348 -388
441
DBD
HR-A/B
HR-C

Figure 4.3 Prediction of the occurrence of coiled-coil structures in heat shock factors from several species. The Coils software version 2.2 (ch.EMBnet.org) was used to predict the probability of amino acid sequences within HSF proteins to adopt a coiled-coil conformation. The software compares the target sequence to a database of known sequences that produce coiled-coil structures by scanning amino acid sequence in 14, 21, or 28 amino acid windows (as indicated) and generating a comparative score. Sequences with a score closer to 1 (y-axis) have a higher probability of adopting a coiled-coil conformation. The number of amino acids in each sequence is indicated on the x-axis. The graph in each panel illustrates the potential for the HSF sequences to adopt a coiled-coil conformation and the region of amino acids where they could occur. Zebrafish HSF1 (**A**), HSF2 (**B**) and HSF4 (**C**) were analyzed to identify regions of possible coiled-coil conformation. The probability of the existence of coiled-coil structures was also assessed in human (**D**), dog (**E**), horse (**F**), chicken (**G**), and zebrafinch (**H**) HSF4 protein sequences.


A
B
C
D
E
F
G
H
window=14
window=21
window=28

The protein sequence of the putative HSF4b is predicted to contain 441 amino acid residues which is 155 amino acids longer than the HSF4a protein. The amino acid length of HSF4b is similar in amino acid length to other members of the heat shock factor family (Fig. 4.2B). Surprisingly, I was able to identify a putative HR/C domain in the HSF4b sequence spanning the 9th and 10th exon by analyzing an alignment of all three zebrafish HSF sequences identified a (Fig. 4.2A). This was also a unique discovery as all mammalian HSF4s identified to date are characterized as lacking this domain (Nakai et al. 1997; Åkerfelt et al. 2007; Abane and Mezger 2010). Analysis of the HSF4b amino acid sequence by SMART and Pfam software identified a vertebrate heat shock transcription factor domain between amino acids 245-422. Further my analysis of this region with UniProtKB and Coils software identified a potential coiled coil structure in this region comprised of amino acids 348-388 which would constitute a putative HR-C domain (Fig. 4.3C). Closer inspection of this region revealed the putative HR-C domain contains 15 hydrophobic amino acids, 10 and 13 of which are conserved in the zebrafish HSF1 and HSF2 protein sequences respectively. Similar coiled coil structures were identified using the Coils program in predicted chicken and in zebrafinch HSF4 proteins (Fig. 4.3G-H). As previously reported, no similar structures could be identified in mammalian HSF4 protein sequences using Pfam or the Coils software (Fig. 4.3D-F).

4.1.3 Phylogenetic analysis of zebrafish heat shock factors protein sequences

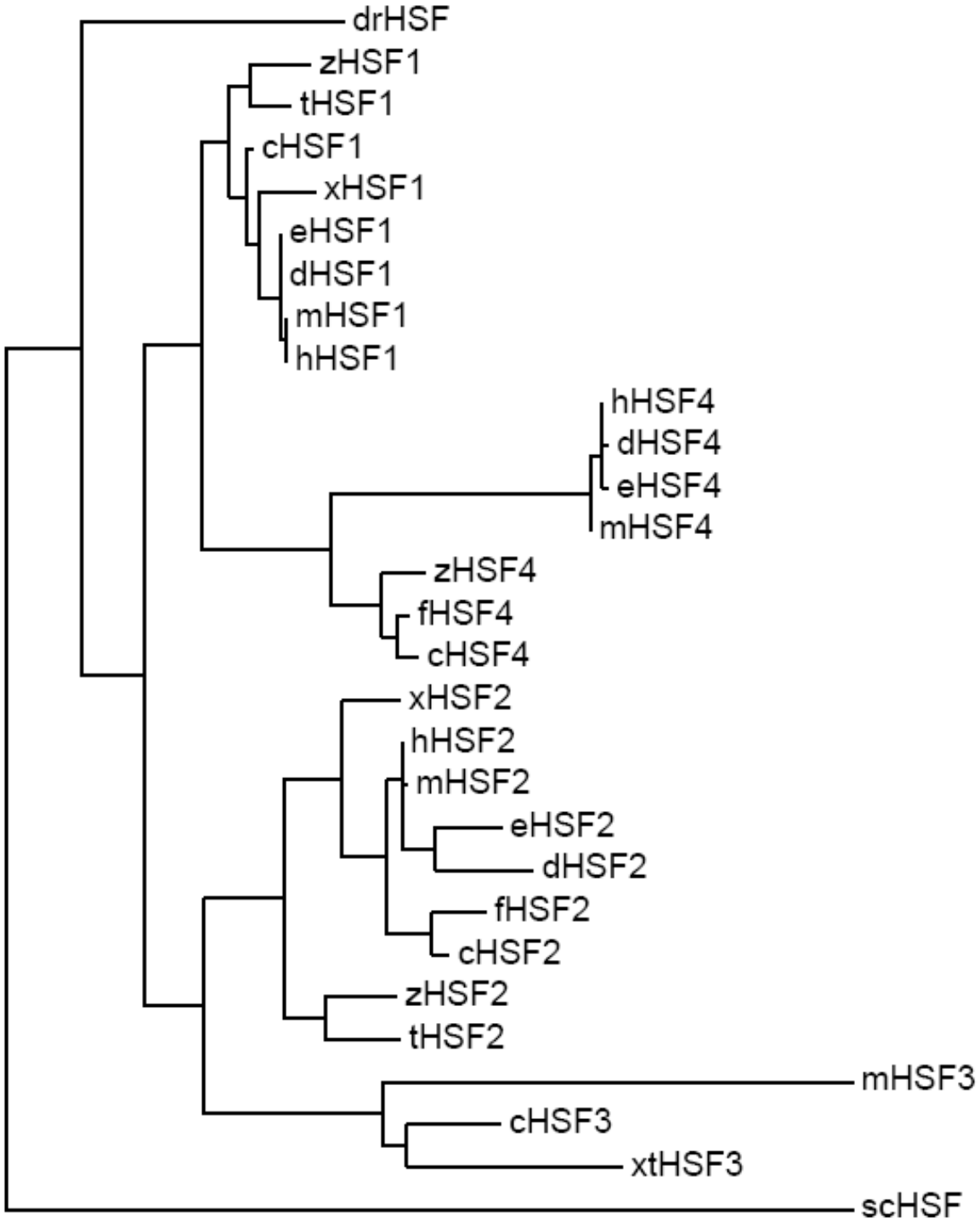
There is a substantial difference in amino acid length of the zebrafish HSF4a protein when compared to known HSF sequences, and the zebrafish HSF4b protein, unlike mammalian HSF4 proteins, appears to contain a putative HR-C domain. These unique characteristics observed in the novel zebrafish proteins led to a thorough *in silico* analysis of the predicted amino acid composition. I performed an extensive phylogenetic comparison of HSF1, HSF2, HSF3 and HSF4 protein sequences from eleven species to identify to which HSF group the novel zebrafish HSF4 protein was most closely related. Phylogenetic analysis shown in Fig. 4.4 was derived by comparing a single isoform of each of the HSF families from each species for simplicity. Preliminary alignments demonstrated that protein isoforms from the same species always clustered together and branching distances between species or HSF families were not different from those shown in Fig. 4.4. The zebrafish HSF4 protein clusters in a distinct group

with the chicken and zebrafinch HSF4 proteins near the mammalian HSF4 sequence cluster (Fig. 4.4). However, the branch for the zebrafish HSF4 protein occurs sooner and is closer to the HSF1 group than the mammalian HSF4 cluster. Additionally, the distance between the zebrafish HSF4 protein and the mammalian HSF4 cluster is greater than that observed between the fish and mammalian proteins for either HSF1 or HSF2. Another interesting observation from this analysis is that the zebrafish HSF4 protein clusters with two avian HSF4 protein sequences, the chicken and zebrafinch. While the chicken and zebrafinch branches are close to each other, similar to what is observed for the HSF2 protein from these same species (Fig. 4.4), the close grouping of the avian proteins to zebrafish proteins appears to be unique to the HSF4 group. In contrast, the grouping of chicken and zebrafinch HSF2 proteins is much closer to that of the mammalian HSF2 while the zebrafish and trout HSF2 branch occurs more distant from the mammalian HSF2 proteins. Similarly, the chicken HSF1 protein is more closely grouped with the mammalian HSF1 proteins while the trout and zebrafish HSF1 proteins branch is closer to the single *Drosophila* HSF than to the mammalian HSF1.

4.1.4 Comparison of zebrafish HSF4 protein domains to other zebrafish and non-zebrafish heat shock factor proteins

The early branching and distinct clustering of the zebrafish HSF4 proteins observed in the phylogenetic analysis (Fig. 4.4) led me to perform a more thorough comparison of the amino acid sequences comprising the HSF4 proteins to other members of the HSF protein family (Tables 4.1-4.4). The putative amino acid sequences of the DBD, HR-A/B and HR-C protein domains of the three zebrafish HSFs were compared between themselves, and to HSF protein sequences from other species (species and accession numbers in Appendix 1). The comparative analysis of these three domains was performed using sequence from a single isoform of each HSF from each animal species as the amino acid sequence of these domains does not vary between isoforms within a species. However, sequences from both isoforms of the zebrafish HSF4 proteins were analyzed as HSF4b contains the HR-C domain which is absent in HSF4a.

Figure 4.4 Phylogenetic analysis of HSF4 protein sequences. An unanchored phylogenic tree of HSF protein sequences from 11 species deposited in Genbank was generated using the Phylogeny.fr software. These sequences include those identified for zebrafish (z), rainbow trout (t), human (h), mouse (m), dog (d), equine (e), *Xenopus laevis* (x), *Xenopus tropicalis* (xt), *Saccharomyces cerevisiae* (sc), *Drosophila* (dr), zebrafinch (f), and chicken (c). Tree branch lengths depict relative evolutionary distances between proteins.


drHSF
zHSF1
tHSF1
cHSF1
xHSF1
eHSF1
dHSF1
mHSF1
hHSF1
hHSF4
dHSF4
eHSF4
mHSF4
zHSF4
fHSF4
cHSF4
xHSF2
hHSF2
mHSF2
eHSF2
dHSF2
fHSF2
cHSF2
zHSF2
tHSF2
mHSF3
cHSF3
xtHSF3
scHSF

Comparative analysis of domain amino acid sequences revealed that both isoforms of the zebrafish HSF4 protein share a high degree of amino acid identity with the DBD of zebrafish HSF1 (85%) and HSF2 (73%) (Fig. 4.2A and Table 4.1). Interestingly, the DBD of the predicted HSF4 proteins also shares 84-85% identity with the DBD of HSF1 proteins from seven other vertebrate species (Table 4.2) but, surprisingly, only a 72% amino acid identity with four mammalian HSF4 proteins (Table 4.1) and a maximum of 73% identity to HSF2 proteins from eight other species (Table 4.3). The HSF3 family is the least similar, sharing only 53-63% amino acid identities in the DBD with any of the zebrafish HSFs (Table 4.4). Of particular interest is the degree of the amino acid sequence identity observed between the zebrafish HSF4 proteins and those predicted for two avian species, the chicken and the zebrafinch (Table 4.1). The zebrafish HSF4a and HSF4b protein sequences had a total amino acid identity of 77% and 70% with the chicken respectively, and 79% and 73% with the zebrafinch respectively, which is higher than what was calculated for either of the HSF4 isoforms in comparison to other zebrafish HSFs (53-57% for HSF1 and 41-47% for HSF2) as well as being higher than that observed for comparisons to any mammalian HSF4 sequence, (52-54%) (Table 4.1). When only the DBD and HR-A/B protein domains are compared between zebrafish and the two avian species, the similarity becomes even more striking.

The DBD of the zebrafish HSF4 protein has a 94% amino acid identity with chicken and a 96% identity with the zebrafinch while the HR-A/B domain has an 82% and 83% identity, respectively. The similarities observed for these protein domains from both avian species and zebrafish HSF4 represent the highest level of amino acid conservation between the HSF4 protein domains in all species compared (Tables 4.1-4.4). Closer analysis of the two avian sequences revealed that the amino acid composition of the avian HSF4 DBD domains also shares a greater degree of identity with HSF1 DBDs than with mammalian HSF4 DBDs (data not shown).

There is a 49% similar amino acid identity between the predicted HR-C domain of the zebrafish HSF4 and the HSF4 sequence from chicken and zebrafinch. Once again, the highest number of amino acid similarities occurs between the zebrafish and avian species, similar to what was observed for the DBD and the HR-A/B domain comparisons. The HSF4b putative HR-C sequence has 24-37% amino acid similarity with the HR-C domains of HSF1, HSF2 and HSF3 in all species analysed.

Table 4.1 Percent amino acid identity of HSF4 protein functional domains between zebrafish and other species. Amino acid sequences of the total protein, the DNA binding domain (DBD), heptad repeat A/B trimerization domain (HR-A/B) and the heptad repeat C (HR-C) domain of HSF4 sequences from human (h), mouse (m), dog (d), Equine (e), chicken (c), and zebrafinch (f), were compared to zebrafish (z) HSF sequences using NCBI BLAST program. Amino acid sequences that displayed higher than 80% identities are shown by shaded boxes. The number of amino acids in each of the three domains analyzed and their positions within each of the zebrafish proteins is shown in figure 4.2B. Percent of amino acid identities calculated to be less than 10% were considered to be not significant (NS). Protein sequences that lacked an HR-C domain are denoted with an NA as not applicable.

Species	Amino acid (#)	zHSF1				zHSF2				zHSF4				
		Total	DBD	HR-A/B	HR-C	Total	DBD	HR-A/B	HR-C	Total HSF4a	Total HSF4b	DBD	HR-A/B	HR-C
zHSF4a	286	30	85	52	NA	27	73	45	NA					NA
zHSF4b	441	53	85	52	27	41	63	45	27					
hHSF4	492	30	75	45	NS	26	63	35	16	53	49	72	56	34
mHSF4	462	26	75	48	NS	20	63	36	NS	52	46	72	59	24
dHSF4	492	28	75	44	NS	27	63	35	NS	54	50	72	56	32
EHSF4	492	28	75	45	NS	27	63	35	NS	52	49	72	55	32
cHSF4	510	37	84	55	NS	31	70	45	32	77	70	94	82	49
fHSF4	504	37	86	53	NS	33	71	45	32	79	73	96	83	49

Table 4.2 Percent amino acid identity of HSF1 protein functional domains between zebrafish and other species. Amino acid sequences of the total protein, the DNA binding domain (DBD), heptad repeat A/B trimerization domain (HR-A/B) and the heptad repeat C (HR-C) domain of HSF1 sequences from human (h), mouse (m), dog (d), Equine (e), *Xenopus laevis* (X), chicken (c), Rainbow trout (t), *Drosophila* (dr) and *Saccharomyces cerevisiae* (sc) were compared to zebrafish (z) HSF sequences using NCBI BLAST program. Amino acid sequences that displayed higher than 80% identities are shown by shaded boxes. The number of amino acids in each of the three domains analyzed and their positions within each of the zebrafish proteins is shown in figure 4.2B. of amino acid identities calculated to be less than 10% were considered to be not significant (NS).

Species	Amino acid (#)	zHSF1				zHSF2				zHSF4				
		Total	DBD	HR-A/B	HR-C	Total	DBD	HR-A/B	HR-C	Total HSF4a	Total HSF4b	DBD	HR-A/B	HR-C
zHSF1	538					38	73	49	35	57	53	85	49	27
hHSF1	529	58	91	75	42	38	73	51	NS	58	63	84	56	32
mHSF1	503	56	91	74	42	38	73	49	NS	56	47	84	55	32
dHSF1	554	61	91	77	48	33	73	51	NS	57	63	84	61	32
eHSF1	590	58	91	77	45	37	73	51	NS	58	47	84	60	29
xHSF1	530	59	93	74	48	37	75	49	32	59	47	84	55	37
cHSF1	491	55	93	75	45	36	72	49	35	58	50	85	56	29
tHSF1	513	66	96	75	77	37	74	53	35	57	48	85	55	NS
drHSF	691	27	56	56	39	20	57	42	13	36	49	56	48	NS
scHSF	833	12	47	21	NS	26	45	22	NS	23	50	44	20	24

Table 4.3 Percent amino acid identity of HSF2 protein functional domains between zebrafish and other species. Amino acid sequences of the total protein, the DNA binding domain (DBD), heptad repeat A/B trimerization domain (HR-A/B) and the heptad repeat C (HR-C) domain of HSF2 sequences from human (h), mouse (m), dog (d), Equine (e), *Xenopus laevis* (X), chicken (c), zebrafinch (f) and Rainbow trout (t) were compared to zebrafish (z) HSF sequences using NCBI BLAST program. Amino acid sequences that displayed higher than 80% identities are shown by shaded boxes. The number of amino acids in each of the three domains analyzed and their positions within each of the zebrafish proteins is shown in figure 4.2B.

Species	Amino acid (#)	zHSF1				zHSF2				zHSF4				
		Total	DBD	HR-A/B	HR-C	Total	DBD	HR-A/B	HR-C	Total HSF4a	Total HSF4b	DBD	HR-A/B	HR-C
zHSF2	489	35	73	49	42					47	41	73	45	27
hHSF2	536	36	72	49	39	60	90	61	77	44	38	73	41	32
mHSF2	517	37	72	49	39	59	89	61	77	44	39	73	41	32
dHSF2	578	33	59	49	39	56	74	61	77	36	36	61	41	32
eHSF2	533	33	58	49	39	56	74	61	77	38	35	58	41	32
xHSF2	515	36	73	45	35	58	90	65	65	41	38	73	39	32
cHSF2	563	36	69	52	39	59	80	61	77	41	38	72	41	32
fHSF2	630	32	66	48	39	57	77	60	77	42	38	70	45	32
tHSF2	511	36	72	55	35	68	95	66	87	47	40	70	43	27

Table 4.4 Percent amino acid identity of HSF3 protein functional domains between zebrafish and other species. Amino acid sequences of the total protein, the DNA binding domain (DBD), heptad repeat A/B trimerization domain (HR-A/B) and the heptad repeat C (HR-C) domain of HSF3 sequences from mouse (m), *Xenopus tropicalis* (Xt), and chicken (c) were compared to zebrafish (z) HSF sequences using NCBI BLAST program. Amino acid sequences that displayed higher than 80% identities are shown by shaded boxes. The number of amino acids in each of the three domains analyzed and their positions within each of the zebrafish proteins is shown in figure 4.2B. of amino acid identities calculated to be less than 10% were considered to be not significant (NS).

Species	Amino acid (#)	zHSF1				zHSF2				zHSF4				
		Total	DBD	HR-A/B	HR-C	Total	DBD	HR-A/B	HR-C	Total HSF4a	Total HSF4b	DBD	HR-A/B	HR-C
mHSF3	492	16	53	29	NS	27	53	34	NS	28	39	50	28	NS
cHSF3	467	31	62	49	NS	34	63	51	45	36	35	63	48	24
xtHSF3	550	20	63	40	42	36	63	44	48	46	40	60	43	34

An interesting observation is that the HR-C domain from the zebrafish shares greater amino acid identitiy with HR-C domains in mammalian and *Xenopus* HSF1 and HSF2 proteins (29-37%) than it does with either the zebrafish HSF1 or HSF2 proteins (27%). During the comparative sequence analysis it was observed that the zebrafish HSF2 protein actually has two coiled-coil domains containing a total of 20 hydrophobic amino acids in HR-C region. This is also a unique characteristic of the zebrafish HSF2 protein as only one coiled coil domain has been characterized in other HSF2 proteins in vertebrates.

Mammalian HSF4 proteins are characterized as HSFs that do not possess a heptad repeat sequence (HR-C) coiled coil domain in the carboxy end of their protein. Analysis of several mammalian HSF4 sequences with both Pfam and the Coils software confirmed that the mammalian HSF4 sequences do not contain the same vertebrate heat shock factor domain or predicted coiled coil domain as observed in the zebrafish HSF4b sequence (section 4.1.2). However, even though a coiled-coil domain does not exist in mammalian HSF4, the HR-C domain of the zebrafish HSF4b still shares 24-32% amino acid identities within the same carboxy region of the mammalian HSF4 sequences which is similar to the percentage identities calculated for HSF1 and HSF2 proteins above.

4.2 Binding affinities of zebrafish HSF1a, HSF2 and HSF4a for heat shock response elements

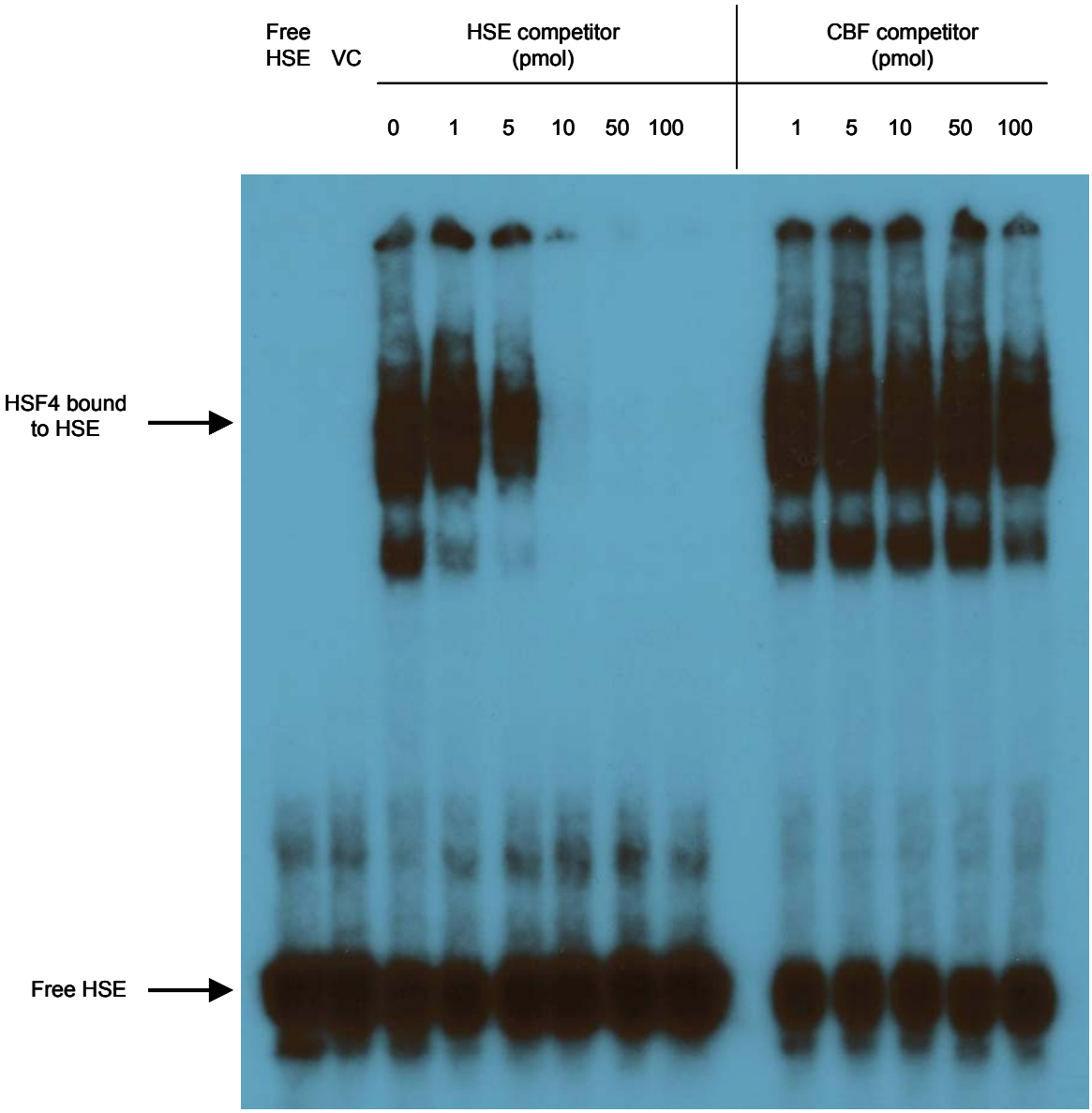
The data presented in sections 4.1.2, 4.1.3 and 4.1.4 strongly suggest that the amino acid sequence of the HSF4 DNA binding domain more closely resembles that of zebrafish and mammalian HSF1 proteins than it does mammalian HSF4. The ability of the HSF4 protein to bind to HSE DNA binding sites was therefore investigated by electrophoretic mobility shift assay (EMSA) using purified HSF4 protein. The purified HSF4 protein was generated and isolated from a pRSET bacterial expression system using the ORF sequence from the *hsf4a* gene as described in section 3.6. Protein purified from cells containing an empty pRSET vector and subjected to IPTG induction was used as a negative control for protein expression. Conditions for the bacterial expression system were optimized to determine the optimal growth temperature for the culture as well as the optimal concentration of IPTG needed for induction of maximal protein expression. Purification of the protein from bacterial culture was also optimized by

determining the concentration of imidizole required for maximum elution of proteins from the protein purification resin.

A dsDNA oligonucleotide designed to be a perfect continuous inverted repeat sequence of nGAAn was used as the HSE target (Table 3.3). The consensus HSE oligonucleotide and other oligonucleotides used for competition assays were labeled with ^{32}P for use as assay probes as described in section 3.8.1. Conditions for the HSE binding reaction were optimized by determining the appropriate dI-dC and protein concentrations required for the EMSA. Preliminary experiments also confirmed that performing the binding reactions at room temperature did not produce different results from those conducted at 4°C. As well, binding reactions performed with protein in buffer containing imidizole did not differ from those obtained using proteins in buffer without imidizole.

Purified HSF4 protein (750 ng) was added to 1 pmol of labeled HSE in a series of EMSA reactions containing increasing concentrations of either an HSE or CCAAT box (CBF) competitor and analysed as described in section 3.8.2. The results shown in Fig. 4.5 demonstrate that the HSF4 protein bound to HSE in a specific manner. The long, extended band produced by the HSF4 bound to HSE is typical of the banding pattern seen in most HSF shift assays. The long band is actually a compilation of several bands likely resulting from several different HSE-HSF4 complexes which migrate at different rates. The *in vitro* nature of the EMSA combined with the lack of post-translational modifications on the HSF4, such as phophorylation, provides an environment where HSFs have been shown to combine into sextamers instead of monomers. The truncated nature of the HSF4 protien suggests that the conformation of the trimer bound to the HSE will be different from that of longer HSF proteins which will likely affect the migration of this complex through the gel. It has also been demonstrated that HSEs containing more than three 5'-nGAAn-3' sequences are capable of binding two HSF trimers instead of one. Additionally, the nature of the inverted repeat sequence that comprises any HSE allows for potential folding of the HSE on itself. All of these factors together create an environment that allows for several different HSE-HSF complexes with different conformations to exist in the EMSA reaction, all of which would migrate at slightly different rates in the non-denaturing gel (Takemori et al. 2009; Yamamoto et al. 2009).

Figure 4.5 Zebrafish HSF4 protein binds to an HSE sequence in a specific manner. Binding reactions for EMSA analysis were prepared containing 1 pmol of ^{32}Plabeled HSE oligonucleotide and 750 ng of either HSF4 protein purified from a bacterial expression system or protein control purified from the same expression system containing an empty pRSET vector (VC). An unlabeled HSE oligonucleotide or CBF oligonucleotide competitor was added (0-100 pmol of competitor per reaction) to binding reactions containing zHSF4 protein. Shifted bands of HSF4 bound to HSE as well as unbound or free HSE are shown with arrows.


Free HSE
VC
HSE competitor (pmol)
CBF competitor (pmol)
0
1
5
10
50
100
1
5
10
50
100
HSF4 bound to HSE
Free HSE

The addition of unlabelled HSE to the reaction at a concentration of 50 times the amount of labeled HSE was sufficient to completely compete the protein away from the ^{32}P-labeled HSE while addition of the same concentration of the CBF oligonucleotide was unable to displace the HSF4 protein from the HSE oligonucleotide.

Recent studies have demonstrated that heat shock factors can have different binding affinities for discontinuous HSE sequences. For example, mammalian HSF1 preferentially binds HSE sites with perfect nGAAn inverted sequences while mammalian HSF4 has similar binding affinities for perfect and discontinuous nGAAn sequences (Fujimoto et al. 2008; Yamamoto et al. 2009; Sakurai and Enoki 2010). Therefore, a comparison of the ability of all three zebrafish HSF proteins to bind to discontinuous HSEs was determined using EMSA. The consensus HSE and five discontinuous HSE sequences, each designed with a different nucleic acid modification (Table 3.3) were used for the analysis. Purified HSF1 and HSF2 proteins were similarly generated as per section 3.6. Reactions were performed by adding 1 pmol of ^{32}P labeled HSE oligonucleotide or modified HSE oligonucleotide to 750 ng of purified protein. As expected, zebrafish HSF4 bound specifically to the perfect HSE, and little difference was observed between the abilities of the three zebrafish HSFs to bind to the perfect HSE (Fig. 4.6). All HSFs were able to bind HSEsn, an HSE containing a single nucleotide change in the consensus HSE sequence (Table 3.3). A relatively small amount of each HSF1 and HSF4 bound the HSEmod1 oligonucleotide while no protein binding was observed for three other modified HSE oligonucleotides. Further analysis of the relative binding affinity of all three zebrafish HSFs was performed by competition assay with unlabelled HSEsn, HSEmod1 and HSEmod2. Addition of 50-100 pmol of HSEsn was able to effectively displace zebrafish HSF2 from the labeled HSE oligonucleotide while a concentration of 300 pmol was required to effectively displace the HSF1 and HSF4 proteins (Fig. 4.7). Further, HSF2 was partially displaced from the labeled oligonucleotide when 1000 pmol of HSEmod1 or HSEmod2 was added to the reaction; the same concentrations of this competamer did not displace any of the HSF1 or HSF4 proteins (Fig. 4.7).

Figure 4.6 Zebrafish HSF proteins can bind to certain discontinuous HSEs. Zebrafish HSF1, HSF2 and HSF4 protein purified from a bacterial expression system (750 μg per reaction) were added to EMSA binding reactions containing either 1 pmol of ^{32}P labeled HSE oligonucleotides or modified discontinuous HSE oligonucleotide (listed in Table 3.3).

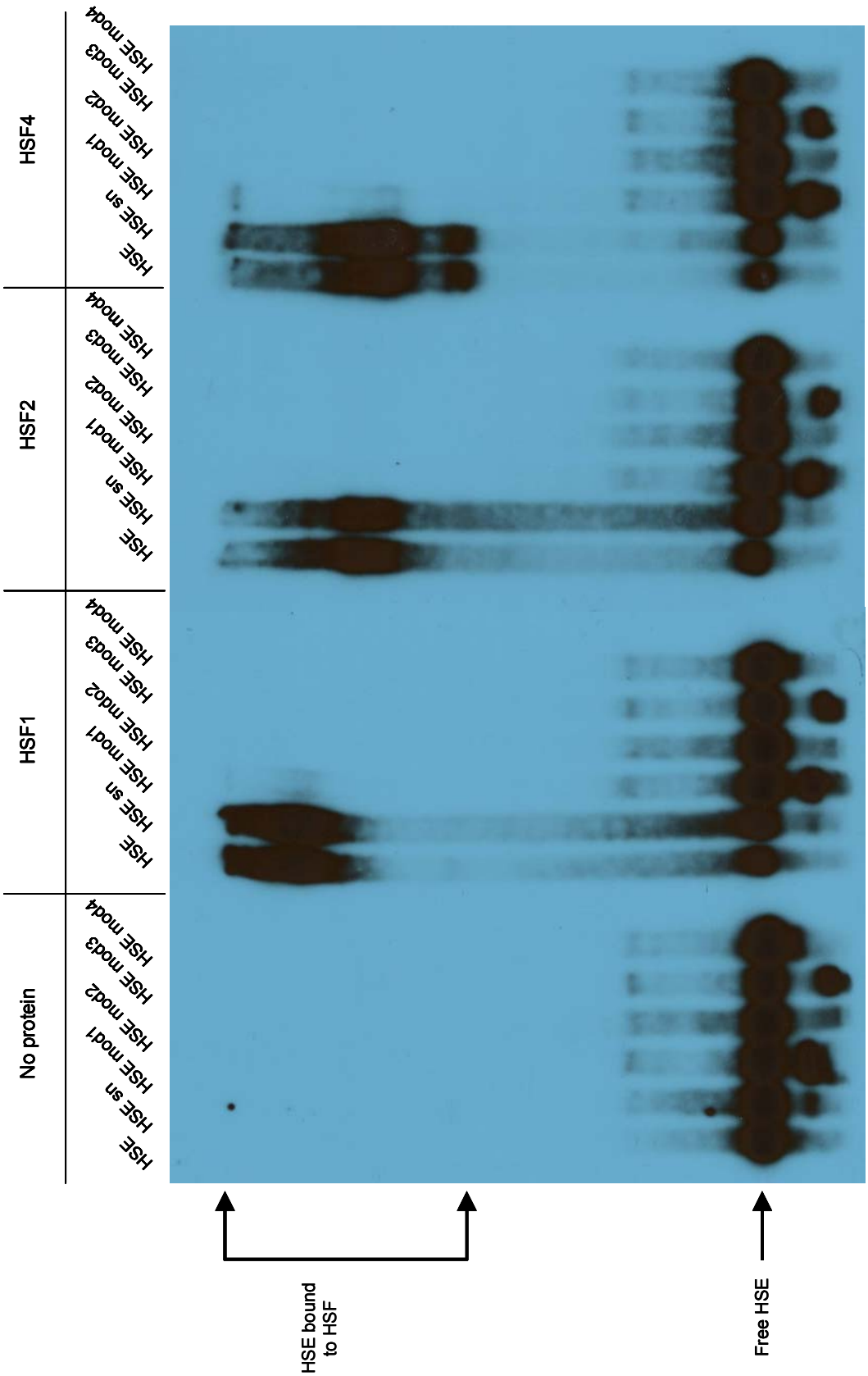
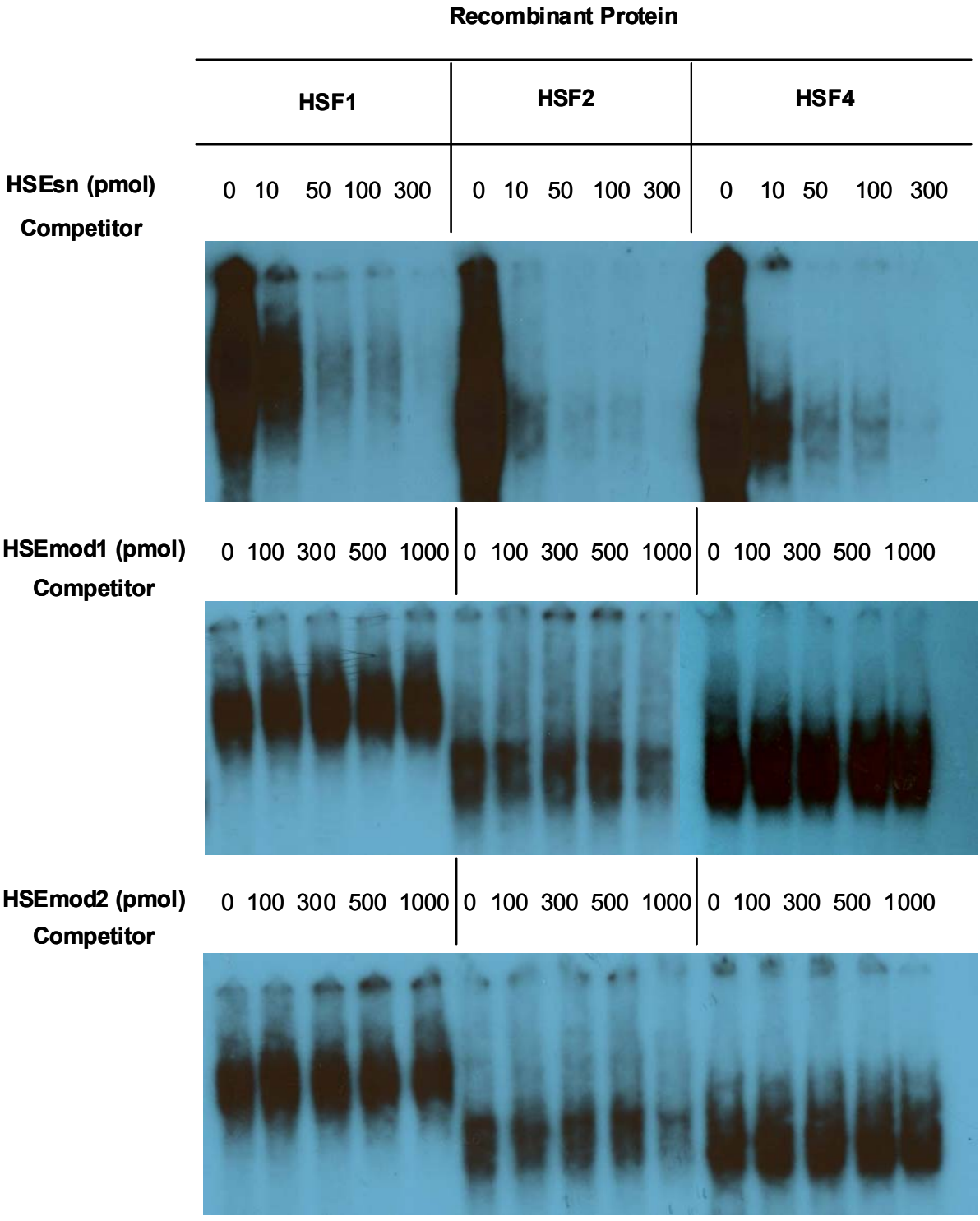

HSF4
HSF2
HSF1
No protein
HSE
HSE sn
HSE mod1
HSE mod2
HSE mod3
HSE mod4
HSE bound to HSF
Free HSE

Figure 4.7 Zebrafish HSF proteins have different binding affinities for discontinuous HSEs. Purified HSF1, HSF2, and HSF4 proteins (750 ng per reaction) were added to EMSA binding reactions containing 1 pmol of ^{32}P-labeled HSE oligonucleotide. An unlabeled modified HSE oligonucleotide was added to the binding reactions as a competitor as indicated. HSEsn oligonucleotide was added to each reaction in quantities ranging from 0 to 300 pmol of competitor per reaction. HSEmod1 or HSEmod2 was added to each reaction in quantities ranging from 0 to 1,000 pmol.


Recombinant Protein
HSF1
HSF2
HSF4
HSEsn (pmol) Competitor
0 10 50 100 300 0 10 50 100 300 0 10 50 100 300
HSEmod1 (pmol) Competitor
0 100 300 500 1000 0 100 300 500 1000 0 100 300 500 1000
HSEmod2 (pmol) Competitor
0 100 300 500 1000 0 100 300 500 1000 0 100 300 500 1000

4.3 Analysis of heat shock factor protein abundance in developing zebrafish

4.3.1 Characterization of zebrafish antibodies for HSF1, HSF2 and HSF4

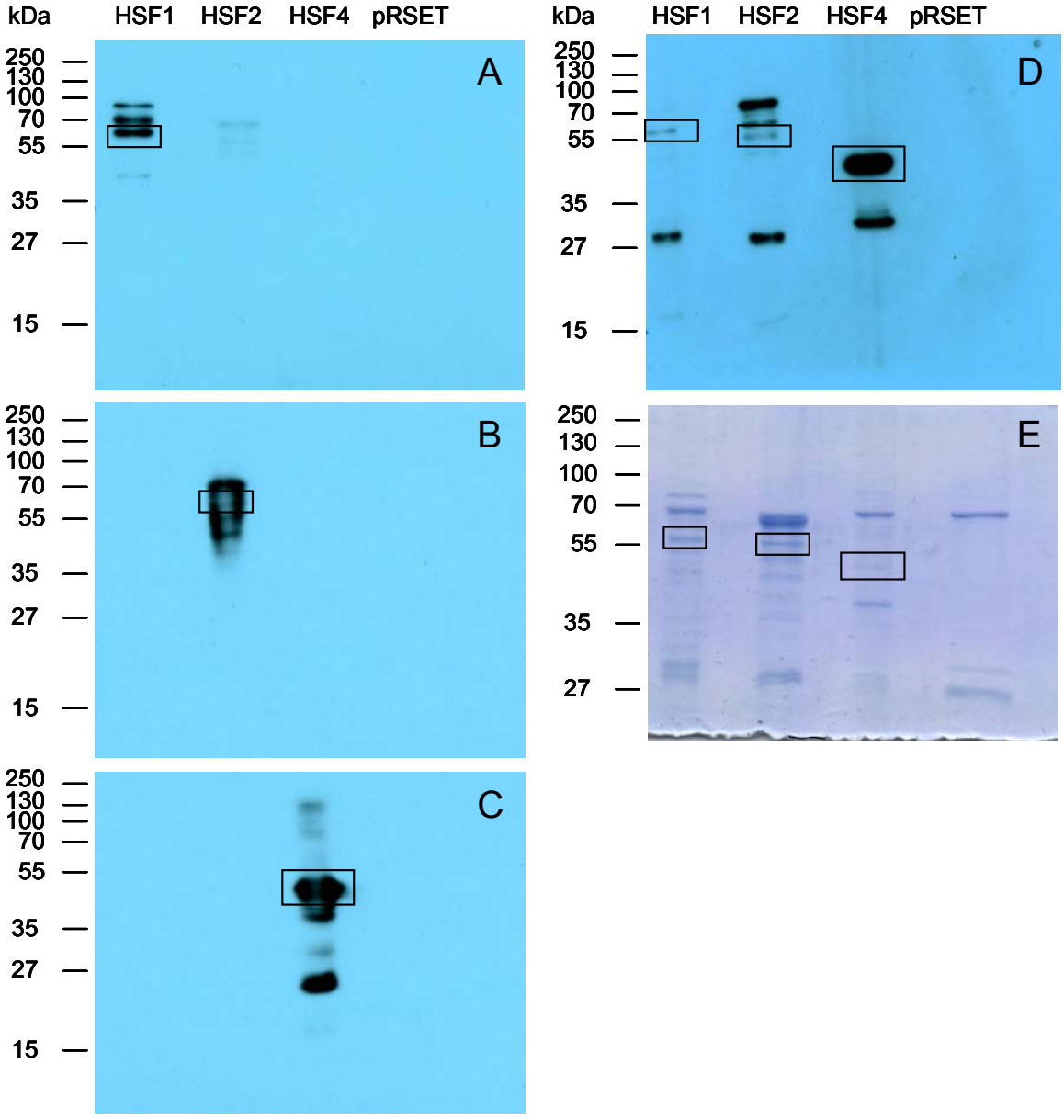
Polyclonal antibodies specific to zebrafish heat shock transcription factors and proteins were custom designed and generated commercially by Affinity BioReagents™ (Colorado). The specificity of the antibodies was determined by SDS-PAGE and Western blot analysis using zebrafish HSF1, HSF2 and HSF4 purified from a bacterial expression system (section 3.6.2). Protein purified from cells containing an empty expression vector and subjected to IPTG induction was used as a negative control for protein expression. The antibodies were characterized by determining the optimal antibody dilution, incubation times and stringency of membrane washes to use for Western blot analysis of the proteins described above.

One microgram of purified protein for each of the three HSFs was loaded into separate lanes on each of four individual SDS-PAGE gels (10% acrylamide). One microgram of purified proteins from the vector control were loaded as a fourth sample on each gel for use as a negative control. Proteins separated by SDS-PAGE electrophoresis were transferred to PVDF membranes. Each blot was then incubated with a primary antibody for one of HSF1, HSF2, HSF4 or the HIS fusion-tag. A fifth SDS-PAGE gel containing 0.5 μg of purified protein was prepared and stained with Coomassie blue to visualize total protein (section 3.7).

The antibody for each of the HSF proteins detected multiple protein bands from the purified proteins including a band of approximately the expected size for each; 60k Da for HSF1, 58 kDa for HSF2 and 35 kDa for HSF4 (Fig. 4.8A-C). The antibody for the HIS tag also detected protein bands of the expected sizes (Fig. 4.8D). None of the four antibodies detected any protein bands in the pRSET negative control confirming that the antibodies are specifically detecting the purified HSF targets. The results were interpreted to suggest that each antibody is specific for its own target since each antibody only detected protein bands in the lane specific for its target (Fig. 4.8A-D) although there may be a small amount of cross-reactivity of the HSF1 antibody with the HSF2 protein (Fig. 4.8A).

Taken together these results demonstrate that each of the custom-made zebrafish specific antibodies can specifically detect its target protein purified from a bacterial expression system. Therefore, the antibodies were determined to be appropriate for use in the analysis of zebrafish HSF proteins.

Figure 4.8 Zebrafish HSF antibodies detect purified target proteins in a specific manner as determined by Western blot analysis. One μg each of purified HSF1, HSF2, HSF4 and 10 μL of pRSET negative control were electrophoresed through four identical 10% SDS-PAGE gels (panels A-D). Separated proteins were transferred to PVDF membranes and incubated with primary antibody for HSF1 (**A**, 1:3000 dilution), HSF2 (**B,** 1:10,000 dilution), HSF4 (**C,** 1:5000 dilution) and HIS-tag antibody (**D**, 1:10,000). Protein bands were detected by a secondary antibody conjugated to horseradish peroxidase and visualized by chemiluminescence as described in section 3.7.3. For all panels, a rectangular box was placed around the band of expected size. A fifth gel with 0.5 μg of each HSF protein and 10 μL of pRSET negative control eletrophoresed through a 10% SDS-PAGE gel was used to visualize the proteins with Coomassie blue (**E**).

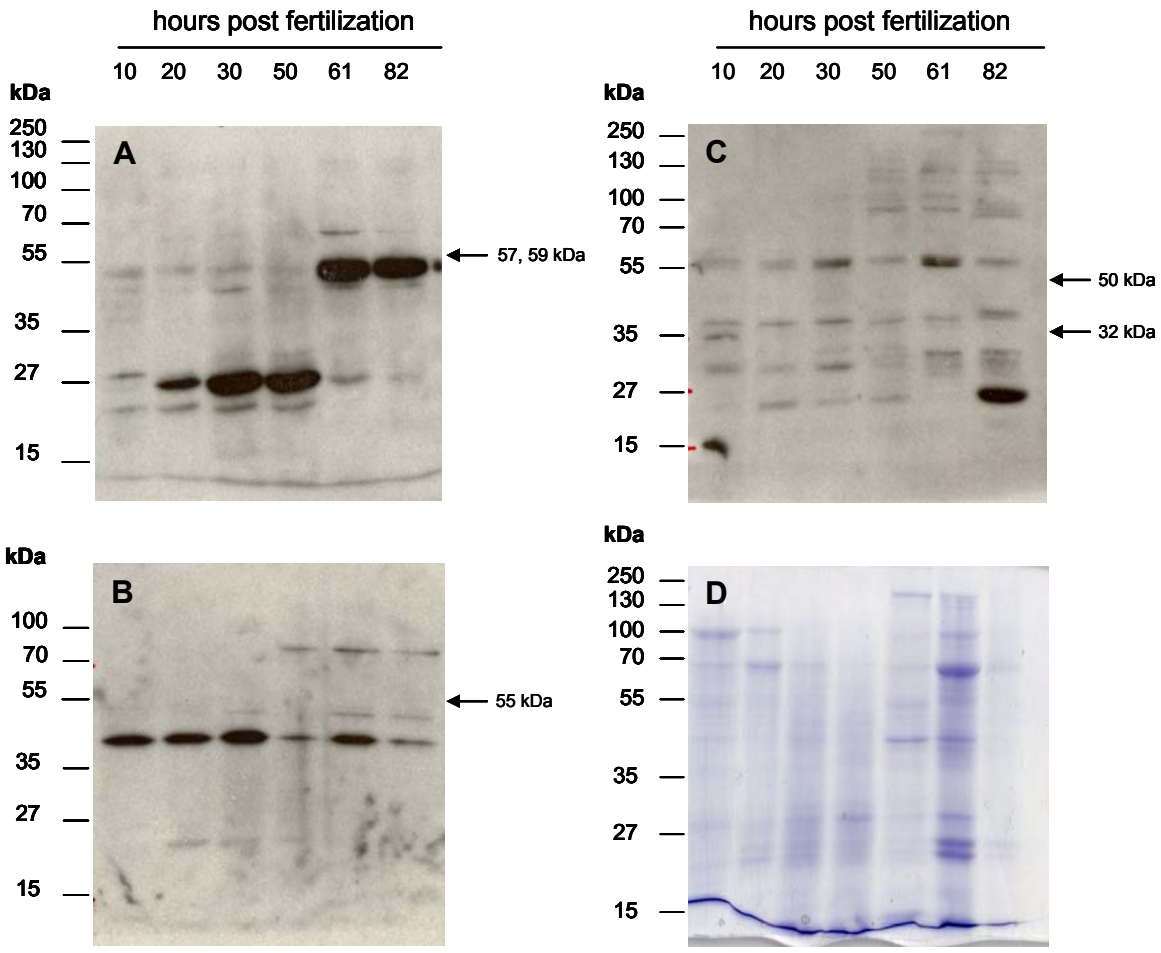

kDa
HSF1
HSF2
HSF4
pRSET
250
130
100
70
55
35
27
15
A
B
C
D
E

4.3.2 Western blot analysis of HSF1, HSF2 and HSF4 abundance in developing zebrafish

Preliminary experiments were performed to optimize the types and concentrations of protease inhibitors required to successfully deyolk and extract proteins from zebrafish embryos. The abundance of HSF proteins in zebrafish embryos was determined by performing Western blot analysis on total protein extracts from zebrafish embryos collected at six different developmental time points between 10 and 82 hpf (Fig. 4.9). Three SDS-PAGE gels containing 10 µg of total embryo protein from each of the time points were prepared, transferred to three separate PVDF membranes and analysed by Western blot (section 3.7.3) using polyclonal antibodies for HSF1, HSF2 and HSF4 described in section 4.3.1.as per section 3.7.3. A fourth identical SDS-PAGE gel was stained with Coomassie blue to visualize total protein content as per section 3.7.3.

Western blot analysis of embryo total protein demonstrated that a protein band corresponding to the expected molecular mass was detected for the HSF1 antibody (Fig. 4.9A). However, the most prominent band for the HSF2 antibody was observed between 40 and 55 kDa, below the expected 55 kDa (Fig. 4.9B). Each of these HSF antibodies also detected other protein bands in addition to the band at the expected molecular mass. The HSF4 antibody in particular detected multiple protein bands in the 30-40 kDa range of the target 35 kDa molecular mass suggesting that the HSF4 antibody may not be specific for the zebrafish HSF4 protein in an *in vivo* sample (Fig. 4.9C). The Western blots could be loosely interpreted to suggest that HSF1 and HSF2 can be detected at all developmental stages analyzed. However, it is difficult to ascertain if HSF4 is present due to the lack of certainty over which band of the several detected by the antibody actually corresponds to HSF4. It is also difficult to make any other definitive conclusions about the abundance of HSF proteins in zebrafish embryos due to lack of specificity of the antibodies.

Figure 4.9 Western blot analysis of zebrafish HSF protein expression during several developmental stages. Protein was extracted from zebrafish embryos at seven different stages of development. Ten μg of total protein from each time point were electrophoresed through four identical 10% SDS-PAGE gels. The proteins from the gels were transferred to PVDF membrane and incubated with the desired antibody: (**A**) HSF1 (1:3000 dilution) to detect HSF1b at 59 kDa and HSF1a at 57 kDa, (**B**) HSF2 (1:10,000 dilution) to detect the protein band at 55 kDa, and (**C**) HSF4 (1:5000 dilution) to detect HSF4b at 50 kDa and HSF4a at 32 kDa. Protein bands were detected by a secondary antibody conjugated to horseradish peroxidase and visualized by chemiluminescence as described in section 3.7.3. The sizes of the expected mass bands for each protein as described in section 3.7.1 are indicated with arrows. An identical gel was stained with Coomassie blue to visualize the protein bands (**D**).


hours post fertilization
10 20 30 50 61 82
kDa
250
130
100
70
55
35
27
15
A
57, 59 kDa
B
55 kDa
C
50 kDa
32 kDa
D

4.4 Analysis of heat shock factor mRNA expression

4.4.1 *In situ* hybridization fails to detect *hsf4*

Whole mount *in situ* hybridization of zebrafish embryos was used in an attempt to localize the expression pattern of the *hsf4* gene during development. Initial experiments were performed using a DIG-labeled antisense RNA probe generated by *in vitro* transcription (section 3.9.1.1). Similarly an antisense probe for the zebrafish *alpha-tropomyosin* gene was generated and used as an assay control (section 3.9.1.1). The expression of *hsf4* was assayed in embryos fixed at 24 and 36 hpf using the method described in section 3.9.2. No observable staining of mRNA was detected in any of these developmental stages even when the colorimetric reaction was allowed to develop for 3 hours (data not shown). The control *alpha-tropomyosin* probe, however, could be visualized in the above mentioned stages after 25-30 minutes (Fig. 4.10A and C).

The method of generating antisense RNA probes by *in vitro* transcription for use *in situ* hybridization is lengthy and can be inconsistent in producing good quality RNA probes. Once generated, the RNA probes are susceptible to degradation and have a short half life. These factors influence the quality of experimental results. For these reasons, additional *in situ* hybridization experiments were performed using a DIG-labeled single stranded DNA (ssDNA) probe prepared as described in section 3.9.1.2, to try and detect *hsf4* mRNA in developing embryos.

A comparison of results obtained for *alpha-tropomyosin* control mRNA expression using RNA probes (generated by *in vitro* transcription) to those obtained using ssDNA probes (generated by PCR) in embryos fixed at 24 and 36 hpf is shown in Fig. 4.10B and D respectively. *Alpha-tropomyosin* mRNA was clearly detected in somitic muscle with RNA probes (panels A and C) and with ssDNA probes (panels B and D). The embryo staining at 24 hpf with the ssDNA probe (panel B) was less intense than that visualized with the RNA probe (panel A). However, at 36 hpf the staining intensity appeared to be equivalent between the two types of probe (panels C and D). These results were interpreted to suggest that both types of probes were suitable to use for detection of mRNA by *in situ* hybridization technique.

Figure 4.10 Comparison of DIG labeled antisense RNA to DIG labeled ssDNA as probes for use in zebrafish whole mount *in situ* hybridization. Antisense RNA probes labeled with DIG residues were used to detect α-tropomyosin mRNA expression in embryos fixed at 24 hpf (**A**) and 36 hpf (**C**) as described in section 3.9. Similarily, DIG labeled ssDNA probe was used to detect α-tropomyosin mRNA expression at 24 hpf (**B**) and 36 hpf (**D**) as described in section 3.9.


B
D
A
C

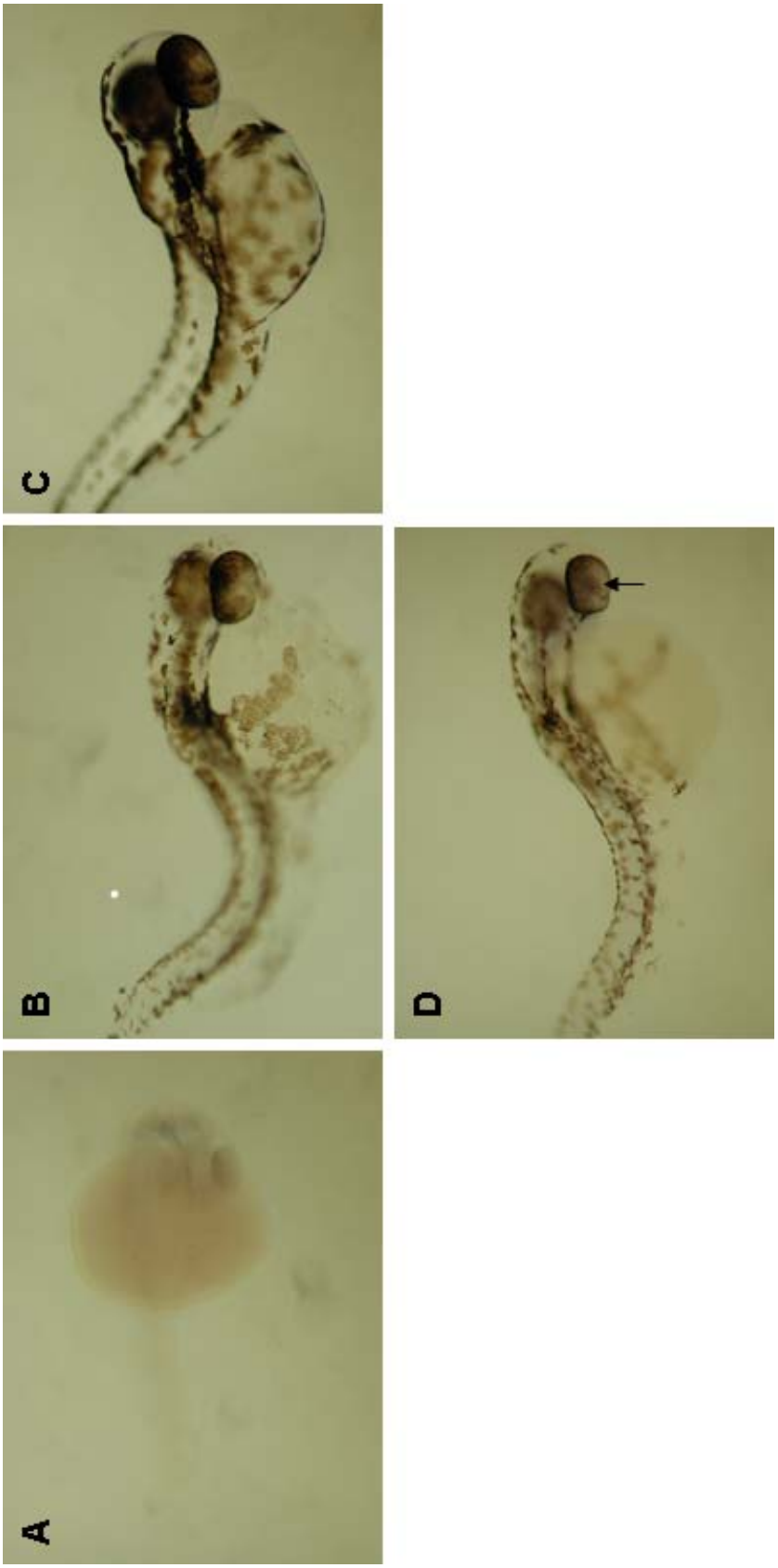
Whole mount *in situ* hybridization, using ssDNA for *hsf4* expression, was performed in embryos fixed at 24, 36, and 72 hours. A faint staining throughout the neural tissue was seen after 4.5 hours of colour reaction in the 24 hpf stages (Fig. 4.11A). The staining appears to be mainly localized in the brain region of the embryos. One *in situ* hybridization experiment performed at 36 hpf did produce embryos with colour visible in the lens of the eye although the mRNA staining was limited to only a few embryos in this experiment (Fig. 4.11D). No staining was evident in the lenses at any other developmental stage analysed (Fig. 4.11B-C). A large amount of dark pigment surrounding the lens can be observed in the eyes at 36 and 72 hpf due to the normal pigmentation that occurs in the eyes at these developmental stages.

In summary, while one experiment suggested that *hsf4* mRNA was present in the lens of the zebrafish embryo at 36 hpf, this was not seen consistently. As well the result could not be repeated in subsequent experiments. These results suggest that *hsf4* mRNA expression is lower than the detection limit for the *in situ* hybridization technique used in these experiments.

4.4.2 Validation of qPCR methodology for analysis of heat shock factor mRNA expression

The results in section 4.4.1 suggested that the levels of *hsf4* mRNA transcripts were too low to be detected by the *in situ* hybridization method. Therefore qPCR, a more sensitive method of detecting mRNA levels, was utilized to detect mRNA for all three HSFs in adult zebrafish tissues as well as in embryonic tissues from several developmental stages during the first four days of development. Initial experiments were performed to validate this method to analyze genes involved in the heat shock response. The validity of two candidate reference genes was also confirmed.

Figure 4.11 Detection of *hsf4* expression in zebrafish embryos by whole mount *in situ* hybridization. Zebrafish embryos fixed at 24 hpf (**A**), 36 hpf (**B, D**) and 72 hpf (**C**) were hybridized with a DIG labeled ssDNA probe specific for *hsf4*. Embryos were screened for the presence of DIG using an Eclipse E600 photomicroscope. Arrow shows colour that was detected in the lens of the eye (**D**).

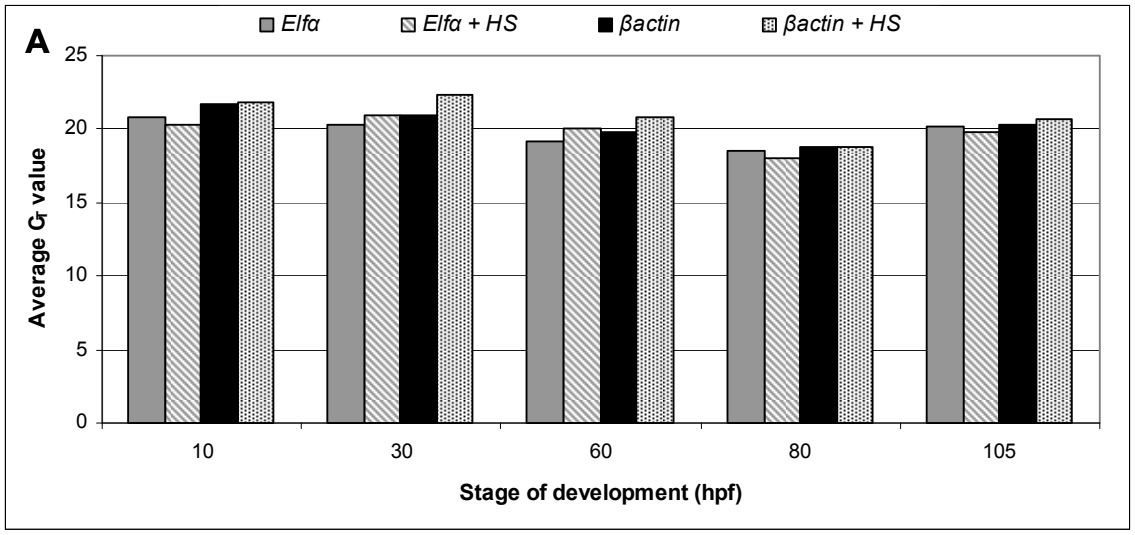

C
B
D
A

4.4.2.1 Determining reference gene stability

Reference genes are genes that have been determined to have relatively stable expression levels across experimental treatments. These genes are typically used to normalize the expression data for genes of interest generated from experiments using qPCR. The genes for *elongation factor alpha* (E*lfα*) and *beta-actin* (*βactin*) were selected as the best choice for qPCR reference genes based on previous literature that analysed several zebrafish genes for use as possible reference genes (Tang et al. 2007; McCurley et al. 2008). In order to validate that these genes were suitable reference genes for the current work, they were analysed by qPCR using cDNA prepared from a series of adult zebrafish tissues and from embryos at different developmental stages. The average C_T values were plotted in Fig. 4.12 to compare the expression of these genes. Both *Elfα* and *βactin* have a similar expression pattern within embryo treatments (Fig. 4.12A) and in each tissue (Fig. 4.12B) with the exception that *βactin* appears to be higher than *Elfα* in liver and squeezed eggs.

The range of mRNA levels for each gene were also similar between tissues, developmental time points and HS treatments in embryos with all the C_T values for both genes in all tissues falling between 16 and 24 (Table 4.5). A small amount of variation can be seen across tissues and embryonic stages which is consistent with observations from other research groups (Tang et al. 2007; McCurley et al. 2008). In adult tissues the C_T values observed for the ovaries and testis were lower those observed for the other tissues. The C_T values for all adult tissues were similar in expression range, mean and percent coefficient of variation (% CV) for both *Elfα* and *βactin* (Table 4.5). The C_T values in embryonic tissue across all developmental stages were also similar in expression range, mean and % CV for the two genes. There was also no difference observed for C_T values of either gene in embryos that were treated with or without heat shock. It can be concluded from these results that both *Elfα* and *βactin* are appropriate reference genes to use in the analysis presented in this thesis. Due to the similarity in the expression patterns of the two genes, only *βactin* was used to normalize gene expression data in qPCR experiments.

Figure 4.12 Analysis of *βactin* and *Elfα* expression in zebrafish adult tissues and developing embryos for use as qPCR reference genes. Average critical threshold values (C_T) for both genes were compared in (**A**) zebrafish embryos from several developmental stages ranging from 10 hpf -105 hpf both with (n=5) or without (n=5) heat shock and (**B**) adult zebrafish tissues; where EggsD refers to a group of eggs dissected from ovarian follicles at various stages of maturity and EggsS refers to mature eggs that were squeezed from a live female (n=4 for each tissue).


A
Elfα
Elfα + HS
βactin
βactin + HS
Average C_T value
25
20
15
10
5
0
10
30
60
80
105
Stage of development (hpf)

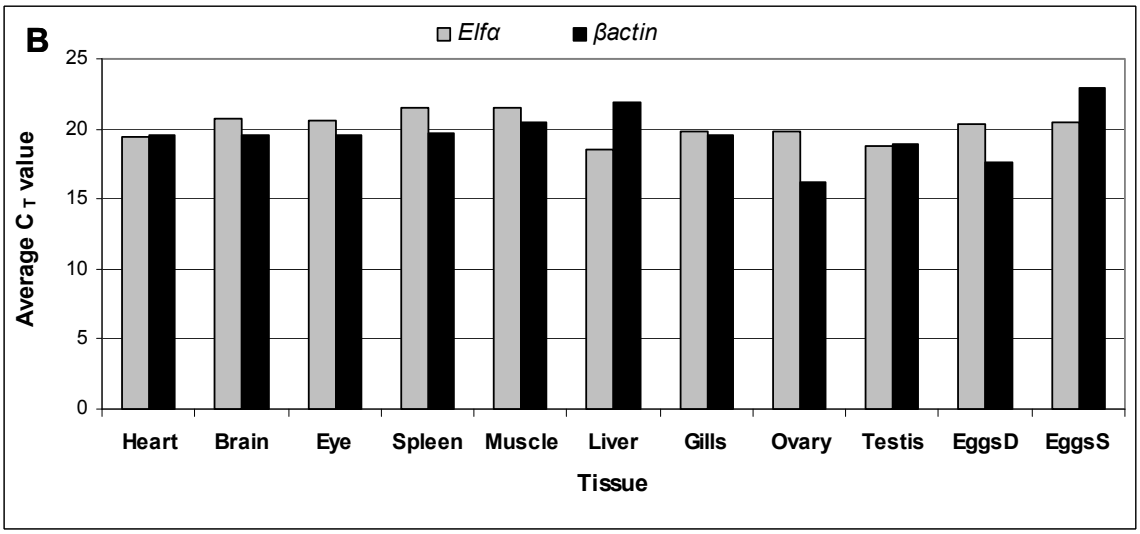

B
Elfα
βactin
Average C T value
25
20
15
10
5
0
Heart
Brain
Eye
Spleen
Muscle
Liver
Gills
Ovary
Testis
EggsD
EggsS
Tissue

Table 4.5 Consistency of *βactin* and E*lfα* mRNA expression in zebrafish adult tissues and developing embryos. Analysis of C_T values obtained by qPCR analysis of each potential reference gene was performed to determine which gene had a more consistent level of expression. The range and mean of C_T values was calculated for each of *βactin* and E*lfα* expression from C_T values obtained for all adult tissues (n=44) and for embryos from all developmental stages both with (n=25) and without heat shock (n=25) to determine the consistency of mRNA expression. The percent coefficient of variation (%CV) was also calculated for each gene in each group as a measurement of variation of mRNA expression.

	Elfα			***βactin***		
	Range of C_T values	Mean C_T	% CV	Range of C_T values	Mean C_T	% CV
Adult Tissues	16.77-23.79	20.30	4.69	15.97-24.53	19.63	9.32
Embryos	17.16-23.12	19.83	4.74	17.18-23.86	20.31	5.67
Embryos after HS	16.57-22.66	19.83	5.39	17.48-23.18	20.86	6.44

4.4.2.2 Analysis of *hsp70* expression as an internal control

Hsp70 is a chaperone that functions to prevent protein aggregation in the cells of an organism under non-stress conditions but also in response to a variety of cellular stresses including heat shock. The *hsp70* gene is expressed in a wide range of tissues and is upregulated dramatically in response to cellular stresses including heat shock. It was therefore exploited as an internal control in qPCR experiments examining tissue specificity and embryo development.

The expression level of the *hsp70* gene was determined in eleven different adult tissues. In all experiments, expression levels of *hsp70* were normalized to *βactin* (section 3.10.2). Relative differences in *hsp70* expression between adult tissues were determined by comparing the normalized *hsp70* expression levels from each tissue to levels of expression in the heart where relative expression of *hsp70* was set equal to 1 (section 3.10.2). Any significant difference between relative expression levels was calculated by the REST analytical software at a significance level of $p \leq 0.05$. Expression of *hsp70* was detected in all eleven adult tissues (Fig. 4.13A). While some variation in *hsp70* expression was observed in different tissues, only mature unfertilized eggs, squeezed from females, had significantly more *hsp70* expression than any of the other tissues.

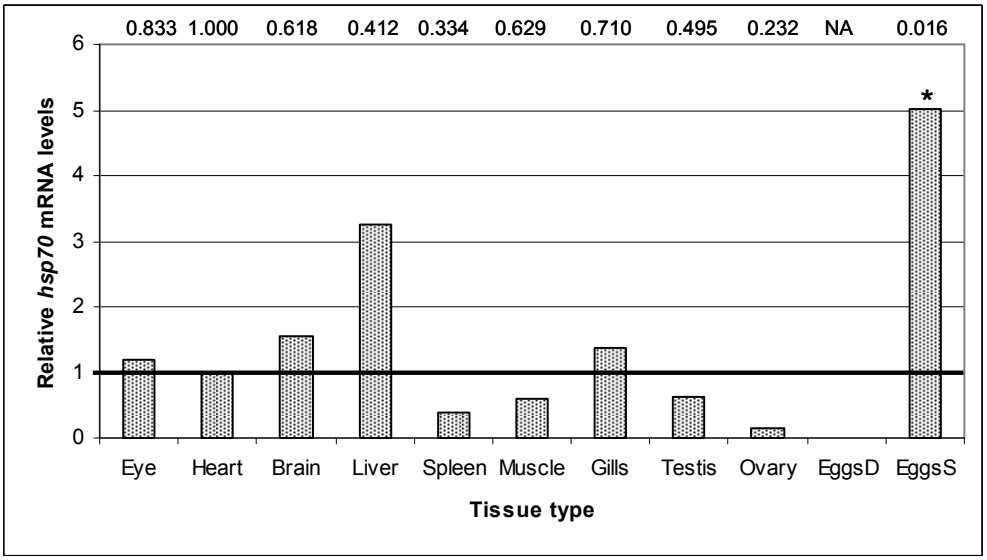
The expression level of *hsp70* in zebrafish embryos at five different developmental stages, both in the presence and absence of heat shock was also determined by qPCR. The five developmental stages assayed span from late gastrulation (10 hpf), through the early stages of organogenesis including eye development (30 hpf), embryo hatching and later organogenesis (60 hpf), young larvae (80 hpf) and larvae that are beginning to feed (105 hpf). Relative differences in *hsp70* expression between embryos at these developmental stages in the absence of heat shock were determined by comparing the normalized *hsp70* expression levels at each time point to levels of expression at 105 hpf where *hsp70* relative expression was set equal to 1. Any significant difference between relative expression levels was calculated by the REST analytical software at a significance level of $p \leq 0.05$. While the data presented in the graph give the appearance of variation in *hsp70* expression across development, the mRNA levels are not significantly different between developmental time points (Fig. 4.13B).

The expression levels of *hsp70* in developing embryos that had been exposed to a 37^{o}C heat shock were also determined. For this analysis, normalized *hsp70* expression levels in heat shocked embryos from each time point were compared relative to the expression levels determined in non-heat shocked embryos at the corresponding time point in which *hsp70* expression was set equal to 1. Any significant difference between relative expression levels was calculated by the REST analytical software at a significance level of $p \leq 0.05$. Expression of *hsp70* was significantly increased in heat shocked embryos at all developmental time points except 30 hpf (Fig. 4.13C). However, the relative expression of *hsp70* in heat shocked embryos at 30 hpf had significance at $p = 0.064$ which is defined as showing a trend to increased expression (section 3.10.2).

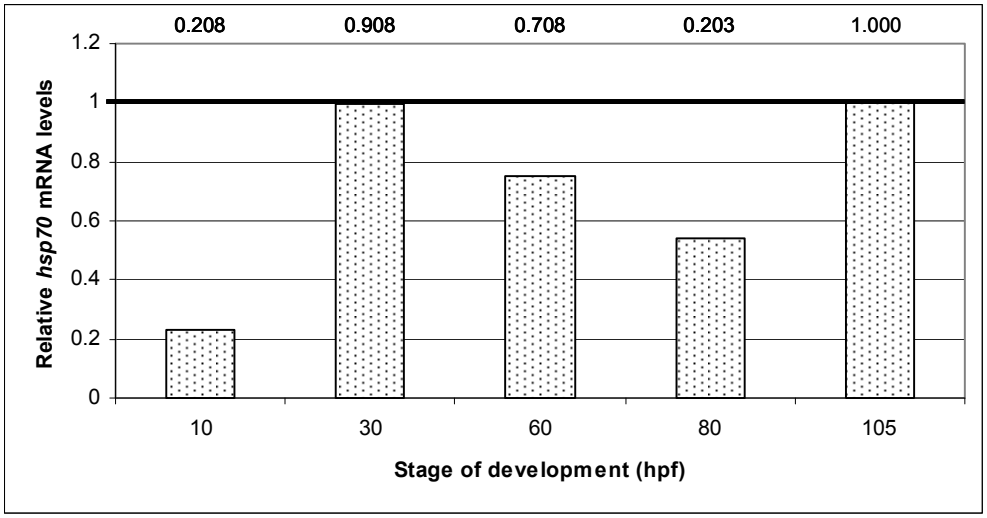
The similarity of *hsp70* expression in zebrafish adult tissues and the increase in *hsp70* expression in embryos exposed to heat determined by qPCR analysis was consistent with expected results (Lindquist 1988; Krone et al. 1997, 2003; Lele et al. 1997). These results confirm that qPCR is a valid method to measure genes involved in the heat shock response. The results also confirm that the method used to heat shock embryos in these experiments was appropriate.

Figure 4.13 Analysis of *hsp70* mRNA by qPCR in zebrafish adult tissues and embryos at different developmental stages. All qPCR mRNA measurements were normalized to the *βactin* reference gene. In all panels a dark line highlights mRNA value set equal to 1 that was used to calculate relative expression values. Significant differences (*) were determined for all experiments at $p \leq 0.05$ using REST analytical software (n=3). Actual probability values for relative differences calculated by REST are given across the top of the graph for each sample analysed. **(A)** Changes in *hsp70* mRNA levels in adult tissues relative to mRNA levels in the heart (heart=1). The EggsD label refers to a group of eggs dissected from ovarian follicles at various stages of maturity while the EggsS label refers to mature eggs that were squeezed from a live female. **(B)** Changes in *hsp70* mRNA levels at each developmental time point relative to 105hpf (105hpf =1). No significant difference from mRNA levels at 105 hpf was observed at $p \leq 0.05$ as calculated by REST. **(C)** Change in *hsp70* mRNA levels after heat shock (HS) for each developmental time point relative to embryos without heat shock (no HS=1 for each time point).

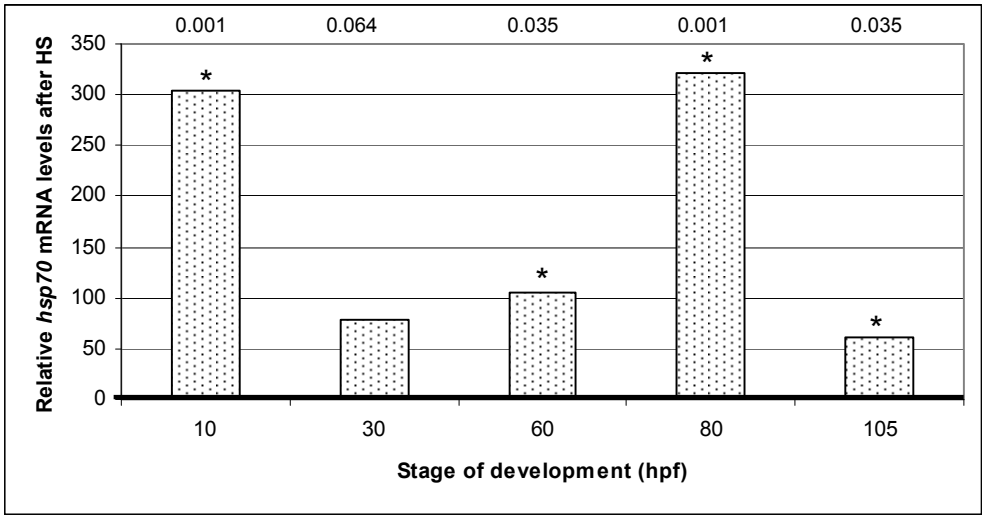
A


0.833 1.000 0.618 0.412 0.334 0.629 0.710 0.495 0.232 NA 0.016
Relative hsp70 mRNA levels
0
1
2
3
4
5
6
*
Eye
Heart
Brain
Liver
Spleen
Muscle
Gills
Testis
Ovary
EggsD
EggsS
Tissue type

B


0.208 0.908 0.708 0.203 1.000
Relative hsp70 mRNA levels
0
0.2
0.4
0.6
0.8
1
1.2
10
30
60
80
105
Stage of development (hpf)

C


0.001 0.064 0.035 0.001 0.035
Relative hsp70 mRNA levels after HS
0
50
100
150
200
250
300
350
*
*
*
*
10
30
60
80
105
Stage of development (hpf)

4.4.3 Analysis of *hsf4* mRNA expression

4.4.3.1 qPCR analysis of *hsf4* expression in adult tissues

The expression of the two isofoms of the *hsf4* gene in mammals has previously been shown to occur primarily in the lens of the eye (Fujimoto et al. 2004, 2008). The identification and subsequent confirmation of the presence of two separate transcripts of the *hsf4* mRNA in zebrafish (section 4.1.1) led me to more fully characterize the expression of this gene in zebrafish tissues. The expression of both *hsf4a* and *hsf4b* was determined by measuring mRNA abundance in eleven tissues from adult zebrafish using qPCR as described in section 3.10. Initial experiments were performed using primers that amplified a region common to both *hsf4* transcripts (Fig. 4.1A) and amplification was quantified using the SYBR green fluorogenic DNA binding dye (Fig. 4.14A).

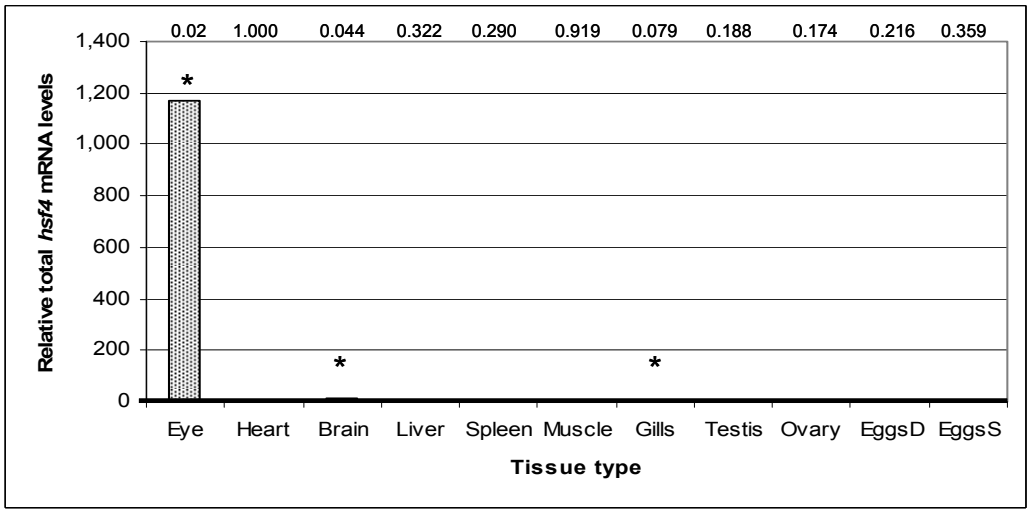
All measurements of *hsf4* mRNA were normalized to *βactin* mRNA values (section 3.10.2). Relative differences in *hsf4* expression between adult tissues were determined by comparing normalized *hsf4* values from each tissue to normalized values from heart tissue where relative expression of *hsf4* was set equal to 1 (section 3.10.2). Any significant difference between relative expression levels was calculated by the REST analytical software at a significance level of $p \leq 0.05$. Results shown in Fig. 4.14A indicate that total *hsf4* expression was detected in all tissues analyzed but was most abundant in the eye. Only the eye and brain had significantly more mRNA levels than the heart while the gill tissue had a trend to decreased levels compared to the heart (section 3.10.2).

A second set of analyses were performed to determine the expression level for each of *hsf4a* and *hsf4b* using primer sets that were specific for each transcript of *hsf4* (Fig. 4.1A). These experiments also employed a more sensitive fluorogenic DNA binding dye called EvaGreen®.

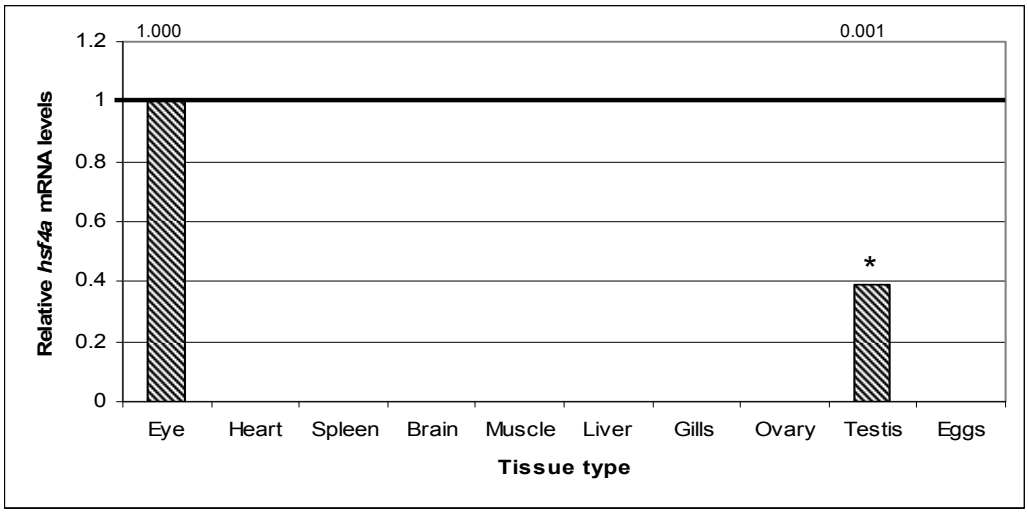
Hsf4a mRNA was also observed predominantly in the eye (Fig. 4.14B). However, unlike *hsf4b* which had a basal level of expression detected in all tissues, *hsf4a* could only be detected in the eye and testis. For this reason, relative expression of *hsf4a* could not be calculated compared to expression in the heart tissue. Therefore, relative expression of *hsf4a* was calculated compared to eye tissue with relative expression in the eye set to 1 (section 3.10.2). The RNA levels of *hsf4a* detected in the testis were significantly less than levels detected in the eye.

Figure 4.14 Analysis of *hsf4* mRNA by qPCR in zebrafish adult tissues. All qPCR mRNA measurements were normalized to the *βactin* reference gene. In panels a dark line highlights values set equal to 1 used to calculate relative expression. Quantitation of total *hsf4* mRNA levels (both splice variants) was determined by qPCR using the SYBR green fluorogenic dye and results are shown in panel A. Quantitation of *hsf4a* and *hsf4b* mRNA independent of each other was determined by qPCR using the EvaGreen® fluorogenic dye and results are shown in panels B and C. Values were expressed relative to the heart in panel A and C. Values were expressed relative to the eye in panel B. Significant differences (*) were determined for all experiments at $p \leq 0.05$ using REST analytical software. Actual probability values for relative differences calculated by REST are given across the top of the graph for each sample analysed. **(A)** Changes in total *hsf4* mRNA levels relative to the heart in adult tissues (heart=1; n=4). **(B)** Changes in *hsf4a* mRNA levels relative to the eye in adult tissues (eye=1; n=3). **(C)** Changes in total *hsf4b* mRNA levels relative to the heart in adult tissues (heart=1; n=3). The EggsD label refers to a group of eggs dissected from ovarian follicles at various stages of maturity while the EggsS label refers to mature eggs that were squeezed from a live female.

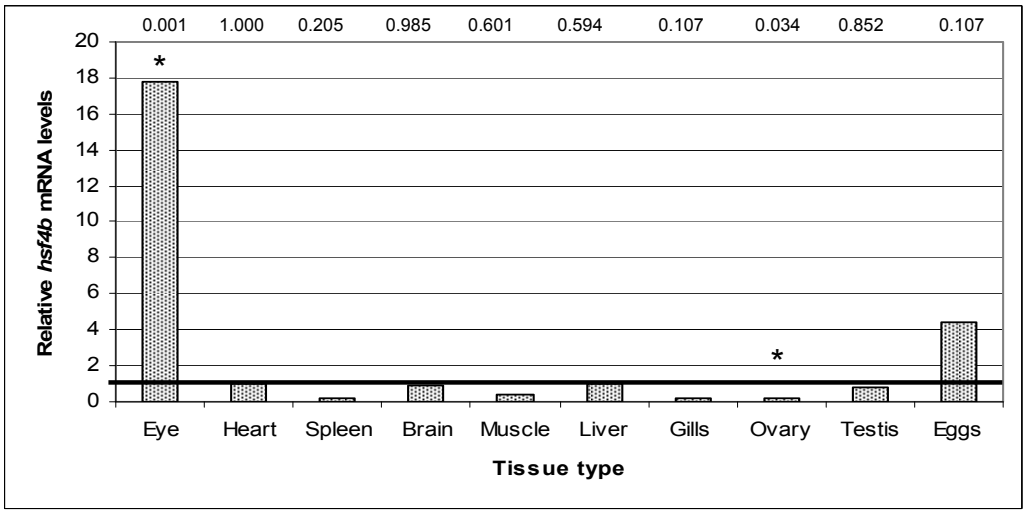
A


0.02 1.000 0.044 0.322 0.290 0.919 0.079 0.188 0.174 0.216 0.359
1,400
1,200
1,000
800
600
400
200
0
Relative total hsf4 mRNA levels
Eye Heart Brain Liver Spleen Muscle Gills Testis Ovary EggsD EggsS
Tissue type

B


1.000 0.001
1.2
1
0.8
0.6
0.4
0.2
0
Relative hsf4a mRNA levels
Eye Heart Spleen Brain Muscle Liver Gills Ovary Testis Eggs
Tissue type

C


0.001 1.000 0.205 0.985 0.601 0.594 0.107 0.034 0.852 0.107
20
18
16
14
12
10
8
6
4
2
0
Relative hsf4b mRNA levels
Eye Heart Spleen Brain Muscle Liver Gills Ovary Testis Eggs
Tissue type

Expression of *hsf4b* was detected in all tissues but was expressed predominantly in the eye with expression levels calculated to be 18 times more than the heart (Fig. 4.14C). There was a trend to increased expression of *hsf4b* in squeezed eggs ($p = 0.107$). The ovary had significantly less *hsf4b* mRNA than the heart while mRNA levels in the gills showed a trend to decreased mRNA levels ($p = 0.107$) (section 3.10.2).

The relative amount of *hsf4b* mRNA was compared to that of *hsf4a* by expressing the *hs4b* mRNA levels in each tissue relative to *hsf4a* levels where relative expression of *hsf4a* in each tissue was set equal to 1. However, the above comparison of values for two genes obtained using a different primer set for each gene can only be considered an approximation as no data correlating C_T values to actual transcript numbers was determined. RNA levels of *hsf4b* were calculatd to be 185 times that of the *hsfa* transcript suggesting that *hsf4b* is the dominant transcript in the eye (data not shown).

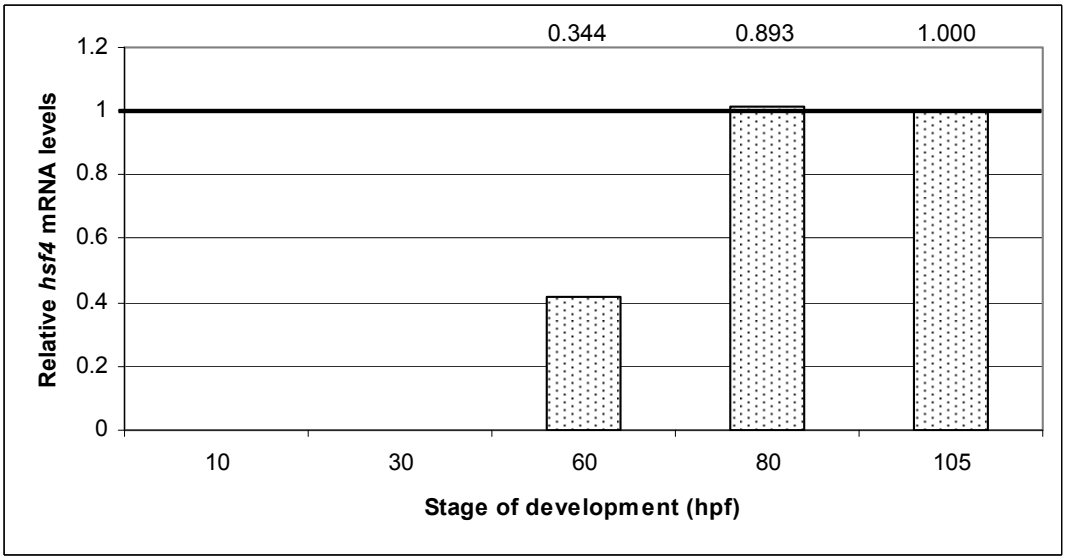
4.4.3.2 qPCR analysis of *hsf4* expression during embryo development

The expression of both transcripts of *hsf4* was further characterized in developing zebrafish embryos at five different developmental time points. Again, initial experiments were performed using primers that amplified a region common to both *hsf4* transcripts (Fig. 4.1A) and amplification was quantified using the SYBR green fluorogenic DNA binding dye (Fig. 4.15). All measurements of *hsf4* mRNA were normalized to *βactin* mRNA values (section 3.10.2). Relative differences in *hsf4* expression between developmental time points were determined by comparing normalized *hsf4* values from embryos at each time point to normalized values from embryos at 105 hpf where relative expression of *hsf4* was set equal to 1 (section 3.10.2). Any significant difference between relative expression levels was calculated by the REST analytical software at a significance level of $p \leq 0.05$ (section 3.10.2).

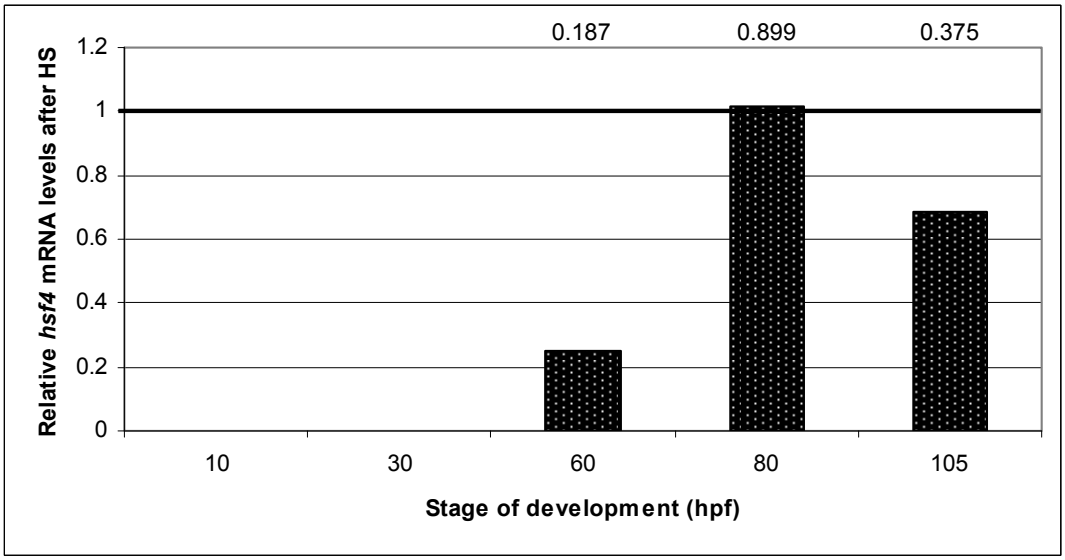
The levels of *hsf4* mRNA could only be confidently quantified in embryos 60 hpf and older both in the absence (Fig. 4.15A) and presence of HS (Fig. 4.15B). Measurements of *hsf4* at 10 and 30 hpf were all below the limit of detection for the qPCR assay and were therefore assigned a value of 0. The levels of *hsf4* detected at 60, and 80 hpf were not significantly different from 105 hpf as determined by REST (Fig. 4.15A). The *hsf4* mRNA levels were also not significantly different in embryos that had been exposed to HS compared to embryos that had no HS at all time points as determined by REST (Fig. 4.15A).

Figure 4.15 Analysis of total *hsf4* mRNA levels by qPCR in zebrafish embryos at different developmental stages. All qPCR mRNA measurements were normalized to the *βactin* reference gene. In panels a dark line highlights values set equal to 1 used to calculate relative expression. Significant differences (*) were determined for all experiments at $p \leq 0.05$ using REST analytical software (n=3). Actual probability values for relative differences calculated by REST are given across the top of the graph for each sample analysed. **(A)** Changes in *hsf4* mRNA levels at each developmental time point relative to 105hpf (105hpf =1). **(B)** Change in *hsf4* mRNA levels after heat shock (HS) for each developmental time point relative to embryos without heat shock (no HS=1 for each time point).

A


0.344
0.893
1.000
1.2
1
0.8
0.6
0.4
0.2
0
Relative hsf4 mRNA levels
10
30
60
80
105
Stage of development (hpf)

B

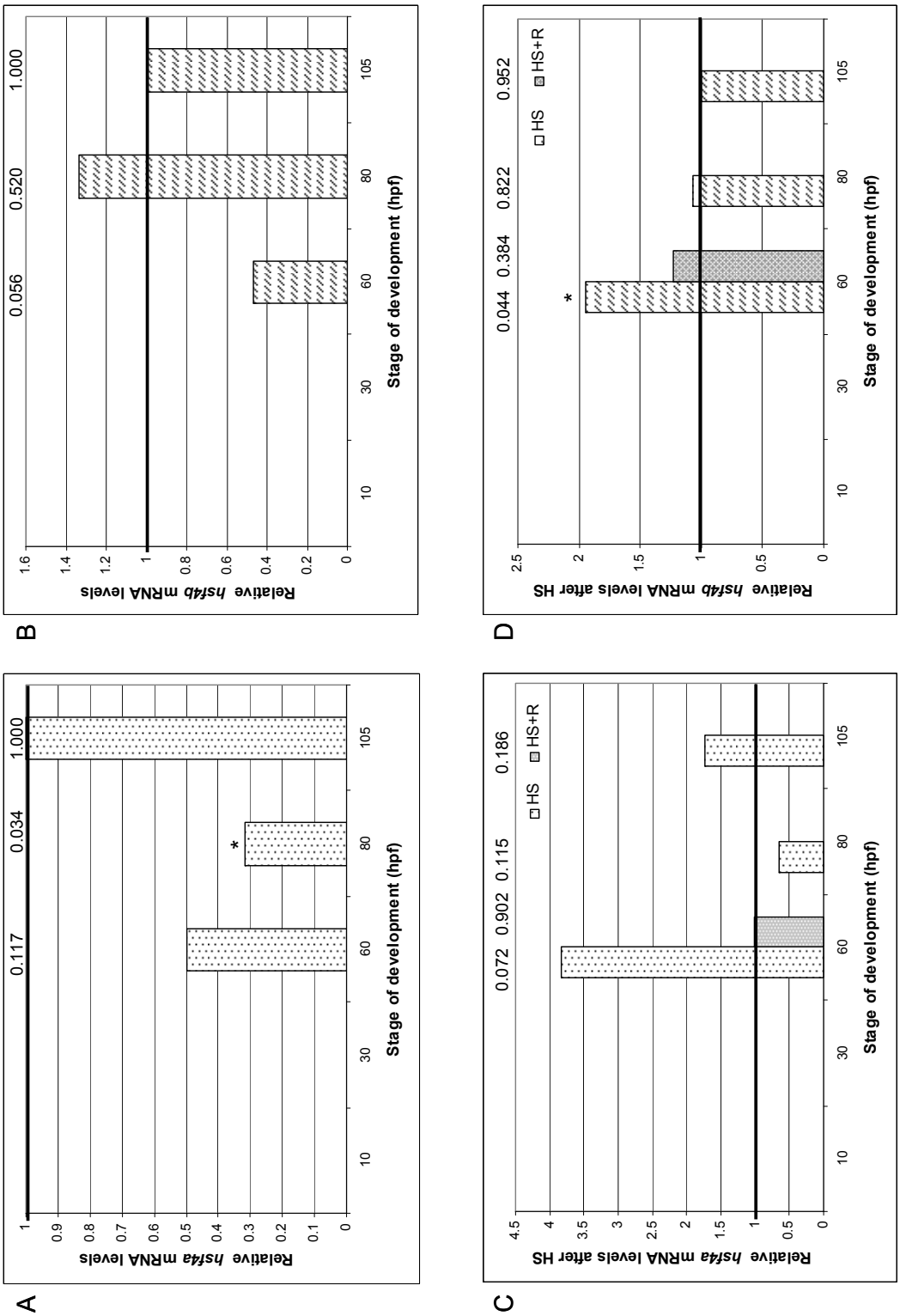

0.187
0.899
0.375
1.2
1
0.8
0.6
0.4
0.2
0
Relative hsf4 mRNA levels after HS
10
30
60
80
105
Stage of development (hpf)

A second set of analyses were performed to determine the expression level for each of *hsf4a* and *hsf4b* transcripts using primer sets specific for each transcript (Fig. 4.1A). These experiments used a more sensitive fluorogenic DNA binding dye called Sso-Fast™ EvaGreen® Supermix. qPCR analysis with this dye demonstrated that both *hsf4a* and *hsf4b* were below the limit of detection at 10 and 30 hpf, similar to previous results from combined *hsf4* transcripts (Fig. 4.16). Levels of mRNA could only be confidently quantified in embryos at 60 hpf or older. The levels of *hsf4a* detected at 60 hpf had a trend to lower expression levels than 105 hpf ($p = 0.115$) and at 80 hpf the mRNA levels were significantly lower than that at 105 hpf (Fig. 4.16A). The mRNA levels for *hsf4b* at 60 hpf ($p = 0.056$) also showed a trend to lower levels than 105 hpf while levels at 80 hpf were not significantly different than those at 105 hpf (Fig. 4.16B).

The effect of heat shock on expression of *hsf4* was also determined at the above developmental time points (Fig. 4.16C and D). For these experiments, relative expression of each *hsf4* transcript in embryos exposed to heat shock was determined by comparing normalized *hsf4a* or *hsf4b* expression in heat shocked embryos to normalized expression of *hsf4a* or *hsf4b* in embryos without heat shock at each time point (section 3.10.2). There was a trend to increased *hsf4a* and a significant increase in *hsf4b* mRNA levels at 60 hpf after heat shock. At 80 hpf there was a trend to decreased *hsf4a* mRNA levels but no significant difference in *hsf4b* levels after heat shock. No significant difference in mRNA levels of either transcript was observed at 105 hpf. The observation that heat shock appeared to increase both *hsf4* transcripts at 60 hpf led to an investigation as to whether mRNA levels would remain elevated after a period of recovery. The mRNA returned to levels similar to those observed in non-heat shocked embryos for both *hsf4a* (Fig. 4.16C) and *hsf4b* (Fig. 4.16D) after a 15 hour recovery at 28°C.

The relative levels of both transcripts were compared to each other. However, again, the comparison of values for two genes obtained using a different primer set for each gene can only be considered an approximation as no data correlating C_T values to actual transcript numbers was determined. In all developmental stages where *hsf4* was detected, the *hsf4b* transcript was calculated to be more abundant than the *hsf4a* transcripts having 9, 30 and 7 fold more mRNA at 60, 80 and 105 hpf respectively. While the higher *hsf4b*:*hsf4a* ratio of mRNA levels was the same in both embryos and adult tissues, fold differences between *hsf4b* and *hsf4a* levels in embryos was much lower than the 185 fold difference observed in adult tissue (section 4.4.3.1).

Figure 4.16 Analysis of mRNA levels for both *hsf4* transcripts by qPCR in zebrafish adult tissues and embryos at different developmental stages All qPCR mRNA measurements were normalized to the *βactin* reference gene. In panels a dark line highlights values set equal to 1 used to calculate relative expression. Significant differences (*) were determined for all experiments at $p \leq 0.05$ using REST analytical software (n=3 for each stage). Actual probability values for relative differences calculated by REST are given across the top of the graph for each sample analysed. **(A)** Changes in *hsf4a* mRNA levels at each developmental time point relative to 105 hpf (105 hpf =1). **(B)** Changes in *hsf4b* mRNA levels at each developmental time point relative to 105 hpf (105 hpf =1). **(C)** Change in *hsf4a* mRNA levels after heat shock (HS) for each developmental time point and after HS plus a 15 hour recovery (HS+R) at 60 hpf, relative to embryos without heat shock (no HS=1 for each time point). **(D)** Change in *hsf4b* mRNA levels after HS and after HS+R at 60 hpf for each developmental time point relative to embryos without heat shock (no HS=1 for each time point).


A
B
C
D
Relative hsf4a mRNA levels
Relative hsf4b mRNA levels
Relative hsf4a mRNA levels after HS
Relative hsf4b mRNA levels after HS
Stage of development (hpf)
HS
HS+R
0.117
0.034
1.000
0.056
0.520
1.000
0.072
0.902
0.115
0.186
0.044
0.384
0.822
0.952
*

4.4.4 Analysis of *hsf1* mRNA expression

4.4.4.1 qPCR analysis of *hsf1* expression in adult tissues

The expression pattern of *hsf1* was also analysed in tissues from adult zebrafish. Previous studies had identified that two transcripts of *hsf1*, *a* and *b*, were expressed in liver, gonads and gills from adult zebrafish (Råbergh et al. 2000). For the purposes of this thesis a more extensive analysis of *hsf1* expression was performed in eleven different tissues using qPCR with SYBR green fluorogenic DNA binding dye as described in section 3.10. One set of primers utilized in these assays was designed to specifically amplify cDNA from the *hsf1b* transcript while a second set of primers was designed to amplify cDNA from both the *hsf1a+b* transcripts (Fig. 4.17).

All measurements of *hsf1a+b* and *hsf1b* mRNA were normalized to *βactin* mRNA values (section 3.10.2). Relative differences in *hsf1* expression between adult tissues were determined by comparing normalized *hsf1a+b* and *hsf1b* values from each tissue to normalized values from heart tissue where relative expression of *hsf1a+b* or *hsf1b* was set equal to 1 (section 3.10.2). Any significant difference between relative expression levels was calculated by the REST analytical software at a significance level of $p \leq 0.05$.

Results shown in Fig. 4.18A indicate that *hsf1a+b* expression was detected in all tissues analyzed. Levels of expression in the brain, liver and testis were similar to those observed in the heart. Expression of *hsf1a+b* was significantly lower in the spleen, muscle, gills, ovary and eggs dissected from the ovary with a trend to lower levels in the eye. Mature eggs squeezed from the ovary had more than 10 fold *hsf1a+b* mRNA levels than the heart and was the only tissue to have significantly more *hsf1a+b* mRNA levels relative to the heart.

The results from the analysis of only *hsf1b* expression are shown in Fig. 4.18B. These results demonstrate that the mRNA levels of *hsf1b* relative to the heart were nearly identical to those of the combined *hsf1a+b*. The exception was that the liver had a trend to lower levels than the heart where *hsf1a+b* levels in the liver had been similar to the heart. The mRNA levels of *hsf1b* in the mature squeezed eggs was also significantly more than the heart but only a six fold increase was seen with the *hsf1b* transcript alone.

Figure 4.17 Diagramatic representation of primer binding sites for expression analysis of zebrafish *hsf1* transcripts. The zebrafish *hsf1* gene contains 13 exons. The sequence from all exons is present in the *hsf1b* transcript while the *hsf4a* transcript contains sequence from all but exon 10 (in grey). Analysis of mRNA levels for the combined *hsf1a+b* transcripts was performed by amplifying a region in exon 9 which is common to both transcripts (indicated by a black arrow). Analysis of the mRNA levels for only the *hsf1b* transcript was performed by amplifying a region in exon 10 specific for that transcript (indicated by a grey arrow).

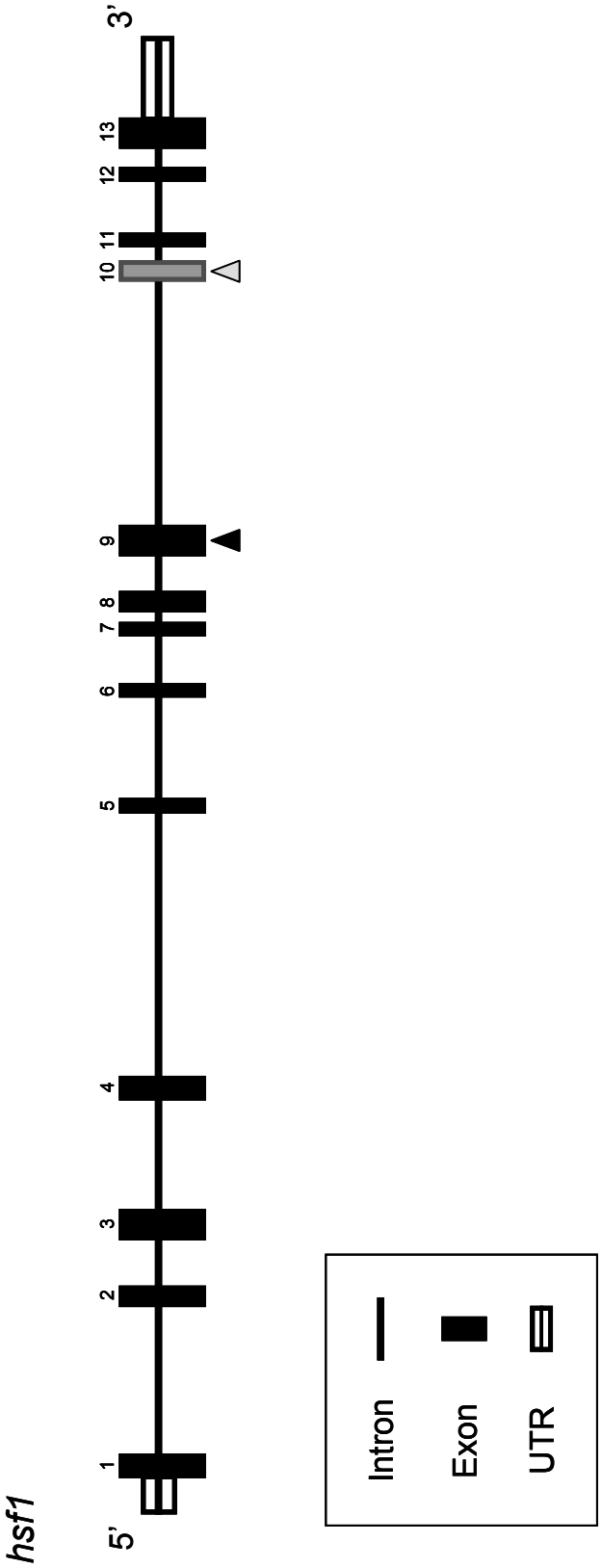
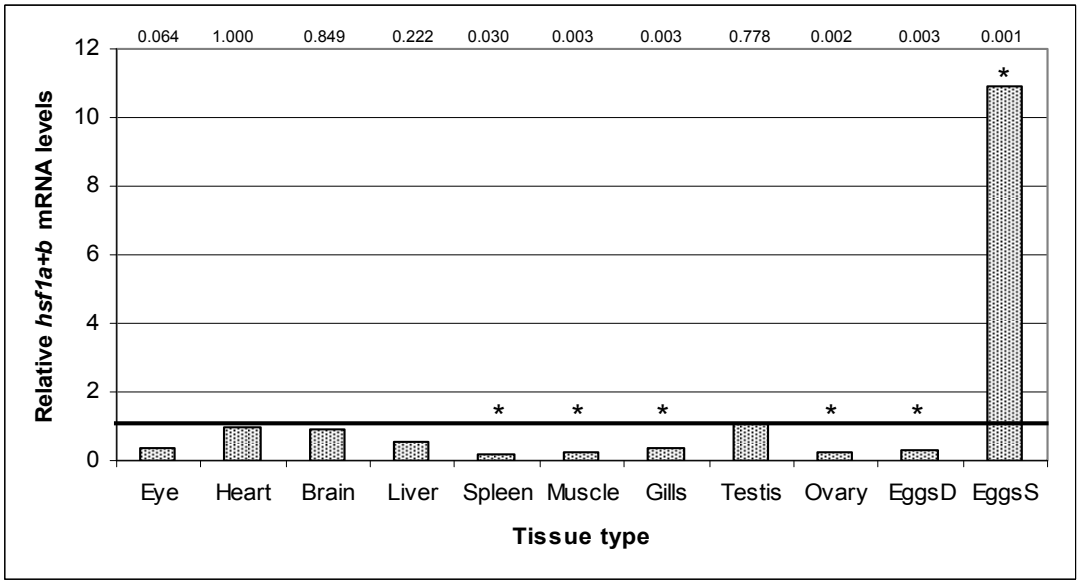
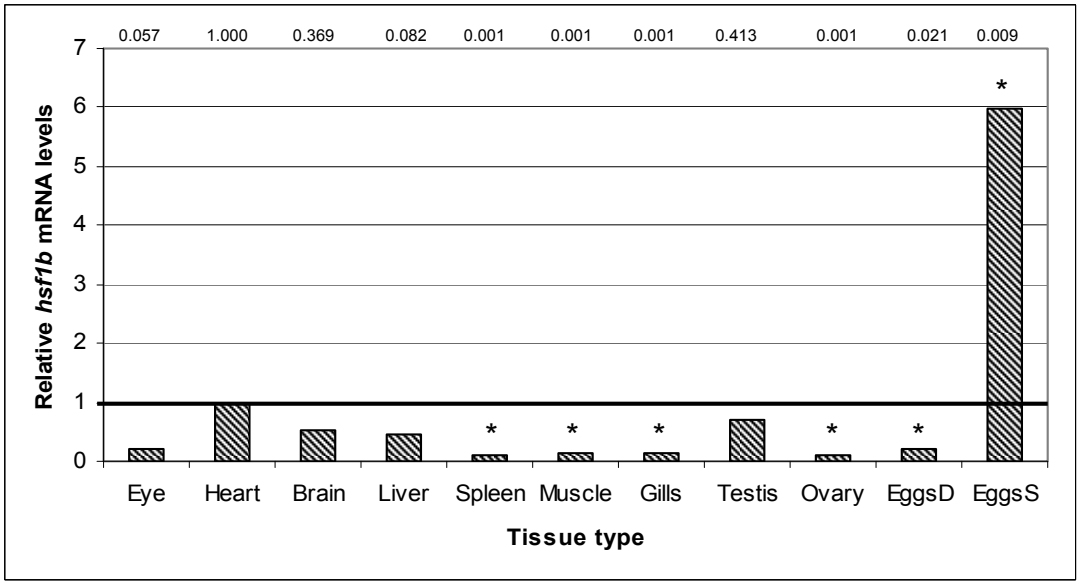

hsf1
5'
3'
1
2
3
4
5
6
7
8
9
10
11
12
13
Intron
Exon
UTR

Figure 4.18 Analysis of *hsf1* mRNA by qPCR in zebrafish adult tissues. All qPCR mRNA measurements were normalized to the *βactin* reference gene. In panels a dark line highlights values set equal to 1 used to calculate relative expression. Values were expressed relative to the heart in panel A and B. Significant differences (*) were determined for all experiments at $p \leq 0.05$ using REST analytical software (n=3). Actual probability values for relative differences calculated by REST are given across the top of the graph for each sample analysed. **(A)** Changes in total *hsf1a+b* mRNA levels relative to the heart in adult tissues (heart=1). **(B)** Changes in total *hsf1b* mRNA levels relative to the heart in adult tissues (heart=1). The EggsD label refers to a group of eggs dissected from ovarian follicles at various stages of maturity while the EggsS label refers to mature eggs that were squeezed from a live female.

A


0.064
1.000
0.849
0.222
0.030
0.003
0.003
0.778
0.002
0.003
0.001
12
10
8
6
4
2
0
Relative hsf1a+b mRNA levels
Eye
Heart
Brain
Liver
Spleen
Muscle
Gills
Testis
Ovary
EggsD
EggsS
Tissue type

B


0.057
1.000
0.369
0.082
0.001
0.001
0.001
0.413
0.001
0.021
0.009
7
6
5
4
3
2
1
0
Relative hsf1b mRNA levels
Eye
Heart
Brain
Liver
Spleen
Muscle
Gills
Testis
Ovary
EggsD
EggsS
Tissue type

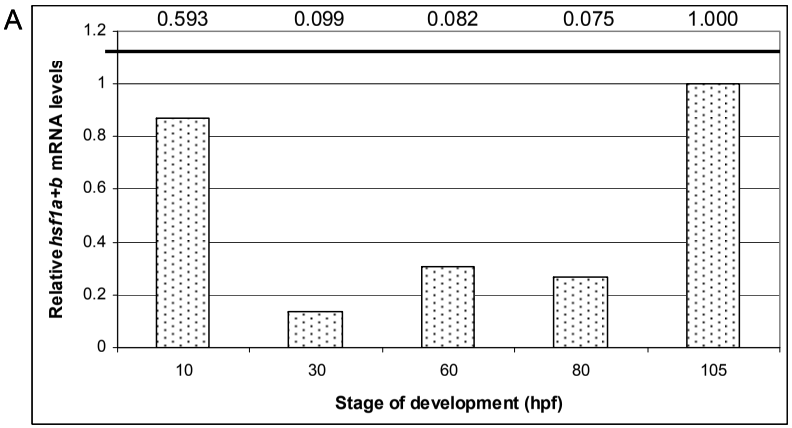
4.4.4.2 qPCR analysis of *hsf1* expression during embryo development

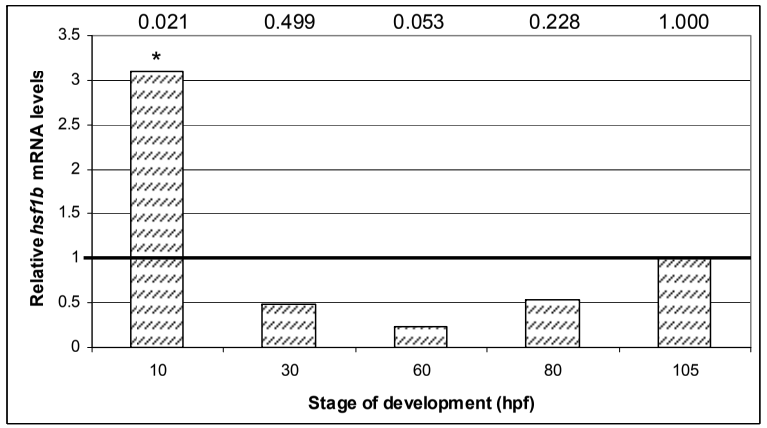
The expression of *hsf1a* has previously been localized mainly to the head region of zebrafish embryos at 36, 60 and 84 hpf by *in situ* hybridization (Yeh et al. 2006). In this study, the expression of the *hsf1* transcripts was further characterized in zebrafish embryos at five different developmental time points by qPCR using SYBR green fluorogenic DNA binding dye as described in section 3.10. The same primers used to characterize both *hsf1* transcripts in adult tissue (section 4.4.4.1 and Fig. 4.17) were used to detect mRNA levels in the embryos. All measurements of *hsf1* mRNA were normalized to *βactin* mRNA values (section 3.10.2). Relative differences in *hsf1a+b* and *hsf1b* expression between developmental time points were determined by comparing normalized *hsf1* values from embryos at each time point to normalized values from embryos at 105 hpf where relative expression of *hsf1a+b* or *hsf1b* was set equal to 1 (section 3.10.2). Any significant difference between relative expression levels was calculated by the REST analytical software at a significance level of $p \leq 0.05$ (section 3.10.2).

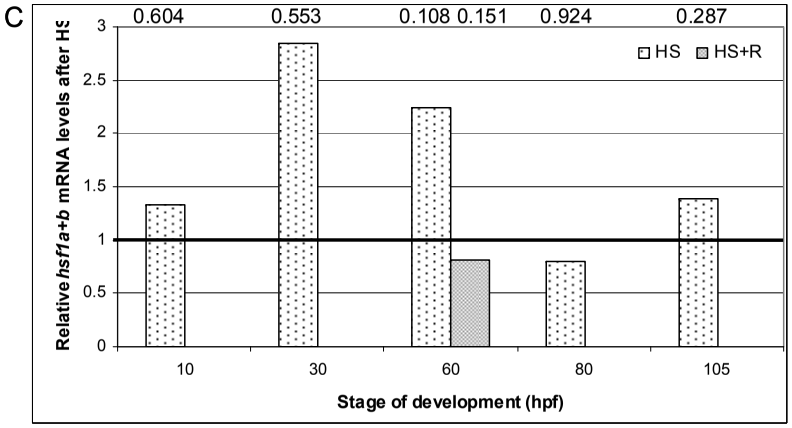
The results in Fig. 4.19A indicate that there was no significant difference observed between *hsf1a+b* levels at 10 hpf and 105 hpf. There was, however, a trend to lower levels of *hsf1a+b* in embryos at time 30, 60 and 80 hpf when compared to 105 hpf. Taken together these results suggest that there are higher levels of mRNA at 10 hpf, followed by a decrease at 30 hpf where levels remain relatively constant until an increase in expression at 105 hpf. The analysis of the *hsf1b* transcript alone indicated that there was a significantly higher level of mRNA at 10 hpf than at 105 hpf. There was no significant difference between *hsf1b* levels in embryos at 30 and 80 hpf when compared to 105 hpf but there was a trend to lower expression levels at 60 hpf (Fig. 4.19B). Taken together the analysis of just the *hsf1b* expression suggests that at 10 hpf there are significantly higher levels of mRNA than any other time point, followed by a decrease at 30 hpf, a further decrease at 60 hpf, followed by increased levels of *hsf1b* at each of the 80 and 105 hpf time points.

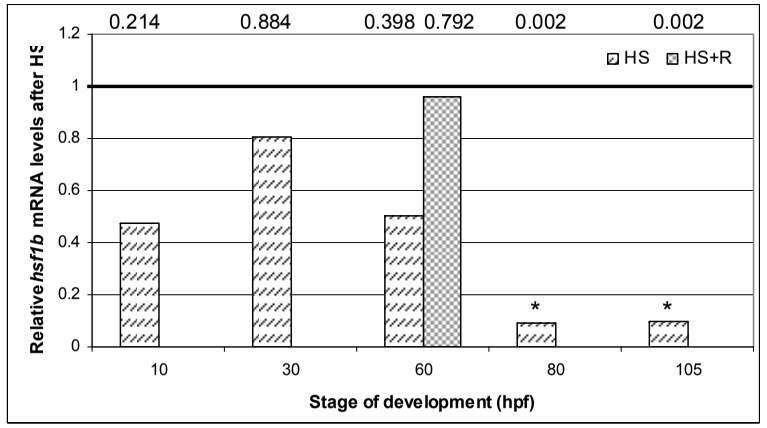
The effect of heat shock on expression of the *hsf1* transcripts was also determined at the above mentioned developmental time points (Fig. 4.19C and D). For these experiments, the relative expression of the *hsf1* transcripts in embryos exposed to heat shock was determined by comparing normalized *hsf1a+b* or *hsf1b* expression in heat shocked embryos to normalized expression of *hsf1a+b* or *hsf1b* in embryos without heat shock at each time point (section 3.10.2).

Figure 4.19 Analysis of mRNA levels for both *hsf1* transcripts by qPCR in zebrafish adult tissues and embryos at different developmental stages. All qPCR mRNA measurements were normalized to the *βactin* reference gene. In panels a dark line highlights values set equal to 1 used to calculate relative expression. Significant differences (*) were determined for all experiments at $p \leq 0.05$ using REST analytical software. Actual probability values for relative differences calculated by REST are given across the top of the graph for each sample analysed. **(A)** Changes in *hsf1a+b* mRNA levels at each developmental time point relative to 105 hpf (105 hpf =1) (n=3). **(B)** Changes in *hsf1b* mRNA levels at each developmental time point relative to 105 hpf (105 hpf =1) (n=5). **(C)** Change in *hsf1a+b* mRNA levels after heat shock (HS) for each developmental time point and after HS plus a 15 hour recovery (HS+R) at 60 hpf, relative to embryos without heat shock (no HS=1 for each time point) (n=3). **(D)** Change in *hsf1b* mRNA levels after HS for each developmental time point and after HS+R at 60 hpf, relative to embryos without heat shock (no HS=1 for each time point) (n=5).


A
0.593
0.099
0.082
0.075
1.000
Relative hsf1a+b mRNA levels
1.2
1
0.8
0.6
0.4
0.2
0
10
30
60
80
105
Stage of development (hpf)


B
0.021
0.499
0.053
0.228
1.000
*
Relative hsf1b mRNA levels
3.5
3
2.5
2
1.5
1
0.5
0
10
30
60
80
105
Stage of development (hpf)


C
0.604
0.553
0.108
0.151
0.924
0.287
HS
HS+R
Relative hsf1a+b mRNA levels after HS
3
2.5
2
1.5
1
0.5
0
10
30
60
80
105
Stage of development (hpf)


D
0.214
0.884
0.398
0.792
0.002
0.002
HS
HS+R
*
*
Relative hsf1b mRNA levels after HS
1.2
1
0.8
0.6
0.4
0.2
0
10
30
60
80
105
Stage of development (hpf)

The results in Fig. 4.19C indicate that heat shock treatment did not significantly change *hsf1a+b* mRNA levels at most of the developmental stages assayed. The one exception was at 60 hpf where there was a trend to increased mRNA levels after heat shock. The analysis of just the *hsf1b* transcript indicated that there was no significant change in mRNA levels after heat shock at 10, 30 and 60 hpf. There was however, a significant decrease in *hsf1b* mRNA levels at 80 and 105 hpf. The observation that heat shock caused a trend to increase the expression of the combined *hsf1a+b* transcripts at 60 hpf led to an investigation as to whether the mRNA levels would stay elevated after a period of recovery. Levels of mRNA appeared to return to levels similar to those observed in non-heat shocked embryos for both *hsf1a+b* (Fig. 4.19C) and *hsf1b* alone (Fig. 4.19D) after a 15 hour period of recovery at 28°C.

The amount of combined *hsf1a+b* mRNA was determined relative to that of just *hsf1b* at each developmental time point where expression of *hsf1b* at each developmental stage was set equal to 1. However, once again, the comparison of values for two genes obtained using a different primer set for each gene can only be considered an approximation as no data correlating C_T values to actual transcript numbers was determined. At most developmental stages *hsf1b* was calculated to be either significantly lower or showed a trend to lower mRNA levels than the combined *hsf1a+b*. The relative levels of *hsf1b* mRNA at 30 hpf were approximately the same as those determined at 10 hpf (0.4 fold and and 0.1 fold lower in embryos without and with HS respectively). However, statistical analysis by REST determined these values to be not significantly different from *hsf1a+b* levels at the 30 hpf time point.

4.4.5 Analysis of *hsf2* mRNA expression

4.4.5.1 qPCR analysis of *hsf2* expression in adult tissues

The expression pattern of *hsf2* was also analysed in eleven tissues from adult zebrafish by qPCR using SYBR green fluorogenic DNA binding dye as described in section 3.10. All measurements of *hsf2* mRNA were normalized to *βactin* mRNA values (section 3.10.2). Relative differences in *hsf2* expression between adult tissues were determined by comparing normalized *hsf2* values from each tissue to normalized values from heart tissue where relative expression of *hsf2* was set equal to 1 (section 3.10.2). Any significant difference between relative expression levels was calculated by the REST analytical software at a significance level of $p \leq 0.05$. The results of the analysis of *hsf2* mRNA levels in adult tissues

demonstrate that *hsf2* mRNA could be detected in all eleven tissues (Fig. 4.20A). There are significantly higher levels of *hsf2* in mature, squeezed eggs than there are in any other adult tissue. As well, there is a trend to higher *hsf2* levels in liver than in the heart. The levels of *hsf2* mRNA were similar to the heart in all other adult tissues assayed.

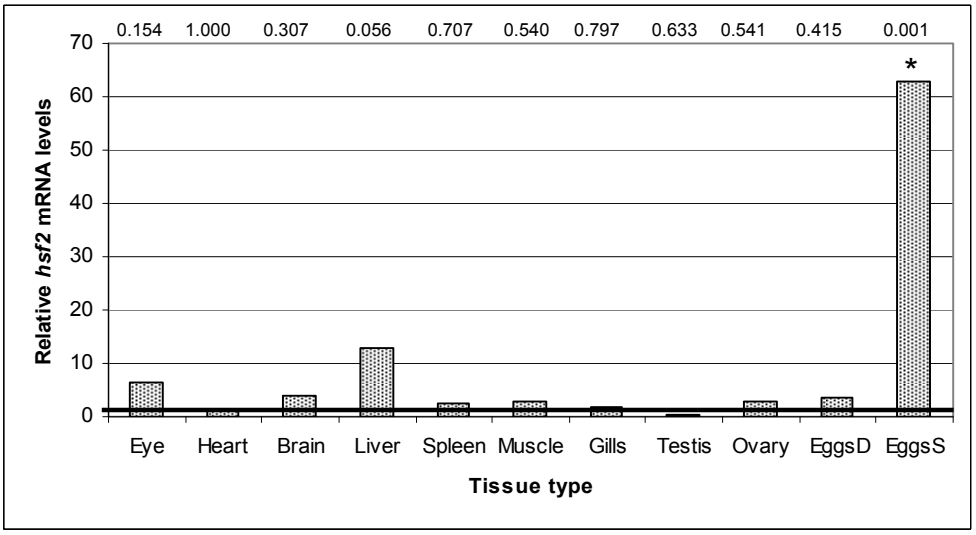
4.4.5.2 qPCR analysis of *hsf2* expression during embryo development

The expression of *hsf2* has previously been localized mainly to the head region of zebrafish embryos at 36, 60 and 84 hpf by *in situ* hybridization (Yeh et al. 2006). This study further characterized the expression of *hsf2* in developing zebrafish embryos at five different developmental time points by qPCR analysis utilizing the SYBR green fluorogenic DNA binding dye as per section 3.10. All measurements of *hsf2* mRNA were normalized to *βactin* mRNA values (section 3.10.2). Relative differences in *hsf2* expression between developmental time points were determined by comparing normalized *hsf2* values from embryos at each time point to normalized values from embryos at 105 hpf where relative expression of *hsf2* was set equal to 1 (section 3.10.2). Any significant difference between relative expression levels was calculated by the REST analytical software at a significance level of $p \leq 0.05$ (section 3.10.2). The results of the analysis of *hsf2* expression in developing embryos indicates there was a trend to higher *hsf2* mRNA in 10 hpf embryos but that *hsf2* mRNA levels in all other stages were not significantly different from those observed for 105 hpf (Fig. 4.20B).

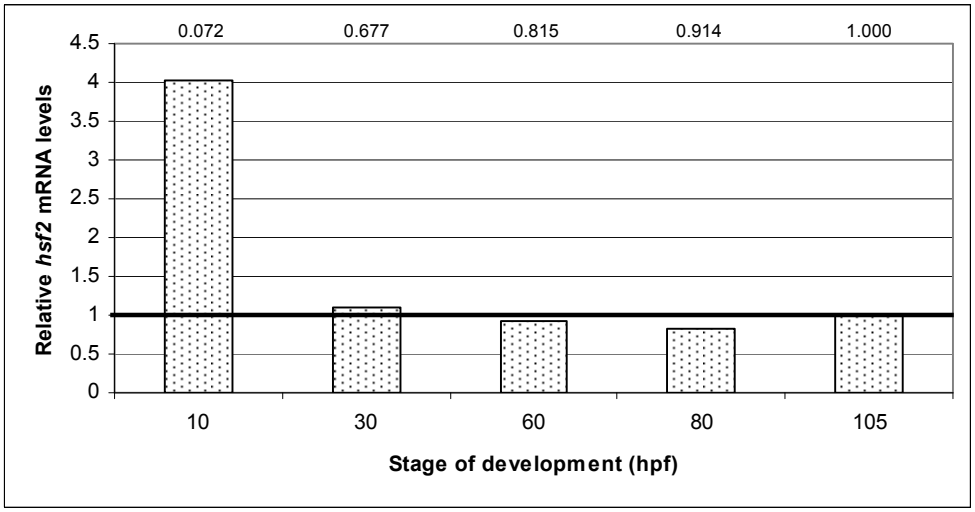
The effect of heat shock on expression of *hsf2* was also determined at the same five developmental time points (Fig. 4.20C). For these experiments, the relative expression of *hsf2* in embryos exposed to heat shock was determined by comparing normalized *hsf2* expression in heat shocked embryos to normalized *hsf2* expression in embryos without heat shock at each time point (section 3.10.2). The results in Fig. 4.20C indicate that there was no significant difference in *hsf2* mRNA levels after heat shock at any of the developmental stages assayed.

Figure 4.20 Analysis of *hsf2* mRNA by qPCR in zebrafish adult tissues and embryos at different developmental stages All qPCR mRNA measurements were normalized to the *βactin* reference gene. In panels a dark line highlights values set equal to 1 used to calculate relative expression. Significant differences (*) were determined for all experiments at $p \leq 0.05$ using REST analytical software (n=3). Actual probability values for relative differences calculated by REST are given across the top of the graph for each sample analysed. **(A)** Changes in *hsf2* mRNA levels in adult tissues relative to mRNA levels in the heart (heart=1). The EggsD label refers to a group of eggs dissected from ovarian follicles at various stages of maturity while the EggsS label refers to mature eggs that were squeezed from a live female. **(B)** Changes in *hsf2* mRNA levels at each developmental time point relative to 105 hpf (105 hpf =1). **(C)** Change in *hsf2* mRNA levels after heat shock (HS) for each developmental time point relative to embryos without heat shock (no HS=1 for each time point).

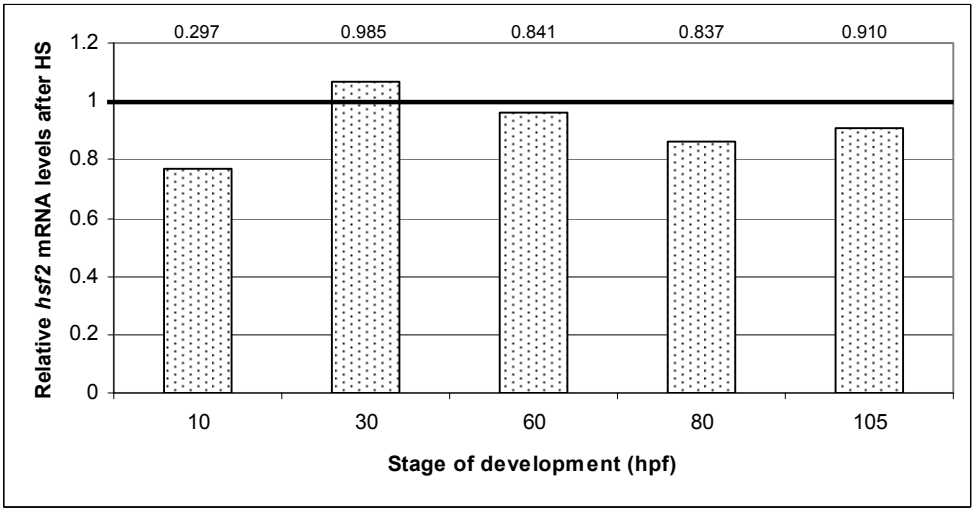
A


0.154 1.000 0.307 0.056 0.707 0.540 0.797 0.633 0.541 0.415 0.001
*
Relative hsf2 mRNA levels
70
60
50
40
30
20
10
0
Eye Heart Brain Liver Spleen Muscle Gills Testis Ovary EggsD EggsS
Tissue type

B


0.072 0.677 0.815 0.914 1.000
Relative hsf2 mRNA levels
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
10 30 60 80 105
Stage of development (hpf)

C


0.297 0.985 0.841 0.837 0.910
Relative hsf2 mRNA levels after HS
1.2
1
0.8
0.6
0.4
0.2
0
10 30 60 80 105
Stage of development (hpf)

4.5 Morpholino knockdown of *hsf4* and *hsf1* expression

4.5.1 Effect of *hsf4* knockdown on the phenotype of zebrafish embryos

The effectiveness of morpholino modified antisense oligonucleotide (MO) microinjection in fertilized zebrafish eggs to knock down *hsp70* and *hsf1* expression has been previously established (Evans et al. 2005, 2007). Thus, a similar approach was used to examine *hsf4* knockdown in zebrafish embryos using MO that were tagged with a 3'-carboxyfluoroscene molecule and targeted to the mRNA sequence spanning the *hsf4* start codon. Preliminary experiments designed to determine if *hsf4*-MO had an effect on embryo phenotype were performed by research technician Nicole Sylvain and are summarized in Table 4.6. Knockdown of *hsf4* expression was performed by injecting embryos with several concentrations of *hsf4*-MO as well as the *hsf4* 5bpmm-MO control ranging from 1.73-5 µg/µL and were compared to Danieau injected and uninjected controls at 48 hours after injection as described in section 3.12. Injected embryos were analysed to determine the maximum *hsf4*-MO concentration for use in subsequent experiments that would produce an observable phenotype without causing significant non-specific phenotypic abnormalities (Evans et al. 2007).

Embryos injected with MO at a concentration of 5 µg/µL had a lower survival rate and higher number of non-specific abnormalities than in embryos injected with lower concentrations (Table 4.6). It was determined that a MO concentration of 3.46 µg/µL was the best for use in subsequent knockdown experiments.

Table 4.6 Efficiency of *hsf4* gene knockdown using four concentrations of modified morpholino oligonucleotides (MO). Total numbers of embryos injected with *hsf4*-MO and *hsf4 5bpmm*-MO as well as uninjected and danieau injected controls as well as the percentage of embryo survival after 48 hours are listed. Efficiency of the MO injections in surviving embryos was determined by calculating percentage of embryos that displayed strong fluoresceine fluorescenc. The number of embryos containing flourescein as well as displaying non-specific deformities was also tallied at each MO concentration to determine the optimal MO concentration for injection.

	Controls		MO Injected							
	Uninjected	Danieu Injected	5.00 µg/µL		3.46 µg/µL		2.595 µg/µL		1.73 µg/µL	
			HSF4	5bpmm	HSF4	5bpmm	HSF4	5bpmm	HSF4	5bpmm
Embryos Injected (#)	461	201	48	37	266	200	267	247	459	309
Survival at 48 hours (%)	79	63	48	38	69	74	70	63	61	75
Deformities (%)	0	0	26	14	9	6	8	8	5	7

A small number of embryos injected with *hsf4*-MO (3.46 µg/µL) were observed to have a moderate small eye phenotype. However, approximately the same number of embryos in the *hsf4* 5bpmm-MO injected group also had a moderate small eye phenotype. No small eye was observed in the uninjected or Danieau injected embryos. Therefore, further analysis of this phenotype in MO injected embryos was performed by measuring the diameter of the lens and surface area of the eye and expressing those values as a ratio to body length. No significant change in ratios between body to eye length, body length to lens diameter or eye surface area to body length were observed at 48 hpf (Table 4.7), indicating that MO microinjection had no visible morphological impact on lens and eye development. The data were interpreted to suggest that the small number of moderate small eye phenotypes that were observed were due to an overall decrease in embryo size that can be attributed to the presence of MO in the embryos and not due to a knockdown of *hsf4* expression.

In addition, no cataract development was observed in injected embryos after 96 hpf as has previously been observed for *hsf1*-MO-injected embryos (Evans et al. 2007). Concentration of MO in injected embryos is known to fall below effective levels beginning at 50-60 hpf (Evans et al. 2005). This fact combined with the observation that *hsf4* mRNA expression is below detectable levels prior to 60 hpf could explain the lack of cataract development in *hsf4* knockdown embryos.

Table 4.7 Knockdown of *hsf4* expression does not result in an altered eye phenotype. Embryos were injected with danieau solution (3.14 nL), *hsf4* morpholino oligonucleotides (MO) or *hsf4* 5bpmm MO (3.46 µg/µL) and several measurements were taken of the embryo eye and body at 48 hours after injection. Ratios of body length to eye length, eye surface area and lens diameter were calculated for each of the treatment groups.

3.46 µg/µL MO	Embryos (#)	Body :Eye length		Body length:lens diameter		Eye length:eye diameter		Eye Surface Area:body length	
		Average	SD	Average	SD	Average	SD	Average	SD
Uninjected	144	12.4	0.73	37.9	3.1	3.07	0.18	4.95	0.74
Danieau Injected	95	12.6	1.1	38.8	2.9	3.09	0.26	4.92	0.77
hsf4-MO	50	12.4	0.94	38.9	3.8	3.17	0.32	5.41	0.57
hsf4 5bpmm-MO	17	12.3	0.65	38.1	3.0	3.10	0.25	5.06	0.38

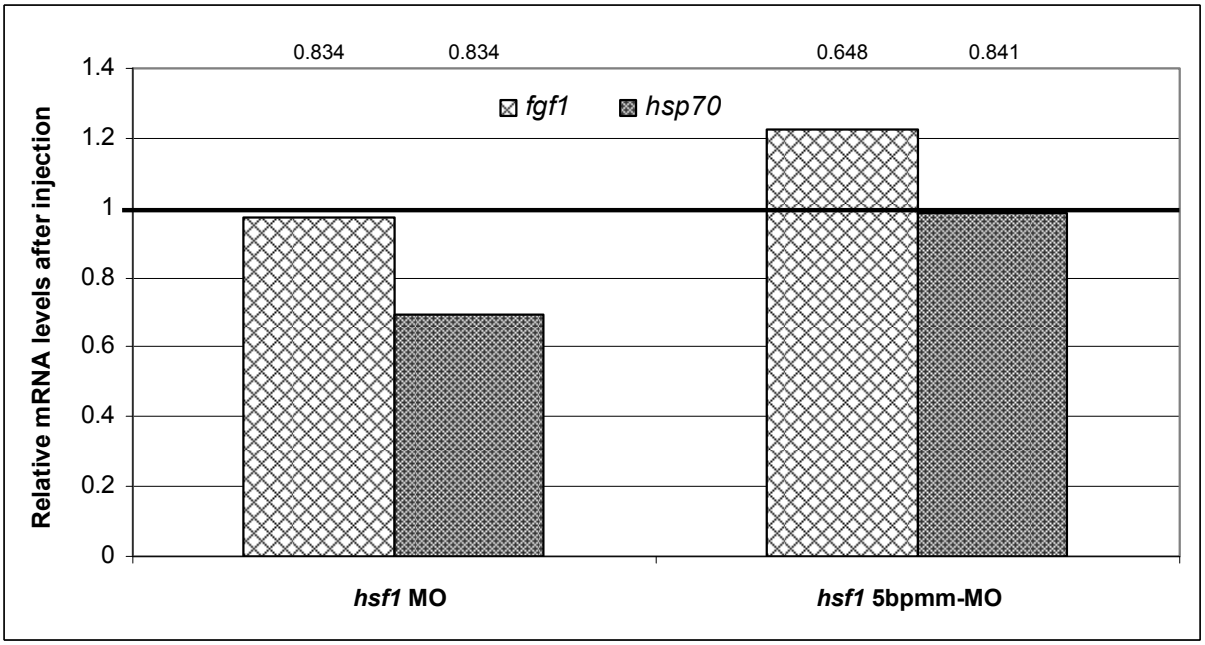
4.5.2 Effect of *hsf1* knockdown on the expression of *hsp70 and fgf1*

Previous studies demonstrated that MO knockdown of *hsf1* or *hsp70* produced a small eye phenotype in zebrafish embryos with a more severe phenotype being observed from the *hsp70* knockdown (Evans et al. 2005, 2007). A knockdown of *hsf1* also results in reduced Hsp70 in developing zebrafish (Evans et al. 2007). The ability of HSF1 to regulate Hsp70 expression in response to several stresses including heat shock is also well established in several species (Pirkalla et al. 2001; Björk and Sistonen 2010). The purpose of these experiments was to validate that a decrease in *hsp70* mRNA, a known target of *hsf1* regulation, resulting from a MO knockdown of *hsf1* in zebrafish embryos could be detected and quantified by qPCR. The effect of *hsf1* knockdown on *fgf1* mRNA levels, a known target of *hsf1* regulation in mammals, was also analysed.

Morpholino sequences used in these experiments were designed to target the ATG translational start site of *hsf1* start codon and also contained a 3'-lissamine tag. Experiments were performed by microinjecting embryos with 3.46 µg/µL *hsf1*-MO and *hsf1* 5bpmm-MO, previously established by Evans et al. 2007, and comparing the results to an uninjected control. Analysis of mRNA levels for *hsp70* and *fgf1*, were determined by qPCR using the EvaGreen® fluorgeneic dye as described in section 3.10. Measurements of mRNA levels from all genes of interest were normalized to *βactin* mRNA values (section 3.10.2). Relative differences in mRNA expression between MO injected and uninjected embryos were determined by comparing normalized values from the genes of interest in injected embryos to the normalized values from uninjected embryos where relative expression of the gene of interest was set equal to 1 (section 3.10.2). Any significant difference between relative expression levels was calculated by the REST analytical software at a significance level of $p \leq 0.05$ (section 3.10.2).

The results of the qPCR analysis shown in Fig. 4.21 demonstrate that no significant change in mRNA levels was observed in embryos injected with the *hsf1*-MO or the control 5bpmm-MO for either *hsp70* or *fgf1*.

Figure 4.21 Analysis of mRNA detected in zebrafish embryos by qPCR after *hsf1* morpholino (MO) knockdown. Zebrafish embryos were injected with either an *hsf1*-MO or *hsf1* 5bpmm-MO control. Expression values for *fgf1* and *hsp70* were determined by qPCR in embryos at 60 hpf. All values were normalized to the *βactin* reference gene. Values for each gene are shown expressed relative to uninjected embryos where the values were set equal to 1 (highlighted by a dark line). No significant difference in mRNA levels was observed at $p \leq 0.05$ as calculated by REST, (n=2). Actual probability values for relative differences calculated by REST are given across the top of the graph for each sample analysed.


0.834
0.834
0.648
0.841
fgf1
hsp70
Relative mRNA levels after injection
1.4
1.2
1
0.8
0.6
0.4
0.2
0
hsf1 MO
hsf1 5bpmm-MO

The results were interpreted to suggest that knockdown of the *hsf1* gene was unsuccessful. The optimal concentration of MO to use in experiments presented here was previously determined using an *hsf1*-MO with the same sequence but tagged with fluorescein instead of lissamine. The MO manufacturer does caution that MOs tagged with lissamine may have lower solubility than MOs containing other end modifications. Therefore, it may be possible that injections performed in this study contained a lower than calculated MO concentration which may be responsible for the lack of *hsf1* knockdown observed in these experiments.

4.5.3 Effect of *hsf4* knockdown on the expression of *hsf1, hsp70* and *fgf1*

Knockdown of *hsf4* using MO failed to produce an observable phenotype (section 4.5.1). Therefore a subsequent set of experiments was performed to identify possible downstream gene targets of the *hsf4* transcription factor that may be affected by *hsf4* knockdown. The effectiveness of MO knockdowns are typically verified by observation of a known phenotype associated with the knockdown, by using Western blot to demonstrate reduced protein levels resulting from MO blocked translation or by a decrease in mRNA or protein levels of a known downstream gene target of the MO. However the *hsf4*-MO knockdown does not produce an observable embryonic phenotype and antibodies designed to detect zebrafish HSF are not able to specifically detect HSF bands from zebrafish embryos (section 4.3.2). These facts combined with the lack of a known target of HSF4 in zebrafish meant that a different method of validating the effectiveness of *hsf4*-MO injections had to be developed. Therefore, these experiments were performed with a second type of MO, termed a splice MO, tagged with 3'-lissamine. Splice MO are designed to bind to a target sequence spanning an intron-exon junction in the target pre-mRNA instead of the ATG translation start site. When the MO binds to the intron-exon junction, it disrupts the nuclear processing of the pre-mRNA typically resulting in the removal of both the exon and the intron instead of just the intron (Draper et al. 2001). This results in mature mRNA products of a reduced length. Analysis of MO effectiveness can then be verified by qPCR using primers that span the knockout region targeted by the MO, as a shorter PCR product will be amplified from the MO injected samples compared to uninjected. As well, the addition of a lissamine tag allowed for visual confirmation that injections were successful and the MO was distributed throughout the injected embryo. Only embryos with uniform fluorescence were selected for subsequent analysis. The splice-MO used in the experiments presented here were

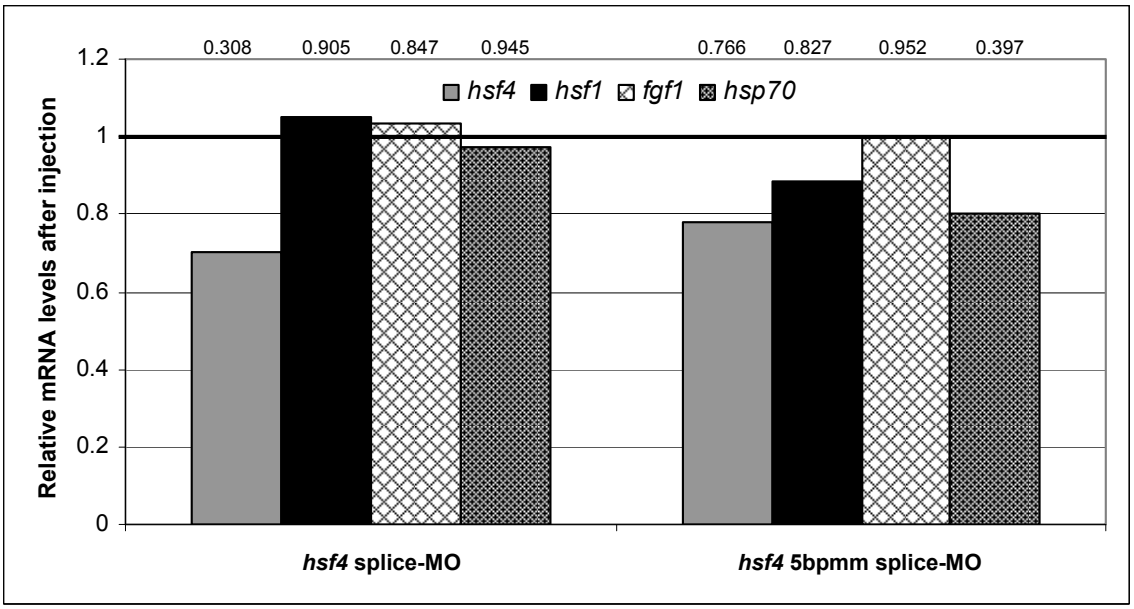
designed to target the *hsf4* mRNA in a way that would cause the 5th exon to be spliced out of the pre-mRNA, resulting in mature mRNA product of reduced length (section 3.12).

Experiments were performed by microinjecting embryos with 3.46 µg/µL *hsf4*-splice MO and *hsf4* 5bpmm-splice MO and comparing the results to uninjected controls. Initially, embryos were screened under a dissecting microscope to determine if a phenotype resulting from the *hsf4* splice-MO could be observed. Several embryos injected with either the MO or 5bpmm appeared to be smaller in size than uninjected controls but no observable phenotype was detected, similar to what was observed in previous knockdown experiments described in section 4.5.1.

Analysis of mRNA levels for *hsf4* and four possible downstream gene targets, *hsf1*, *hsp70* and *fibroblast growth factor 1* (*fgf1*), were determined by qPCR as described in section 3.10. All qPCR reactions were performed using the EvaGreen® fluorogenic dye. Measurements of mRNA levels from all genes of interest were normalized and relative differences in mRNA expression between MO injected embryos and uninjected embryos were determined as described previously for the *hsf1*-MO injections in section 4.5.2. Any significant difference between relative expression levels was calculated by the REST analytical software at a significance level of $p \leq 0.05$ (section 3.10.2).

The results of the qPCR analysis shown in Fig. 4.22 demonstrate that no significant change in mRNA levels for any of the genes of interest was observed in embryos injected with the *hsf4*-splice MO. The same result was observed for embryos injected with *hsf4* 5bpmm-splice MO. The results were interpreted to suggest that gene knockdown with the *hsf4* splice-MO were unsuccessful. It is possible that the lissamine tagged *hsf4*-splice MO also had a lower solubility, resulting in a lower than expected MO concentration similar to that described above for the *hsf1* knockdown experiments (section 4.5.2).

Figure 4.22 Analysis of mRNA detected in zebrafish embryos by qPCR after *hsf4* morpholino (MO) knockdown. Zebrafish embryos were injected with either an *hsf4* splice-MO or *hsf4* 5bpmm splice-MO. Expression values for *hsf4*, *hsf1a+b*, *fgf1* and *hsp70* were determined by qPCR in embryos at 60 hpf. All values were normalized to the *βactin* reference gene. Values for each gene are shown expressed relative to uninjected embryos where the values were set equal to 1 (highlighted by a dark line). No significant difference in mRNA levels was observed at $p \leq 0.05$ as calculated by REST, (n=3). Actual probability values for relative differences calculated by REST are given across the top of the graph for each sample analysed.


0.308
0.905
0.847
0.945
0.766
0.827
0.952
0.397
hsf4
hsf1
fgf1
hsp70
Relative mRNA levels after injection
1.2
1
0.8
0.6
0.4
0.2
0
hsf4 splice-MO
hsf4 5bpmm splice-MO

5.0 DISCUSSION

Our lab has been examining the developmental roles of Hsps and HSFs in zebrafish (*Danio rerio*), a well established developmental model system. Recent studies by our group have shown that regulation of *hsp70* gene expression by HSF1 is important for development of the zebrafish eye (Blechinger et al. 2002; Evans et al. 2005, 2007). Both HSF1 and HSF4 play a role in mammalian lens development (Fujimoto et al. 2004; Min et al. 2004; Smaoui et al. 2004; Mellersh et al. 2006; Fujimoto et al. 2008; Shi et al. 2009; Abane and Mezger 2010), suggesting that the zebrafish *hsf4* identified in an expressed sequence tag (EST) library could play a similar role. The zebrafish genome contains genes encoding at least three heat shock factors; two isoforms of HSF1 (Råbergh et al. 2000), HSF2 (Yeh et al. 2006) and the novel HSF4. The focus of this thesis was to characterize the expression and function of the novel zebrafish HSF4 and determine if it had a role in eye development similar to that observed in mammalian species. As well, tissue specific and developmental expression patterns of *hsf4* were compared to those of *hsf1* and *hsf2* in this study.

Two transcript variants of zebrafish *hsf4* were identified in this study. The first sequence identified from an EST library, *hsf4a*, encodes a severely truncated HSF4 in which the entire carboxy portion of the protein is absent. This is the first time an HSF protein of such a short length has been identified in animals, and it represents a unique subgroup of vertebrate HSFs. The second *hsf4b* transcript, predicted in the Ensembl database, encodes a putative protein more similar in length to other HSFs; however this protein contains an HR-C inhibition domain that is not present in mammalian HSF4 proteins. A comparative *in silico* protein sequence analysis of zebrafish HSF4 to zebrafish HSF1 and HSF2, as well as HSFs from a wide range of organisms was performed. An examination of this analysis as well as DNA binding properties of HSF4 in comparison to other zebrafish HSFs suggests that HSF4 is not functionally equivalent to mammalian HSF4 in many respects and therefore supports the first hypothesis that zebrafish HSF4 represents a novel class of HSFs in vertebrates.

Several experiments were performed to characterize the expression of the three known zebrafish HSFs across a series of developmental time points and in adult tissues. One of the objectives was to characterize HSF protein expression levels in tissues and embryos using custom designed antibodies for zebrafish HSFs. Even though each HSF antibody was able to specifically detect the corresponding HSF protein by Western blot when purified HSF proteins

were used (Fig. 4.7), analysis of proteins from whole zebrafish extracts was inconclusive (Fig. 4.8). For that reason, this discussion of HSF expression in zebrafish embryos and tissues focuses solely on mRNA expression analysis in this thesis. The analysis of HSF mRNA expression demonstrated that both *hsf4* transcripts are expressed in a developmentally regulated manner and are predominantly expressed in the eye. A basal expression level of *hsf1* and *hsf2* could be detected in all tissues and at all developmental stages, however a specific pattern of expression was also observed for each of these genes supporting the second hypothesis that HSFs are expressed in distinct developmental patterns in zebrafish embryos.

While no analysis of protein accumulation could be performed in zebrafish due to lack of specific antibodies, other studies have demonstrated that HSF RNA and protein levels are roughly correlated. Several experiments have shown in various cell and animal models, that heat shock stimulates HSF1 or HSF3 activation primarily by inducing trimerization and DNA binding of constitutively expressed HSFs without changes in RNA and protein levels (Nakai and Morimoto 1983; Sarge et al. 1993; Rabindran et al. 1994; Tanabe et al. 1997; Fujimoto et al. 2010; Björk and Sistonen et al. 2010). Analysis of HSF2 and HSF4 expression has also demonstrated that both mRNA and protein for each HSF exist concurrently in the tissues in which they are expressed (Nakai and Morimoto 1983; Tanabe et al. 1999; Fujimoto et al. 2008). Specifically for HSF2, a correlation between increased mRNA levels and increased DNA binding of the protein has been shown (Björk and Sistonen et al. 2010). Since a high degree of homology is observed between HSF amino acid sequences in all animal species, it is likely that regulatory mechanisms are also similar. Therefore, while we can not definitively rule out regulation of protein expression at the translational level, the available evidence would suggest that regulation of HSFs is predominantly at the mRNA level.

Previous studies have demonstrated that maternal RNA from *hsp70* and *hsf1* exist in developing embryos (Bensaude et al. 1983; Bienz et al. 1984; Ovsenek and Heikkila 1990; Gordon et al. 1997; Christians et al. 2000; Metchat et al. 2009; Bierkamp et al. 2010; Åkerfelt et al. 2010). Analysis of zebrafish *hsf2* expression demonstrated that it also exists in zebrafish embryos as maternal RNA, which is the first such evidence of maternal *hsf2* expression in any species.

5.1 *In silico* and EMSA analysis suggest that zebrafish *hsf4* evolved from an *hsf1* duplication event.

Two HSFs have previously been characterized in zebrafish. In 2000, Råbergh et al. reported the identification of two HSF1 isoforms in adult zebrafish tissues and in 2006 Yeh et al. identified HSF2 in zebrafish embryos. Initial experiments in this study were designed to characterize a unique zebrafish HSF4 isoform that was originally identified in an EST screen of an embryonic cDNA library (Strausberg et al. 2002), and subsequently through annotation of zebrafish genomic sequence data. More recently the ORF of a second, longer putative *hsf4* transcript was predicted from the same gene (Ensemble Zv8, release 59), although it was not known if it represented a true gene product. Both *hsf4* transcripts are splice variants of the same gene as the sequence from the first eight exons is identical between the two transcripts with the *hsf4b* containing an additional four exons downstream of the first eight (Fig. 4.1A). The gene structure of zebrafish *hsf4* is similar to other *hsfs* in that the first three exons encode the DBD and exons 4-6 encode the HR-A/B domain (Fig. 4.1D), both of which are preserved in all previously characterized *hsf* genes (Prikkala et al. 2001; Yeh et al. 2006). As well, the splice site in the 8th exon used to produce the two transcripts is similar to the human *hsf4* gene in that the splice variants for both human *hsf4* transcripts are generated from splice sites in the 8th and 9th exons (Tanabe et al. 1999). However, the splice site in the 8th exon of the human gene results in a frameshift to generate the two human transcripts while the splice site in the zebrafish gene results in a stop codon in the *hsf4a* transcript. A UGA stop codon is present 859 nucleotides into the open reading frame of the transcript which encodes the truncated HSF4a protein. One of the more interesting observations made from the sequence analysis is how the longer, more abundant *hsf4b* transcript is generated considering the presence of this stop codon. Immediately upstream of the stop codon in the *hsf4a* sequence is an AGG codon for arginine which is conserved in the *hsf4b* transcript. The 5'-AGG UAG sequence, which is followed by an extensive 3' UTR, for *hsf4a* is entirely derived from the 8th exon and the presence of the stop codon was confirmed by sequence analysis (section 4.1.1, Fig. 4.1B). However, the mRNA sequence for the *hsf4b* transcript is generated from a splicing event that occurs in between the two guanine nucleotides in the AGG codon, hence the AG nucleotides are retained from the 8th exon and the third guanine is acquired from the 9th exon to complete and conserve the AGG codon before the sequence continues in frame. Interestingly, in this scenario, the second guanine from the AGG in the 8th

exon and the uracil from the *hsf4a* stop codon are apparently recognized by the spliceosome as a GU splice donor signaling the start of an intron which is then spliced out in order to generate the *hsf4b* transcript (Fig. 4.1B). This is not a completely unique mechanism as other genes have been reported to produce splice variants using this method. The original sequencing of the human *β-globin* gene determined that the AG splice donor site for the first intron occurs within an AGG codon where AG originates in exon 1 and the final guanine is added from the second exon, which parallels the same intra-codon splicing on an AGG in the *hsf4b* transcript (Lawn et al. 1980). Another example comes from analysis of the sequence of multiple transcripts of equine *chorionic gonadotropin* secreted by the corpora luteum and fetal gonads during equine pregnancy. These experiments demonstrated that the B isoform of this hormone is encoded by a splice variant of *chorionic gonadotropin* that results from a splicing event between a CTC codon in the tenth exon with the remaining nucleotides originating from the eleventh exon. This mid-codon splicing event causes a frameshift in the sequence resulting in a truncated mRNA with a UGA stop (Saint-Dizier et al. 2004). The specific mechanism by which the 5'-AGG UGA-3' sequence would be processed as a codon followed by a stop codon in one gene transcript but also as a partial codon followed by a splice donor site in second alternate transcript has not been characterized. There are several alternative splicing mechanisms that have been characterized and most include a combination of exon and intron splicing enhancers and silencers which bind to specific sequences in the RNA to facilitate or block recruitment of spliceosome components (Chen and Manley 2009). One possible mechanism by which the zebrafish *hsf4* would be alternatively spliced is by the presence of an intron or exon splicing silencer in the *hsf4a* pre-mRNA which would block the recognition of the GT splice site and allow the UGA stop codon to be recognized (Chen and Manley 2009).

The *hsf4a* transcript is characterized by a UGA stop codon located 1003 bp into the mRNA sequence that has been assumed to be the stop codon for the shorter *hsf4a* transcript. However, this is only an assumption based on the fact that it is the first stop codon encountered in the mRNA. Further analysis of the *hsf4a* 3' UTR reveals that there are two other in frame UGA stop codons downstream at 1018 and 1095 bp followed by a TAA stop at 1101. The UGA stop codon in some genes has been demonstrated to be "leaky". Natural suppressor tRNAs have an altered anticodon that allows the translational machinery to recode the nonsense stop codon into a different sense codon (Beier and Grimm 2001). One of the best characterized mechanisms

in eukaryotes involves the recoding of the UGA stop codon into a codon for a selenocysteine (Berry et al. 1993). Selenoproteins are characterized as proteins that contain multiple selenocysteines encoded by several UGA codons in the 3' UTR (Berry et al. 1993). Several studies that have identified selenoproteins in zebrafish including a specific family called SelJ that is expressed almost exclusively in the lens with some expression in the brain and neural tissues at early stages of development (Kryukov and Gladyshev 2001; Castellano et al. 2004, 2005; Shchedrina et al. 2007). It is then tempting to hypothesize that the series of UGA codons in the *hsf4a* 3' UTR might actually code for a series of selenocysteine residues. If this were true, it would result in a protein that is 321 amino acids in length instead of the 286 amino acid that has been predicted. There have been several documented cases of proteins having dual roles as enzymes as well as transcription factors (Green and Paget 2004; Commichau and Stülke 2008; Latrasse 2013). There would be several interesting implications in HSF4a function if it were a selenoprotein. The majority of selenoproteins have been demonstrated to have enzymatic activity involved in redox reactions (Castellano et al. 2005). This is particularly interesting considering redox reactions have been suggested to play a role in activating HSF trimerization (Ahn and Thiele 2003). An enzymatic function related to HSF4a would be a unique observation as no HSFs have been reported to have this function to date. However, other HSFs, specifically HSF1 and HSF2, have been demonstrated to be involved in chromatin modulation but so far no enzymatic function has been associated with this role (Hong and Sarge 1999; Xing et al. 2005; Abane and Mezger 2010; Björk and Sistonen 2010; Sakurai and Enoki 2010; Vihervaara et al. 2013).

The predicted amino acid sequences of both isoforms of zebrafish HSF4 show strong identity to other HSFs which is in agreement with the high levels of amino acid conservation that is characteristic of this protein family (Fig. 4.1B and C; Tables 4.1-4.4). The lack of an HR-C domain in zebrafish HSF4a suggests that, similar to mammalian HSF4, zebrafish HSF4a may also be constitutively trimerized in the cell (Nakai et al. 1997; Björk and Sistonen 2010). However, unlike the mammalian proteins, the lack of an HR-C domain is due to the severe C-terminal truncation of zebrafish HSF4a compared to the mammalian and avian forms of HSF4. Indeed, this truncation makes the zebrafish HSF4a the smallest HSF of any class identified to date. The function of the C-terminal region of the mammalian HSF4 proteins is at present unknown, but the relatively high degree of sequence conservation across this region (73-88%) of

the mammalian proteins suggests that it may be of functional or structural importance. Clearly, any such potential function in this region would not be present within the truncated zebrafish form of HSF4a.

Surprisingly, the zebrafish HSF4b protein and both isoforms of the avian HSF4 contain an HR-C domain and, therefore, may not be constitutively trimerized as has been described for the mammalian protein (Fig. 4.2B). Given the presence of an HR-C domain in the carboxy portion of zebrafish HSF4b it is likely that this domain interacts with other cellular proteins to maintain HSF4b in an inactive state as has been described for HSF1 and HSF2 proteins. The HSF4b HR-C domain shares only 24 to 27 percent amino acid identity when compared to other HR-C domains in zebrafish or mammalian HSFs (Table 4.1). This suggests that some proteins that interact with HSF4b will be unique for the specific regulation of this protein. The degree of amino acid identity shared with the avian HSF4 species is higher at 49% which suggests there may be some similarity in the regulatory aspects of trimer inhibition between zebrafish and the avian species.

The phylogenetic analysis of complete HSF protein sequences from eleven species demonstrates that the zebrafish and avian proteins form a separate subgroup of the HSF4 cluster, where zebrafish HSF4 groups very closely with chicken and zebrafinch predicted HSF4 protein sequences (Fig. 4.3). The close relationship of zebrafish to avian HSF4 appears to be unique to HSF4 proteins as the branching distances between these species is much greater for HSF1 and HSF2. While this group of zebrafish and avian HSF4 proteins does cluster closer to the mammalian HSF4 proteins than to any of the HSF1, HSF2, or HSF3 groups, the zebrafish/avian group branches from HSF1 much sooner in the tree. The zebrafish and avian HSF4 protein sequences both share a high level of amino acid identity in the DBD to HSF1 proteins (section 4.1.4). This feature, combined with the presence of an HR-C domain in the zebrafish and avian proteins sequences, are probable reasons for the shorter branching distance observed between these proteins and HSF1 (section 4.1.3). Similarly, the greater amino acid divergence in the DBD between zebrafish and mammals and the lack of an HR-C domain in all mammalian HSF4 sequences could explain the large branching distance between these proteins. However, the truncated nature of HSF4a is still novel among this group of HSF4 proteins. Interestingly, another truncated vertebrate HSF4 protein has been predicted to exist in the anole lizard (*Anolis carolinensis*) by the Ensembl software. The anole lizard *hsf4* gene contains eight exons that

encode for a protein predicted to contain 282 amino acids and also lacks an HR-C domain, very similar to zebrafish *hsf4a*. The DBD of the lizard putative HSF4 protein also shares 96% amino acid identities with the zebrafish, similar to what is observed between the zebrafish and avian species. Phylogenic analysis of these proteins indicates that the lizard HSF4 clusters very close to the zebrafish and avian species and that the branching distance between this protein and HSF1 is the same as that observed for zebrafish HSF4 (data not shown). However, unlike other species, only a single *hsf4* transcript has been predicted for the anole lizard. In combination, these data suggest that the HSF4 protein family arose from a duplication event of an ancestral *hsf1* gene, and through the process of evolution, mammals have lost the HR-C domain in the carboxy terminal of the protein.

Heat shock factor binding sites which contain slight variations in the classic HSE 5'-nGAAn-3' inverted repeat sequence are referred to as discontinuous HSE sequences (Ovsenek and Heikkila 1990; Sakurai and Takemori 2007; Yamamoto et al. 2009). Several studies have demonstrated that discontinuous HSE sequences affect the binding affinity of specific HSFs although recent studies of genome wide and cell-specific HSF binding patterns have suggested that other as yet unidentified mechanisms have a role preferential HSF binding (Jing et al. 2013; Vihervaara et al. 2013). Variations in the amino acid sequence within protein domains between different HSFs provide these proteins with the ability to have variable functions. Specific amino acid regions of the human HSF1 DBD and HR-A/B domain regulate the ability of this HSF1 to bind to discontinuous HSE sequences (Sakurai and Takemori 2007; Takemori et al. 2009). Enoki and Sakuri (2011) demonstrated that sequence differences in the HR-C domain of zebrafish HSF1 isoforms affected oligomerization which resulted in different binding affinities for discontinuous HSEs between the two isoforms. Similarly other studies of human and mice HSFs have demonstrated that different HSFs will recognize variations of HSE sequences, with HSF1 preferring a continuous inverted 5'-nGAAn-3' repeat sequence while HSF2 and HSF4 are capable of binding to disordered versions of this sequence (Fujimoto et al. 2008; Takemori et al. 2009; Yamamoto et al. 2009, Sakuri and Enoki 2010). In mouse lens tissue, 222 binding sites for HSF4 were identified and of those binding sites only 6 had perfect 5'-nGAAn-3' repeats in the HSE sequences (Fujimoto et al. 2008). Some binding sites had one or no GAA sequences. Modified HSEs with an ambiguous HSE sequence (5'-nGnnn-3') could partially compete mouse HSF4 from a concensus HSE in an EMSA but the same ambiguous sequence could not compete

away mouse HSF1 or HSF2 (Fujimoto et al. 2008). Our results from similar EMSA experiments using zebrafish HSF proteins in HSE binding assays demonstrate that the HSF4 protein does bind to a perfect HSE in a specific manner (Fig. 4.4). However, competition assays demonstrated that HSF4 binds to discontinuous HSEs in a more restricted manner very similar to the binding pattern observed for HSF1 (Fig. 4.5; Fig. 4.6), and unlike the results previously reported in mammalian systems (Fujimoto et al. 2008). This observation is supported by our analysis of amino acid similarities (section 4.1.3) in which we showed the amino acid composition of zebrafish HSF4 DBD was more similar to HSF1 (85%) than to any of the mammalian HSF4 sequences (72%). As well, other studies have demonstrated that there are similar patterns of PTMs including phosphorylation on key serine residues and a SUMO modification on a lysine residue in both human HSF1 and HSF4. These PTMs all contribute to the inhibition of the transactivation domain indicating a functional similarity and adding further evidence to the suggestion that these proteins are more closely related to each other than to other HSFs (Björk and Sistonen 2010; Xu et al. 2012). The phylogenetic analysis (section 4.1.3) also supports the suggestion that HSF1 and HSF4 are closely related and also demonstrates that the branching distance of zebrafish and avian HSF4 is much closer to HSF1 sequences than the mammalian HSF4. The high level of identity observed between the fish and avian HSF4 proteins with HSF1, plus the identification of an HR-C domain in the zebrafish HSF4, coupled with the relatively low affinity of HSF4 for discontinuous HSE sequences, suggests that target genes for the zebrafish and avian HSF4 transcription factors may not be fully shared with those regulated by mammalian HSF4. To date, no target genes for HSF4 in fish and avian species have been identified.

5.2 Zebrafish *hsf4* is expressed in a developmental and tissue specific manner

Expression of both *hsf4* transcripts was detected in both zebrafish embryos and adult tissues (Fig. 4.13). In all tissues where *hsf4* mRNA was detected, calculations suggested *hsf4b* was always more abundant than *hsf4a*, the difference being especially pronounced in the eye where the *hsf4b* mRNA levels were calculated to be 185 fold higher than *hsf4a*. This is consistent with the fact that *hsf4b* levels were also demonstrated to be higher than *hsf4a* in mouse tissues (Nakai et al. 1997; Tanabe et al. 1999; Somasundaram and Bhat 2004). Analysis of HSF4 isoforms in mice demonstrated that the less abundant HSF4a isoform was actually a

repressor of several genes while the more abundant HSF4b activated gene transcription. Since the relative abundance of the two transcripts of *hsf4* in zebrafish appears to be similar to that reported for mice, it would be interesting in future studies to determine if the less abundant zebrafish HSF4a also functions as a repressor.

The high level of *hsf4* expression in the eye relative to other tissues supports the third hypothesis that HSF4 has a role in zebrafish lens development. These results are also consistent with other studies that have identified an important role for mammalian HSF4 in lens development and differentiation of lens fibers (Fujimoto et al. 2004; Min et al. 2004; Morange 2006; Shi et al. 2009; Abane and Mezger, 2010; Zhou et al. 2011; Jing et al. 2013). Mutations in *hsf4* genes have been linked to several human families with hereditary congenital cataracts as well as several dog breeds that also frequently exhibit juvenile cataracts (Bu et al. 2002; Smaoui et al. 2004; Ke et al. 2006; Mellersh et al. 2006; Engelhardt et al. 2007). Several of these mutations occur in the DNA binding domain of the protein and affect the ability of HSF4 to bind to HSE sequences (Enoki et al. 2010). Cataracts are also observed in mouse models carrying mutations introduced into the *hsf4* gene (Fujimoto et al. 2004). The onset of cataract formation in the lenses of mammals that have a mutation in *hsf4*, is due in part to improper regulation of the crystallin genes, especially γ-crystallin genes, by HSF4 (Somasundaram and Bhat 2004; Shi et al. 2009). HSF4 has also been shown to act cooperatively with HSF1 in lens epithelial cells to regulate cell proliferation as well as having a role in the regulation of *hsp27* during lens development (Fujimoto et al. 2004; Min et al. 2004). While no studies have specifically examined the regulation of *hsp70* expression by HSF4, several recent studies have shown that an absence of *hsf4* expression results in an increase in *hsp70* mRNA levels (Fujimoto et al. 2004; Kim et al. 2011; Cui et al. 2012). These studies suggested that HSF4 would represent a good candidate to be involved in zebrafish lens development and in co-operative regulation of lens-specific *hsp70* gene expression together with HSF1 (Evans et al. 2007).

In adult tissues, *hsf4b* was prominently expressed in the eye but a basal level of expression was detected in all tissues assayed (Fig. 4.13C). Expression of *hsf4a* was also most prominent in the eye, but significant expression in the testis was also detected (Fig. 4.13B). Initial studies on human *hsf4a* identified expression in human heart, brain, skeletal muscle, and pancreas while *hsf4b* was detected in the brain and the lung (Nakai et al. 1997; Tanabe et al. 1999). Further studies clarified that the eye was the main site of *hsf4* expression in several mammalian species

(Bu et al. 2002; Fujimoto et al. 2004; Mellersh et al. 2006; Engelhardt et al. 2007; Abane and Mezger 2010). To date no other study has identified *hsf4* expression in the testis; in fact, experiments performed by Fujimoto et al. (2004) demonstrated that classical *hsp* and *fgf-1* expression was unchanged in the testis of *hsf4* knockout mice. Several studies characterizing the expression and function of HSF1 and HSF2 have been performed in testis tissue. These studies have shown that HSF2 is more abundant than HSF1 but both proteins are required for successful spermatogenesis (although this has only been demonstrated in mammals) (Fiorenza et al. 1995; Wang et al. 2003; Åkerfelt et al. 2010; Björk and Sistonen 2010; Björk et al. 2010). A role for HSF4 in the testis remains unknown. HSFY is another member of the HSF family which has been shown to be exclusively expressed in the testis of mammals. The HSFY gene is located within the male specific region of the Y chromosome in humans and while its function is currently unknown, it has been shown to be temporally expressed in spermatocytes during spermatogenesis and has been linked to reduced fertility (Forresta et al. 2001; Skaletsky et al. 2003; Shinka et al. 2004; Tessari et al. 2004; Kichine et al. 2012). Zebrafish do not have sex chromosomes and to date an orthologue of the HSFY gene has not been identified on any chromosome in zebrafish. Several recent studies have suggested that sex differentiation in zebrafish is polygenic and several clusters of sex determining genes have been identified within regions of four separate chromosomes (Anderson et al. 2012; Liew et al. 2012; Howe et al. 2013). Environmental factors, such as hypoxia, high temperatures, and crowding, also have a secondary role in affecting sex differentiation in zebrafish (Anderson et al. 2012; Howe et al. 2013). The lack of sex chromosomes in the zebrafish, combined with the unique expression of the shorter *hsf4a* transcript in the testis, could suggest that zebrafish HSF4a may share some of the as yet unknown functions performed by the mammalian HSFY during spermatogenesis. However, the lack of all heptad repeat sequences in HSFY which are typically found in HSFs is a very important structural difference between it and zebrafish HSF4a (Shinka et al. 2004). Additionally, HSFY does not appear to trimerize and cannot bind traditional HSEs (Shinka et al. 2004). Therefore, the function of the HSFY protein is most likely through an as yet uncharacterized mechanism which is completely unique to the HSF family. However, these structural differences do not entirely preclude zebrafish HSF4a from sharing some functions of mammalian HSFY. It is possible that the zebrafish *hsf4*a is related to the mammalian HSFY autosomal ancestor and therefore a functional similarity between the two genes may still exist.

The expression of both *hsf4* splice variants is developmentally regulated in zebrafish embryos and the results of the expression analysis reported here support the hypothesis that HSFs have distinct developmental expression patterns. Neither transcript of *hsf4* could be detected by qPCR in embryos until 60 hpf (Fig. 4.15A and B), unlike *hsf1* (Fig. 4.17A and B) and *hsf2* (Fig. 4.18B) which were shown to be expressed at all developmental stages. The developmental regulation of *hsf4* expression is consistent with observations in mice lenses in which *hsf4* mRNA levels were not detected in mice younger than the E12.5 stage embryos, at which point *hsf4* was detected in both lens epithelial and fibre cells (Fujimoto et al. 2004; Abane and Mezger 2010). Additionally, levels of *hsf4* mRNA were also shown to increase and peak in post-natal mouse lenses after which mRNA levels declined somewhat (Somasundaram and Bhat 2004; Fujimoto et al. 2004, 2008; Björk and Sistonen 2010).

The inability to detect *hsf4* expression in zebrafish embryos prior to 60 hpf was unexpected considering the development of the zebrafish lens begins between 18-20 hpf and the expression of *hsp70* can be detected at 30 hpf (Blechinger et al. 2002). However it was consistent with previous studies in which MO knockdown of *hsf4* expression in zebrafish failed to produce an eye phenotype at 48 hpf, the point in development when both *hsp70* and *hsf1* knockdown results in a clear small lens phenotype (Evans et al. 2005; Evans 2006). Cui et al. (2013) were recently able to show localization of *hsf4* expression to the lenses of zebrafish embryos at 24, 36, 48 and 60 hpf using a whole mount *in situ* hybridization technique. Their results showed a small but detectable level of *hsf4* mRNA in the lens at 24 and 36 hpf with increasing mRNA levels at 48 hpf and 60 hpf. The large increase in mRNA levels observed by Cui et al. (2013) at 60 hpf would support our qPCR results which demonstrate that *hsf4* levels go from undetectable to well within the limits of detection at this developmental stage. However, Cui et al. (2013) also report detection of *hsf4* expression in younger embryos by *in situ* hybridization, which is in stark contrast to the results reported in this thesis in which attempts to detect *hsf4* in whole zebrafish embryos by *in situ* hybridization were largely unsuccessful. Expression of *hsf4* in the lens was only detected in a single experiment at 36 hpf which could not be consistently repeated at any other time point (Fig. 4.10). Our initial conclusion from these experiments was that our *in situ* hybridization technique was not sensitive enough to detect *hsf4* mRNA in whole mount embryos. The methodology cited by Cui et al. (2013) is taken from an updated method claiming more sensitive *in situ* hybridization detection for zebrafish (Thisse and Thisse 2008). This updated

methodology is nearly identical to the method used in this thesis with a couple of notable exceptions. The authors expose embryos to proteinase K for a longer period of time prior to DIG-probe hybridization compared to our method which could have increased the permeability of the tissues. Increased permeability of tissues, especially densely packed tissues like the lens, would allow for greater access of the probe to mRNA, thereby increasing the sensitivity of the assay. However, it is important to note that the method used in this thesis is based on the method used by Blechinger et al. (2002) in which the proteinase K treatment was adequate for *hsp70* mRNA to be detected in the zebrafish lens. The experiments performed by Cui et al. (2013) also used a DIG-labeled RNA probe while the experiments presented in this thesis used a DIG-labeled cDNA probe. The use of an RNA probe could have increased the assay sensitivity as RNA-RNA hybridizations are typically more sensitive than cDNA-RNA hybridizations. Initial *in situ* hybridization analysis presented in this thesis as an internal control demonstrated that both RNA and cDNA probes could successfully detect *α-tropomyosin* expression in zebrafish embryos (Fig. 4.9). However, *α-tropomyosin* is expressed at high levels throughout somatic muscle during all developmental stages examined. While a cDNA-RNA hybridization may have been sufficient to detect *α-tropomyosin*, it may not have been sensitive enough to detect the low amount of *hsf4* transcript that exists in zebrafish lens. Additionally the analysis performed by Cui et al. (2013) used an overnight incubation with an anti-DIG antibody to detect the DIG labeled probe which is longer than the 3-4 hour incubation period used in our analysis. This may have also improved the sensitivity of their assay.

Cui et al. (2013) also state that they detected *hsf4* mRNA in 24 and 36 hpf embryos with qPCR but did not report C_T values or show other data to support this. Our analysis of *hsf4* mRNA levels in embryos at 10 and 30 hpf with qPCR data did occasionally generate C_T values that could have been used to calculate mRNA expression. However, the majority of the C_T values generated for these time points were either larger or the same as the non-template control or had very large deviations between the technical replicates. These scenarios are typical for samples were mRNA levels are below the limit of detection for the assay and C_T values of this nature are not to be used for quantitative analysis as described in the MIQE guidelines (Bustin 2000; Bustin et al. 2009). The fact that qPCR analysis was performed on whole embryo extracts while *hsf4* mRNA is primarily expressed in only lens tissue may account for mRNA levels of

hsf4 that were below the limit of detection. For these reasons we can only report with confidence that mRNA levels were quantifiable in embryos 60 hpf and older.

The low level of *hsf4* mRNA expression in developing zebrafish embryos prior to 60 hpf (Fig. 4.15) may provide an explanation for the lack of phenotypic effect observed in the MO knockdown studies (Evans et al. 2005; Evans 2006). Cui et al. (2013) also used *hsf4*-MO knockdown technology to determine the effects of reduced *hsf4* expression in zebrafish embryos. While these embryos also did not develop a small lens phenotype or cataract, they did have an apparent retention of nuclei in the lenses of embryos at 72 hpf which was not observed at 48 hpf or 96 hpf. While MO are a very effective tool to knockdown gene expression in zebrafish, they become diluted in the cells as the embryos develop, and typically MO induced phenotypes are severely reduced by 70 hpf and are absent by 90 hpf (Evans et al. 2005; Cui et al. 2013). Development of the zebrafish eye lens begins between 18-20 hpf, with *hsp70* expression beginning by 30 hpf (Blechinger et al. 2002; Greiling and Clark 2009, 2012). Lens development and growth continue into later embryogenesis and larval/juvenile stages. Indeed, lens growth in fish continues into adulthood.

The results of the mRNA expression analysis, together with the lack of phenotype in the MO knockdown embryos prior to 72 hpf, would suggest that if *hsf4* plays a significant role in the zebrafish eye, it would occur in later stages of development. Mice carrying deletions of *hsf4* exhibit normal lens formation in early embryogenesis, with defects only becoming apparent during late gestation and postnatally (Fujimoto et al. 2004; Shi et al. 2009). In mice, lens vesicles are completely separated from the overlying epithelium by E11, suggesting that HSF4 also plays no role in earlier events of lens formation in this species. *Gamma-crystallin* genes have been shown to be one of the primary gene targets of HSF4 in mouse lens development and HSF4 has also been shown to regulate expression of other crystallin genes such as *α*B-*crystallin* (Fujimoto et al. 2004; Min et al. 2004; Somasundaram and Bhat 2004; Shi et al. 2009; Jing et al. 2013). A proteomic analysis of protein expression in zebrafish lens demonstrated that expression of the more than 50 different α, β, and γ-crystallins occurs in a sequential manner over the course of lens development (Greiling et al. 2009). Only 13 of the 36 zebrafish *γ-crystallin* genes are expressed in the lens by 5 day post fertilization. Expression of the remaining members of the *γ-crystallin* gene family occurs much later, between 3 weeks and 6 months of age (Greiling et al. 2009). Assuming that HSF4 has a role in regulating crystallin genes, and especially the *γ-*

crystallin gene family, in zebrafish as it does in mice, the sequential and delayed expression of crystallin gene families during lens development may partially explain the increase in *hsf4* mRNA expression at later developmental stages. In any event, these data suggest that HSF4 and HSF1 do not work cooperatively to regulate early aspects of zebrafish lens formation, such as *hsp70* induction observed at 30 hpf. This does not rule out the possibility of such co-operative function in later stages of zebrafish lens development.

The development of functional vision in zebrafish is essential for survival of the animal and has been determined to occur at the time the fish develop the ability to capture prey (Easter and Nicola 1996). Zebrafish begin to show a startle response to shadow images and begin to distinguish between light and darkness between 69-72 hpf (Easter and Nicola 1996; Gestri et al. 2011). This behaviour is quickly followed by the development of eye movements that are indicative of tracking prey and the ability to distinguish shapes at 73-80 hpf. The development of prey tracking behaviour also coincides with retinal image formation and the maturation of extraocular muscles (Easter and Nicola 1996; Gestri et al. 2011). As well, at 72 hpf, the OFZ in the lens has expanded to all but the outermost layers of the lens fibres and cell nuclei are only visible in a few cells that lie in the posterior section (near the retina) of the lens (Greiling and Clark 2012). The end of 72 hpf or the third day of development also marks the completion of the hatching stage of zebrafish development, after which the fish are referred to as larvae (Kimmel et al. 1995). After hatching, the fish are no longer protected from the environment by their chorion and it stands to reason that the completion of functional vision, which is necessary for survival, would coincide with this stage of development. Expression analysis of *hsf4* transcripts in the results presented in this thesis demonstrate that there is a trend for both transcripts to be less abundant at 60 hpf when compared to mRNA levels at 105 hpf. As well, calculations suggest the level of *hsf4b* mRNA was always higher than *hsf4a* mRNA levels (section 4.4.3.2). However, in embryos at 80 hpf the abundance of the *hsf4b* transcript appears to increase to levels similar to 105 hpf while the *hsf4a* mRNA levels remain lower than at 105 hpf. This increase in only the *hsf4b* transcript also effectively increases the b:a ratio of *hsf4* transcripts from approximately 7 fold at 60 hpf to embryos having approximatley 30 fold more at 80 hpf before the ratio decreases to 9 fold more *hsf4b* at 105 hpf, similar to that seen at 60 hpf. The fact that the increase in *hsf4b* is observed at the same developmental stage at which hatching is completed and functional vision is achieved in zebrafish is intriguing. In mouse development, a peak in HSF4 expression

was observed in lenses immediately after birth after which the levels of HSF4 declined (Björk and Sistonen 2010). It may be possible that the increase in *hsf4b* observed in 80 hpf zebrafish is occurring at a similar developmental time point as the peak observed in mice lenses, the role of which in development is still unknown. It has also been shown that HSF1 and HSF4 are necessary for the function of sensory organs in mammals after they have been exposed to the environment for the first time (Åkerfelt et al. 2010). Based on this observation, there could be a need for increased *hsf4b* expression by the zebrafish to maintain or further develop functional vision at 80 hpf as the time period surrounding this developmental stage would the first time larvae are exposed to an environment outside of the chorion.

The primary role of HSF4 has been determined to be a transcriptional activator or repressor of genes during development of the lens. So far gene targets regulated by HSF4 are themselves parts of cellular mechanisms that are either involved in the differentiation of lens epithelial cells into fibre cells (*fgfs*, *RhoA* and *Rac1*) or provide structural proteins for the lens (*γ and αB-crystallins*, *SKAP*, *Bfsp*) (Fujimoto et al. 2004, 2008; Somasundaram and Bhat 2004; Shi et al. 2009; Enoki et al. 2010; Zhou et al. 2011; Hu et al. 2013; Jing et al. 2013; Li et al. 2013). These proteins are required for the development of a functionally transparent lens with the appropriate refractive index through which light can be transmitted and focused to form an image on the retina. Acquisition of lens transparency to allow image formation on the retina is a fundamental component of eye development and achieving functional vision. The expression of various crystallin genes occurs sequentially throughout the growth period of the fish and unique crystallins may begin to be expressed at this stage. Previous studies in mammals have shown that the *γS-crystallin* gene family is one of the primary targets of HSF4 (Fujimoto et al. 2004; Min et al. 2004; Shi et al. 2009). However, these proteins would not be related to the acquisition of vision in zebrafish at this stage as they are not expressed until six weeks of age and later (Greiling et al. 2009).

The removal of DNA and organelles from lens fibres during maturation is an important process that allows the lens to achieve transparency and funnel light through to the retina (Greiling et al. 2009; Greiling and Clark 2012). Two different pathways, an attenuated apoptosis pathway and the ubiquitin-proteosome pathway, have been implicated in the elusive mechanism of organelle and DNA clearing during lens development (Wride 2011; Greiling and Clark 2012). The Hsp70 chaperone is expressed under non-stress conditions in the lens of developing

zebrafish (Nakai et al. 1997; Fujimoto et al. 2004; Evans et al. 2005; Evans 2006; Kim et al. 2011; Cui et al. 2012). Knockdown of this gene during zebrafish development results in a small eye and lens phenotype (Evans et al. 2005; Evans 2006). There has been some evidence that this chaperone may have a role in both of the above pathways. The Hsp70 chaperone protects proteins from mis-folding and aggregation during stresses that can eventually lead to apoptosis and therefore has been suggested to have a protective role in the lens during the attenuated apoptosis process linked to protein clearance in the lens (Mathew and Morimoto 1998; Pirkkala et al. 2001; Evans et al. 2005; Evans 2006; Abane and Mezger 2010). In the ubiquitin-proteosome pathway, Hsp70 has been shown to be required for the formation of a functional 26S proteosome (Awasthi and Wagner 2005; Grune et al. 2011). The *volvox* zebrafish mutant, known to have a non-functioning Psmd6 subunit of the 26S proteosome, has a small lens phenotype similar to that observed in *hsp70* knockdown zebrafish lending further support for a role for Hsp70 in this pathway (Evans et al. 2005; Evans 2006; Imai et al. 2010).

Recent evidence suggests that HSF4 may have role in facilitating degradation of DNA in lens fibres during eye development by regulating the expression of genes such as *Rad 51 recombinase*, *DLAD* and *DNase IIβ* (He et al. 2010; Cui et al. 2012, 2013). Several studies have also demonstrated that HSF4 regulates *hsp70* expression in the lens, presumably at later stages of lens development as expression of *hsp70* is detected earlier (28 hpf) than *hsf4* (60 hpf) in the lenses of zebrafish (Blechinger et al. 2002; Fujimoto et al. 2004; Kim et al. 2011; Cui et al. 2012, 2013). In zebrafish embryos, a developmentally specific increase in Hsp70 was observed between 72-84 hpf in the absence of heat shock (Yeh and Hsu 2000, 2003). These elevated levels of Hsp70 at 84 hpf were eight fold higher than those observed in embryos between 12-60 hpf and 96-108 hpf stages of development (Yeh and Hsu 2003). These studies also demonstrated that the largest increase in Hsp70 levels in response to heat shock also occurred at 84 hpf. While the analysis of *hsp70* expression presented in this thesis does not show a significant difference in basal mRNA levels at 80 hpf when compared to other developmental stages, levels of *hsp70* mRNA after heat shock do appear to be highest at 10 hpf and 80 hpf (Fig. 4.13C). Interestingly, the sudden and stage specific increase in Hsp70 expression observed by Yeh and Hsu (2000) occurs at the same time as the increase in *hsf4b* mRNA levels observed at 80 hpf. It is therefore tempting to speculate that higher levels of *hsf4b* at 80 hpf result in increased levels of HSF4b in the embryo for the purpose of increasing transcription of target genes, including *hsp70*, which

encode for proteins required for lens clearing. An increase in this suite of proteins at the 80 hpf stage may be necessary to achieve a final clearance of organelle and DNA fragments from lens fibre cells to establish a completely transparent lens necessary for the acquisition of functional vision in zebrafish larvae which coincides with this stage of development.

Developing zebrafish embryos exposed to heat shock had an increase in the mRNA levels of both *hsf4* transcripts at 60 hpf (Fig. 4.15C and D). These increases returned to non-heat shocked levels after a recovery period. The same response to heat shock was not seen in older embryos at 80 and 105 hpf. While we did not observe a response to heat shock in embryos younger than 60 hpf, we can not rule out that a similar response may occur. H*sf4* expression in younger embryos has been shown to occur earlier than 60 hpf in other studies even though expression was below the limit of detection in the qPCR analysis presented in this thesis (Cui et al. 2013). To date, no other studies have investigated *hsf4* expression in response to heat stress. Heat shock had no effect on *hsf1a+b* (Fig. 4.17C) and *hsf2* (Fig. 4.18C) mRNA levels in zebrafish in our experiments but there was a decrease in *hsf1b* in response to heat shock which will be discussed further in section 5.5. The results for *hsf1* and *hsf2* mRNA in our studies are consistent with previous studies which have demonstrated that *hsf1* and *hsf2* mRNAs are maintained at constant levels during heat shock with an exception being zebrafish *hsf1b* which has been shown to decrease in response to heat shock (Lindquist 1986; Råbergh et al. 2000; Pirkkala et al. 2001; Airaksinen et al. 2003a, 2003b). Interestingly, avian HSF3 is the only other HSF where an increase in expression has been demonstrated in response to severe and prolonged heat shock (Tanabe et al. 1997).

Studies investigating the effect of heat shock on potential HSF4 targets demonstrated that heat shock could up-regulate *hsp70*, *hsp90* and *hsp27* in mammalian cells transfected with human *hsf4b* (Tanabe et al. 1999). As well, HSF4b was able to re-establish thermotolerance in an *hsf* null yeast cell line. The same response could not be repeated in either cell line transfected with *hsf4a* (Tanabe et al. 1999). A later study profiling HSF binding sites in mouse lenses demonstrated heat shock induced HSF4 binding to the promoter region of several non-classical heat shock genes, some of which were also bound by HSF1 (Fujimoto et al. 2008). They were also able to demonstrate that in lenses from *hsf4* null mice, HSF1 was not able to bind to these same sites upon heat shock. Further analysis demonstrated that HSF4 binding was associated with decreased levels of methylated histone H3K9 and relaxed chromatin structure, both in

mouse lenses and in cell culture. It was determined that the absence of H3K9 methylation regulation by HSF4 in *hsf4* null mouse lenses was responsible for decreased HSF1 binding in these regions (Fujimoto et al. 2008). It may be that the increase in both zebrafish *hsf4* transcripts observed at 60 hpf may be necessary to directly regulate gene expression related to proteostasis in response to stress. At this stage in lens development, fibres are rapidly undergoing denucleation and organelle clearing, a process that occurs on a large scale between 50-72 hpf. As discussed earlier, several studies demonstrate that HSF4 can regulate expression of *hsp70*, a chaperone involved in regulating proteostasis in response to stresses, in the lens (Fujimoto et al. 2004; Kim et al. 2011; Cui et al. 2012, 2013). Another possibility is that an increase in HSF4 may be related to an increased need of the embryo to modify chromatin to allow HSF1 access to classical heat shock genes required to maintain proteostasis during the stress in the lens. In either case, the data presented here suggest that HSF4 is important for the stress response in embryos at early stages of development while the same is not true for older embryos as heat shock did not increase *hsf4* mRNA levels in embryos at 80 or 105 hpf. The possible role of HSF4 regulation of the heat shock response in early developmental stages is an interesting one that should be investigated further.

5.3 Expression of *hsf1* in zebrafish embryos and tissues is similar to mammalian expression patterns

The expression levels of two *hsf1* transcripts were analyzed in zebrafish across developmental stages and in adult tissue by determining the mRNA levels of combined *hsf1a+b* transcripts and the *hsf1b* transcript alone (Fig. 4.17). Both transcripts of *hsf1* are splice variants of the same gene with *hsf1b* being identical in sequence to *hsf1a* except for a 78 bp sequence immediately upstream of the HR-C domain (Råbergh et al. 2000). Calculations suggest that the mRNA levels for the combined *hsf1a+b* are always more abundant than *hsf1b* transcripts alone for both embryos and adult tissues, indicating that *hsf1a* is always present in the same tissues as *hsf1b* (section 4.4.4.1 and 4.4.4.2). Both the combined *hsf1a+b* mRNA and *hsf1b* mRNA alone was expressed in all tissues, with the spleen, muscle, gills, ovary and dissected eggs having significantly lower expression of both transcripts than the heart (Fig. 4.16). This observation suggests that there is some tissue specific regulation of the expression of *hsf1* splice variants. These observations are consistent with the ubiquitous expression with variable mRNA levels of

both *hsf1* transcripts observed in all tissues in mice (Fiorenza et al. 1995). In mice, the highest *hsf1* mRNA levels were observed in the testis, heart and skeletal muscle while the weakest expression was observed in the liver and kidney (Fiorenza et al. 1995). The mRNA levels of *hsf1* observed in zebrafish were similar to those seen in mice in that, with the exception of mature eggs, the highest *hsf1* levels were also observed in the heart and testis. However, the levels of mRNA in zebrafish skeletal muscles were significantly lower than the heart while expression levels in liver appear to be similar to the heart. Mature eggs (squeezed) had significantly higher levels of both mRNAs (Fig. 4.16), the implications of which will be discussed in more detail in section 5.5. Both transcripts were also expressed in all stages of embryo development with higher levels of *hsf1* transcripts in embryos at 10 hpf (Fig. 4.17). Again, this observation will be discussed in greater detail in section 5.5. One of the most striking observations about *hsf1* expression was seen in embryos after heat shock treatment. The mRNA levels of *hsf1b* decreased or showed a trend to decrease after heat shock treatment at 60, 80 and 105 hpf (Fig. 4.17D). The decrease in *hsf1b* mRNA levels appears to return to non-heat shock levels after a 15 hour recovery period at normal growth conditions.

Very few investigations have looked at the differential expression and function of the two HSF1 isoforms in any species. Previous studies in zebrafish have shown that heat shock resulted in a tissue specific decrease in *hsf1b* in liver and gills which was accompanied by an increase in Hsp70 (Råbergh et al. 2000). A decrease in the *hsf1b* transcript level in response to heat stress was also shown in the ZF4 cell line (Airaksinen et al. 2003a). However, while zebrafish exposed to heavy metal toxicity stress using cadmium and copper also had increased Hsp70 expression in liver and gonad tissues, the decrease in *hsf1b* was not observed leading the authors to suggest that the change in ratios of *hsf1* mRNA levels may be the result of a stress specific mechanism (Airaksinen et al. 2003a, 2003b). This mechanism may come as a result of heat shock inducing an HSF1a and b heterotrimer as Ojima and Yamashita (2004) were able to demonstrate that the two isoforms in trout could form heterotrimers. While the results presented in this thesis did not examine the effect of heat shock on *hsf1* mRNA levels in specific tissues, whole zebrafish embryos showed a similar response of decreased *hsf1b* expression after heat shock across all developmental time points. This suggests that a response mechanism regulating HSF1 in embryos during development is similar to that observed in liver and gills. While heat shock caused a decrease in *hsf1b* mRNA levels, there was no significant change observed in the level

of *hsf1a+b* mRNA, suggesting that levels of *hsf1a* may have increased as *hsf1b* levels decreased (Fig. 5.17C). This suggests that heat shock may affect mechanisms that regulate *hsf1* alternative splicing. Further studies are required to elucidate the specific role and mechanism by which *hsf1b* has in regulating heat stress and to determine if a similar mechanism is observed in other vertebrate species.

5.4 Pattern of *hsf2* expression in zebrafish embryos and tissues differs from mammalian expression patterns.

The expression of *hsf2* was also analyzed in zebrafish across developmental stages and in adult tissue by determining the mRNA levels using qPCR. Expression of *hsf2* was detected at all developmental stages with a trend to higher levels of mRNA only at 10 hpf (Fig. 4.18B) which will be discussed in greater detail in section 5.5. The results presented here are similar to those obtained from studies analyzing the expression of *hsf2* in post-implantation mouse embryos from stage E8.5 to E15.5 in which *hsf2* levels were highest in the earlier developmental stages (Rallu et al. 1997; Min et al. 2000). However, *hsf2* mRNA and protein levels gradually decreased in abundance throughout the mouse embryo and began to localize to the central nervous system in later developmental stages. This expression pattern differs from the seemingly steady state levels of *hsf2* mRNA levels observed in later developmental stages in zebrafish.

Expression of *hsf2* was also detected in all tissues analyzed with only the mature eggs having significantly higher mRNA levels (Fig. 4.18A) which, again, will be addressed in section 5.5. The ubiquitous expression of *hsf2* in all adult tissues with varying levels of steady state mRNA between tissues is consistent with what has been observed in mice (Fiorenza et al. 1995; Björk and Sistonen 2010). While the expression of *hsf2* in mice has been shown to occur in most tissues, the pattern of expression is distinct with higher levels of mRNA in testis and skeletal muscle. The results presented in this thesis are quite different than those seen in mice as our analysis showed a trend to higher mRNA levels in the eye and liver but no significant increase in the testis compared to the heart. This is an intriguing result considering the body of recent literature showing that HSF2 has a critical role in spermatogenesis and male fertility in mammals (Wang et al. 2003; Åkerfelt et al. 2010; Björk et al. 2010). Spermatogenesis in mammals and fish are similar in several ways. Males from both groups have bilateral testes which contain spermatogonia within seminiferous tubules. The spermatogonia undergo meiosis to form

spermatocytes that eventually become mature sperm capable of fertilizing eggs from a female (Menke et al. 2011). The stages of spermatogenesis in zebrafish are similar to other vertebrates, however mature sperm develop much faster in zebrafish, taking only six days to develop from spermatogonia to mature sperm while this process takes 5-7 weeks in most mammals (Leal et al. 2009). One of the primary differences between mammals and fish is that the process of spermatogenesis in fish occurs in cysts formed from cytoplasmic projections from one or two Sertoli cells surrounding a single spermatogonium (Leal et al. 2009; Menke et al. 2011). These cysts line the seminiferous tubules within the testis that are structurally long tubes contained within the body of the fish. In mammals, multiple spermatogonium and Sertoli cells are contained within seminiferous tubules which are tightly coiled within multiple lobes of the testis (Johnson and Everett 1995). Mammalian testes are not maintained within the animal's body but are held in a scrotal sac positioned outside of the body to facilitate thermoregulation and keep the testes cooler than the core body temperature during spermatogenesis (Waites 1991; Skandhan and Rajahariprasad 2007). Thermoregulation of the testis during spermatogenesis is crucial to the maintenance of sperm quality in mammals (Waites 1991; Skandhan and Rajahariprasad 2007). However, this is a characteristic that seems to be specific to mammalian sperm development as other species such as birds, amphibians and fish retain their male reproductive organ inside the body cavity and spermatogenesis occurs successfully through a range of temperatures including those similar to the core body temperature (Paniagua et al. 1990; Beaupre et al. 1997; Leal et al. 2009). The expression of HSF2 during spermatogenesis has only been characterized in mammalian species and the exact nature of its function is still under investigation. It may be possible that HSF2 is required to regulate genes required for thermoregulation during spermatogenesis in mammals which would not be required during zebrafish spermatogenesis, explaining the lower level of *hsf2* expression observed in our results.

HSF2 has also been demonstrated to have a role in chromatin modulation of the Y chromosome in mice. Several recent studies have revealed that there are multiple binding sites for HSF2 on the mouse Y chromosome in both mitotic and non-mitotic stages of cell division, suggesting that HSF2 may be required for transcription of genes from the Y chromosome (Åkerfelt et al. 2008, 2010; Vihervaara et al. 2013). Zebrafish do not contain sex chromosomes and therefore chromatin modulation of the type found in mouse testis would not occur in

zebrafish, offering another possible reason for a lower abundance of *hsf2* mRNA observed in zebrafish testis.

5.5 *hsf1* and *hsf2* are present as maternal RNA in zebrafish eggs.

Analysis of the tissue specific expression of HSFs revealed that mRNA for both *hsf2* and *hsf1* had significantly more transcripts in mature eggs (squeezed eggs) than in any other tissue including immature eggs (dissected eggs) and the entire ovary (Fig. 4.16; Fig. 4.18A). Complementing this is the observation that there was a trend to higher *hsf1* and *hsf2* transcript levels in embryos at 10 hpf. These transcript levels were drastically reduced at 30 to 80 hpf (Fig. 4.17A and B; Fig. 4.18B). Transcript levels of *hsp70* were also significantly higher in mature eggs than in any other tissue, however the same was not true in 10 hpf embryos (Fig. 4.12). Together, these observations strongly suggest that mRNAs encoding for both isoforms of HSF1 and HSF2 as well mRNA encoding Hsp70 are deposited as maternal RNA in zebrafish eggs at a later stage in folliculogenesis when eggs are nearing maturation.

In zebrafish, the process of folliculogenesis and oogenesis is divided into five stages, these being primary growth, cortical alveolus or pre-vitellogenic, vitellogenic, maturation and finally a mature egg (Knoll-Gellida et al. 2006). The developmental process from primary stage follicles to post-vitellogenic stage follicles takes 10 days after which the eggs are arrested prior to final maturation and ovulation (Knoll-Gellida et al. 2006; Clelland and Peng 2009). During oogenesis, maternal mRNA and proteins are densely packed into eggs. However, there is little known about the exact timing and mechanism for deposition of maternally derived molecules in zebrafish embryos (Clelland and Peng 2009). Maternal RNA and proteins are important for early stages of embryo development as there is no activation of zygotic genes. Maternal RNA and proteins are primarily responsible for cellular functions and early cleavage stages of development, prior to the initiation of zygote directed gene transcription (Knoll-Gellida et al. 2006; Tadros and Lipshitz 2009).

In zebrafish, the mid-blastula transition (MBT) occurs at cell cycle 10 which is approximately 3 hpf (Kane and Kimmel 1993; Tadros and Lipshitz 2009). The MBT, which is characterized as a lengthening of the cell cycle and a loss of cell cycle synchrony, occurs at roughly the same time as the major wave of zygotic gene activation (Kane and Kimmel 1993; Tadros and Lipshitz 2009). The maternal to zygote transition (MZT) refers to the developmental

time period where regulation of cell cycle and metabolism transitions from maternal regulation to regulation dependent solely on zygote derived RNA and protein. The MZT spans the time where maternal mRNA and proteins are being degraded and the zygote begins to regulate its own gene expression and ends when all maternal RNA has been degraded and regulation of development has been completely transferred to the zygote (Tadros and Lipshitz 2009). In zebrafish the MZT spans the time period of 3-8 hpf which encompasses the MBT.

Maternal RNA in zebrafish has been divided into three different groups characterized by whether the majority of the RNA degradation has occurred before, during or after zygote gene activation (Tadros and Lipshitz 2009). The results presented in this thesis suggest that maternal *hsf* expression would be placed into the third group as mRNA levels were still high in embryos at 10 hpf (Fig. 4.17A and B; Fig. 4.18B). However, *hsp70* most likely would belong to either the first or second group as *hsp70* mRNA levels at 10 hpf were not significantly different from those observed at any other developmental stage analyzed (Fig. 4.12B). Previous studies support our observation that both *hsf1* and *hsp70* are expressed as maternal transcripts. Expression of *hsp70* was detected in oocytes from mouse, *Xenopus* and Atlantic cod (Bensaude et al. 1983; Metchat et al. 2009; Skjaerven et al. 2011) and *hsf1* transcripts have been identified in both *Xenopus* and mouse oocytes (Bienz 1984; Bienz et al. 1984; Ovsenek and Heikkila 1990; Gordon et al. 1997; Christians et al. 2000; Metchat et al. 2009; Bierkamp et al. 2010; Åkerfelt et al. 2010). In both species, HSF1 has been shown to bind and regulate *hsp70* expression in oocytes (Gordon et al. 1997; Åkerfelt et al. 2010). In mice, a knockout of maternal *hsf1* resulted in reduced viability of oocytes and zygotes, possibly due to disruptions in mitochondrial function and *hsp90α* expression (Christians et al. 2000; Metchat et al. 2009; Bierkamp et al. 2010). In zebrafish, an expression and proteomics analysis of maternally derived RNA and proteins from mature eggs found 329 unique mRNA transcripts (Knoll-Gellida et al. 2006). While the study identified transcripts for *hsp60*, *hsp8*, *hsp5* and *hsp90β*, they did not report identifying *hsf1*, *hsf2* or *hsp70* maternal RNA. However, it must be noted that the study only discussed an analysis of the 58 genes that culminated in the 17% most abundant transcripts.

While the presence of maternal *hsf1* and *hsp70* mRNA has been reported in oocytes from several species, the current dogma surrounding *hsf2* is that it is not a maternally expressed transcript and expression begins only at the eight cell stage in mice (Kallio et al. 2002; Abane and Mezger 2010). The results presented in this thesis demonstrating that *hsf2* is present at high

levels in zebrafish eggs is, to the best of my knowledge, the first description of maternally expressed *hsf2*. The reason for the discrepancy in the presence of maternal *hsf2* between zebrafish and mammals is unknown. It is tempting to suggest that this may be a characteristic of species that develop *in ovo* and it would be interesting to try and determine if *hsf2* is expressed maternally in avian species. In *Xenopus* oocytes, no HSF:HSE complexes could be identified with an HSF2 antiserum (Gordon et al. 1997). This result does not negate the presence of *hsf2* mRNA in *Xenopus* oocytes but does suggests that maternal HSF2 is either not present in *Xenopus* oocytes or that HSF2 was not actively trimerized and complexed with HSEs at the time when the analysis was performed. The maternal expression of *hsf2* mRNA may also be a characteristic specific to telosts. Further investigation into the expression of *hsfs* in oocytes of teleosts and other species is required to determine to what extent *hsf2* is maternally expressed in the animal kingdom.

Another interesting observation about the presence of maternal *hsp70*, *hsf1* and *hsf2* in zebrafish eggs is that there is a significant increase in the mRNA levels for these genes in mature eggs vs immature eggs or the ovary. This suggests that these maternal transcripts are deposited into the oocyte at late stages of folliculogenesis most likely during the maturation stage after the vitellogenic stage. Reviews on maternal RNA deposits in several species including zebrafish have noted that maternal RNA is capable of remaining stable in the oocyte for a prolonged time, suggesting transcripts remain for days, weeks or even months prior to being degraded. (Knoll-Gellida et al. 2006; Tadros and Lipshitz 2009). These statements could be interpreted to suggest that maternal RNA is deposited in eggs at an early stage of oogenesis. Our results would suggest that this is a concept that requires further study as high levels of maternal RNA for all transcripts are only observed in mature oocytes. As well, the majority of maternal RNA appears to be degraded by approximately 8-10 hpf which suggests that *hsf* maternal transcripts are only required to be stable from the point of being deposited into a mature follicle to degradation at 8 hpf.

6.0 CONCLUSIONS

Several conclusions about HSF expression and function in zebrafish development can be drawn from the data presented in this thesis. We have identified that zebrafish express two alternatively spliced transcripts of the *hsf4* gene, with the *hsf4a* transcript comprised of sequence from eight exons and the *hsf4b* transcript comprised of the same eight plus an additional four exons. The similarity observed in the amino acid composition of protein domains between zebrafish HSF4 and HSF1, as well as the close relationship of these proteins observed in the phylogenetic analysis, suggest that zebrafish *hsf4* evolved from an *hsf1* duplication event and that some domains in the protein remains structurally more similar to HSF1 than the more distant mammalian HSF4. The truncated structure of the HSF4a protein as well as the presence of an HR-C domain in the HSF4b isoform suggests that zebrafish and mammalian HSF4 proteins may not share all the same functions.

Zebrafish *hsf4* is developmentally expressed in zebrafish embryos and is predominantly expressed in adult eye tissue. This evidence strongly suggests HSF4 has important roles in eye development as has been reported for mammals. While a basal level of *hsf4b* can be detected in all adult tissues tested, mRNA levels are undetectable for both *hsf4* transcripts prior to 60 hpf. This observation, combined with the surge of *hsf4b* expression at 80 hpf, suggests that HSF4 has a more prominent role in later stages of eye development in zebrafish and may be involved in regulating genes crucial for the development of functional vision.

Further analysis of tissue specific expression of zebrafish HSFs determined that there are significantly higher levels of *hsf1*, *hsf2* and *hsp70* mRNA in eggs that have been squeezed from zebrafish than in any other tissue analyzed including the ovary. This result suggests that the mRNAs for these three genes come from the maternal genome and are deposited in eggs at later stages of oogenesis. The presence of maternal *hsf1* and *hsp70* RNA has been reported in mammalian eggs. However, this is the first time the presence of maternal RNA for *hsf2* has been identified.

The role HSFs play in regulating different developmental mechanisms is a relatively young area of research. To date most analysis of HSF function and conclusions drawn about the roles HSFs have in development have been based on observations made primarily on HSFs from mammalian species (Abane and Mezger 2010). Currently the HSF4 protein has been globally characterized as a HSF that lacks an HR-C domain. However, the observation that zebrafish

HSF4b and avian HSF4 isoforms both have an HR-C domain means that the current definition of HSF4 protein structure will have to be redefined. Analysis of *hsf2* expression presented in this thesis has also demonstrated that the high level of *hsf2* mRNA observed in mammalian testis is not observed in the zebrafish. As well, the expression of maternal *hsf2* RNA in zebrafish eggs is, again, unlike what has been observed in mammals. The data presented in this thesis suggest that further analysis of HSF expression and function from multiple classes of animals is needed to fully understand the evolution of these proteins and provide a better understanding of the roles HSFs have during development.

7.0 FUTURE DIRECTIONS

While a full characterization of *hsf4* expression in zebrafish tissues and during development was presented in this study, very little investigation was performed on the activity of HSF4 as a transcription factor in zebrafish. One of the limitations in these studies was the lack of antibodies that could specifically detect individual HSFs in zebrafish embryos. The antibodies currently available to us were developed against synthetic peptides 13-15 amino acids long. Unfortunately, Western blot analysis of HSF proteins in zebrafish embryos with these antibodies was not specific enough to derive any conclusions from the experiments. However, during the initial characterization of the zebrafish antibodies, full length HSF1, HSF2 and HSF4 proteins were isolated from a bacterial expression system for the purpose of determining antibody specificity. The use of these full length proteins as antigens for antibody development would likely generate antibodies with better specificity. New antibodies with greater specificity would allow us to characterize HSF protein expression and localization using Western blot, immunohistochemistry or chromatin immunoprecipitation assay techniques.

The results from the thesis determined that the majority of *hsf4* expression occurs in the zebrafish eye, similar to what has been observed in mammals. However, phylogenic analysis of HSF proteins across several species would suggest that zebrafish HSF4 did not diverge from HSF1 to the same extent that mammalian HSF4 diverged. As well, *in silico* comparisons of the zebrafish HSF4 sequence to other HSFs determined that functional domains of this protein (DBD and HR-C) more closely resemble the amino acid composition of HSF1 than mammalian HSF4. These observations suggest that the downstream targets regulated by HSF4 in zebrafish may not be entirely the same as those identified in mammals. One way to determine if there is a functional difference between the HSF4 transcription factors in different species is to determine if HSF4 regulates the same genes in both zebrafish and mammals. The quickest way to study functional differences is to knockdown *hsf4* using morpholino antisense technology and determine the effect on gene expression of known downstream targets of mammalian HSF4 such as *fgf* family members, *α* and *γ-crystallins* and potentially *hsp70*. Preliminary experiments using MO were performed in this study with inconclusive results occurring most probably due to a solubility problem with the Lissamine tag on the MO. Therefore, future *hsf4* knockdown experiments could be performed using *hsf4*-MO linked to a different marker such as Fluorsceine. There are several limitations to the analysis of *hsf4* expression that can be determined with MO

technology specifically the late onset of detectable *hsf4* mRNA in embryos. In our lab, the effects of gene knockdown by MO begin to be reduced by 60 hpf when MO begin to be diluted and phenotypes begin to disappear by 72 hpf. This leaves only a very narrow window to observe any effects of *hsf4*-MO on zebrafish development.

Recently zinc finger nuclease and transcription activator-like effector nuclease (TALEN) technology has been used successfully to introduce heritable mutations and create gene specific zebrafish knockout lines (Doyon et al. 2008; Ekker 2008; Meng et al. 2008; Foley et al. 2009; Bedell et al. 2012). While generating *hsf4* knockout lines for zebrafish using this technology would take a longer time to develop, they would be more useful for studying the function of this gene throughout zebrafish development as well as specifically looking at effects of *hsf4* knockout on eye and lens physiology. Mutations of specific functional domains could also be introduced using this technology. Unlike mammals, the zebrafish HSF4b protein has an HR-C inhibition of trimerization domain. It would be interesting to see how the function of the zebrafish protein was altered if this domain was made non-functional by mutation. Microarray analysis of *hsf4* binding sites using specific tissues, such as the eye, from both *hsf4* knockout strains and wildtype zebrafish would allow for a detailed identification of HSF4 targets in zebrafish. As this thesis suggests, there may be differences between mammalian and zebrafish HSF4 target genes. Identification of HSF4 targets in the zebrafish lens using chromatin immunoprecipitation techniques such as ChIP-on-chip or ChIP sequencing. Comparison of the results to those obtained from mouse lens would help to isolate unique zebrafish HSF4 targets for further studies (Fujimoto et al. 2008). Given that *hsf4a* is expressed in the testis, it is likely some of these novel targets would be involved in spermatogenesis. Studies investigating the effects *hsf4* knockdown on sperm quality would be interesting to pursue.

In mammals, HSF4a functions as a repressor of gene function (Tanabe et al. 1999). It is therefore likely that one of the zebrafish isoforms may also act as a repressor. Overexpression of each *hsf4* transcript using *in vitro* transcribed mRNA for each splice variant independently microinjected into embryos, could allow us to determine if either HSF4 isoform is capable of repressing gene expression of potential target gene candidates (Finkbeiner et al. 2011). The similarity between both isoforms of zebrafish HSF4 is high, and therefore it would likely be impossible to generate isoform specific knockout lines for HSF4. However, analysis of the effect of each HSF4 isoform on zebrafish development could be pursued by using the same

technique used for overexpression in wild type fish to introduce mRNA from a single *hsf4* splice variant into zebrafish embryos with inactivated *hsf4* genes generated using TALEN technology.

It has been previously shown that *hsf1* mRNA exists in embryos as maternal RNA (Gordon et al. 1997; Christians et al. 2000; Metchat et al. 2009; Bierkamp et al. 2010; Åkerfelt et al. 2010). The results presented in this thesis demonstrate that *hsf2* exists in embryos as maternal RNA as well. Again, this highlights a difference in *hsf* expression between zebrafish and mammals. It would be interesting to determine if the presence of multiple *hsfs* as maternal RNA is specific to aquatic species or if they also exist as maternal RNA in other species that develop *in ovo*. Quantitation of mRNA levels for *hsf3* in unfertilized eggs from chicken and *Xenopus* could provide some insight into this question.

Analysis of the shared synteny between zebrafish *hsf4* and other mammalian Hsfs, specifically *HSFY*, may also be interesting. This type of analysis may help determine if zebrafish *hsf4* is related to the mammalian *HSFY* or, alternately, may help to identify additional zebrafish genes on different chromosomes that are similar to *HSFY*.

8.0 REFERENCES

Abane R, Mezger V (2010) Roles of heat shock factors in gametogenesis and development. FEBS Journal 277 (20):4150-4172

Ahn SG, Liu PCC, Klyachko K, Morimoto RI, Thiele DJ (2001) The loop domain of heat shock transcription factor 1 dictates DNA-binding specificity and responses to heat stress. Genes & Development 15 (16):2134-2145

Ahn SG, Thiele DJ (2003) Redox regulation of mammalian heat shock factor 1 is essential for Hsp gene activation and protection from stress. Genes & Development 17 (4):516-528

Airaksinen S, Jokilehto T, Råbergh CMI, Nikinmaa M (2003a) Heat- and cold-inducible regulation of HSP70 expression in zebrafish ZF4 cells. Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology 136 (2):275-282

Airaksinen S, Råbergh CMI, Lahti A, Kaatrasalo A, Sistonen L, Nikinmaa M (2003b) Stressor-dependent regulation of the heat shock response in zebrafish, *Danio rerio*. Comparative Biochemistry and Physiology A-Molecular & Integrative Physiology 134 (4):839-846

Åkerfelt M, Henriksson E, Laiho A, Vihervaara A, Rautoma K, Kotaja N, Sistonen L (2008) Promoter ChIP-chip analysis in mouse testis reveals Y chromosome occupancy by HSF2. Proceedings of the National Academy of Sciences of the United States of America 105 (32):11224-11229

Åkerfelt M, Morimoto RI, Sistonen L (2010) Heat shock factors: integrators of cell stress, development and lifespan. Nature Reviews Molecular Cell Biology 11 (8):545-555

Åkerfelt M, Trouillet D, Mezger V, Sistonen L (2007) Heat shock factors at a crossroad between stress and development. Annals of the New York Academy of Sciences 1113 (1):15-27

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215:403-410

Amin J, Ananthan J, Voellmy R (1988) Key features of heat shock regulatory elements. Molecular and Cellular Biology 8 (9):3761-3769

Anderson JL, Rodriguez Mari A, Braasch I, Amores A, Hohenlohe P, Batzel P, Postlethwait JH (2012) Multiple sex-associated regions and a putative sex chromosome in zebrafish revealed by RAD mapping and population genomics. PLoS ONE 7 (7):e40701

Anonymous (2006) TOPO TA cloning kit for sequencing. Five-minute cloning of *Taq* polymerase-amplified PCR products for sequencing. Life Technologies Corporation, Carlsbad, California

Anonymous (2008) DIG application manual for filter hybridization. Roche Diagnostic GmbH, Mannheim, Germany

Anonymous (2009) REST 2009 software user guide, 12/2009. Qiagen GmbH, Netherlands

Anonymous (2010a) pRSET A, B, and C. For high-level expression of recombinant proteins in *E. coli*. Life Technologies Corporation, Carlsbad, California

Anonymous (2010b) Real-time PCR applications guide. Bio-Rad Laboratories, Hercules, California

Awasthi N, Wagner BJ (2005) Upregulation of heat shock protein expression by proteasome inhibition: An antiapoptotic mechanism in the lens. Investigative Ophthalmology & Visual Science 46 (6):2082-2091

Barna J, Princz A, Kosztelnik M, Hargitai B, Takács-Vellai K, Vellai T (2012) Heat shock factor-1 intertwines insulin/IGF-1, TGF-ß and cGMP signaling to control development and aging. BMC Developmental Biology 12:32

Beaupre CE, Tressler CJ, Beaupre SJ, Morgan JLM, Bottje WG, Kirby JD (1997) Determination of testis temperature rhythms and effects of constant light on testicular function in the domestic fowl (*Gallus domesticus*). Biology of Reproduction 56 (6):1570-1575

Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug RG, Tan WF, Penheiter SG, Ma AC, Leung AYH, Fahrenkrug SC, Carlson DF, Voytas DF, Clark KJ, Essner JJ, Ekker SC (2012) *In vivo* genome editing using a high-efficiency TALEN system. Nature 491 (7422):114-118

Beier H, Grimm M (2001) Misreading of termination codons in eukaryotes by natural nonsense suppressor tRNAs. Nucleic Acids Research 29 (23):4767-4782

Bensaude O, Babinet C, Morange M, Jacob F (1983) Heat-shock proteins, 1st major products of zygotic gene activity in mouse embryo. Nature 305 (5932):331-333

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2005) GenBank. Nucleic Acids Research 33 (Database issue):D34-38

Berry MJ, Banu L, Harney JW, Larsen PR (1993) Functional-characterization of the eukaryotic SECIS elements which direct selenocysteine insertion at UGA codons. EMBO Journal 12 (8):3315-3322

Bienz B, Zakuthouri R, Givol D, Oren M (1984) Analysis of the gene coding for the murine cellular tumor-antigen p53. EMBO Journal 3 (9):2179-2183

Bienz M (1984) *Xenopus hsp* 70 genes are constitutively expressed in injected oocytes. EMBO Journal 3 (11):2477-2483

Bierkamp C, Luxey M, Metchat A, Audouard C, Dumollard R, Christians E (2010) Lack of maternal Heat Shock Factor 1 results in multiple cellular and developmental defects, including mitochondrial damage and altered redox homeostasis, and leads to reduced survival of mammalian oocytes and embryos. Developmental Biology 339 (2):338-353

Björk JK, Sandqvist A, Elsing AN, Kotaja N, Sistonen L (2010) miR-18, a member of Oncomir-1, targets heat shock transcription factor 2 in spermatogenesis. Development 137 (19):3177-3184

Björk JK, Sistonen L (2010) Regulation of the members of the mammalian heat shock factor family. FEBS Journal 277 (20):4126-4139

Blechinger SR, Evans TG, Tang PT, Kuwada JY, Warren JT, Krone PH (2002) The heat-inducible zebrafish *hsp70* gene is expressed during normal lens development under non-stress conditions. Mechanisms of Development 112 (1-2):213-215

Bu L, Jin Y, Shi Y, Chu R, Ban A, Eiberg H, Andres L, Jiang H, Zheng G, Qian M, Cui B, Xia Y, Liu J, Hu L, Zhao G, Hayden MR, Kong X (2002) Mutant DNA-binding domain of HSF4 is associated with autosomal dominant lamellar and Marner cataract. Nature Genetics 31 (3):276-278

Bustin SA (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. Journal of Molecular Endocrinology 25 (2):169-193

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT (2009) The MIQE Guidelines: *M*inimum *I*nformation for Publication of *Q*uantitative real-time PCR *E*xperiments. Clinical Chemistry 55 (4):611-622

Castellano S, Lobanov AV, Chapple C, Novoselov SV, Albrecht M, Hua D, Lescure A, Lengauer T, Krol A, Gladyshev VN, Guigo R (2005) Diversity and functional plasticity of eukaryotic selenoproteins: identification and characterization of the SelJ family. Proceedings of the National Academy of Sciences of the United States of America 102 (45):16188-16193

Castellano S, Novoselov SV, Kryukov GV, Lescure A, Blanco E, Krol A, Gladyshev VN, Guigo R (2004) Reconsidering the evolution of eukaryotic selenoproteins: a novel nonmammalian family with scattered phylogenetic distribution. EMBO Reports 5 (1):71-77

Chang YH, Ostling P, Akerfelt M, Trouillet D, Rallu M, Gitton Y, El Fatilmy R, Fardeau V, Le Crom S, Morange M, Sistonen L, Mezger V (2006) Role of heat-shock factor 2 in cerebral cortex formation and as a regulator of p35 expression. Genes & Development 20 (7):836-847

Chen M, Manley JL (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nature Reviews Molecular Cell Biology 10 (11):741-754

Cheng SH, Shakespeare T, Mui R, White TW, Valdimarsson G (2004) Connexin 48.5 is required for normal cardiovascular function and lens development in zebrafish embryos. Journal of Biological Chemistry 279 (35):36993-37003

Choi MR, Jung KH, Park JH, Das ND, Chung MK, Choi IG, Lee BC, Park KS, Chai YG (2011) Ethanol-induced small heat shock protein genes in the differentiation of mouse embryonic neural stem cells. Archives of Toxicology 85 (4):293-304

Christians E, Davis AA, Thomas SD, Benjamin IJ (2000) Embryonic development: Maternal effect of Hsf1 on reproductive success. Nature 407 (6805):693-694

Clelland E, Peng C (2009) Endocrine/paracrine control of zebrafish ovarian development. Molecular and Cellular Endocrinology 312 (1-2):42-52

Commichau FM, Stulke J (2008) Trigger enzymes: bifunctional proteins active in metabolism and in controlling gene expression. Molecular Microbiology 67 (4):692-702

Condon TP (1999) Generation of single-stranded DNA hybridization probes by PCR using a short, synthetic DNA template. Biotechniques 26 (1):18-19

Cui X, Wang L, Zhang J, Du R, Liao S, Li D, Li C, Ke T, Li DW, Huang H, Yin Z, Tang Z, Liu M (2013) HSF4 regulates DLAD expression and promotes lens de-nucleation. Biochimica et Biophysica Acta-Molecular Basis of Disease 1832 (8):1167-1172

Cui X, Zhang J, Du R, Wang L, Archacki S, Zhang Y, Yuan M, Ke T, Li H, Li D, Li C, Li DW, Tang Z, Yin Z, Liu M (2012) HSF4 is involved in DNA damage repair through regulation of Rad51. Biochimica et Biophysica Acta-Molecular Basis of Disease 1822 (8):1308-1315

Cutforth T, Rubin GM (1994) Mutations in *Hsp83* and *cdc37* impair signaling by the sevenless receptor tyrosine kinase in Drosophila. Cell 77 (7):1027-1036

Dai C, Whitesell L, Rogers AB, Lindquist S (2007) Heat shock factor 1 is a powerful multifaceted modifier of carcinogenesis. Cell 130 (6):1005-1018

de Rooija D, de Boerb P (2003) Specific arrests of spermatogenesis in genetically modified and mutant mice. Cytogenetic and Genome Research 103:267-276

de Thonel A, Le Mouel A, Mezger V (2012) Transcriptional regulation of small HSP-HSF1 and beyond. International Journal of Biochemistry & Cell Biology 44 (10):1593-1612

Dereeper A, Audic S, Claverie JM, Blanc G (2010) BLAST-EXPLORER helps you building datasets for phylogenetic analysis. BMC Evolutionary Biology 10:8

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Research 36 (Web Server issue):W465-W469

Dix DJ, Allen JW, Collins BW, Mori C, Nakamura N, Poorman-Allen P, Goulding EH, Eddy EM (1996) Targeted gene disruption of Hsp70-2 results in failed meiosis, germ cell apoptosis, and male infertility. Proceedings of the National Academy of Sciences of the United States of America 93 (8):3264-3268

Dix DJ, Garges JB, Hong RL (1998) Inhibition of *hsp70-1* and *hsp70-3* expression disrupts preimplantation embryogenesis and heightens embryo sensitivity to arsenic. Molecular Reproduction and Development 51 (4):373-380

Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, Amora R, Hocking TD, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Amacher SL (2008) Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nature Biotechnology 26 (6):702-708

Draper BW, Morcos PA, Kimmel CB (2001) Inhibition of zebrafish *fgf8* Pre-mRNA splicing with morpholino oligos: A quantifiable method for gene knockdown. Genesis 30 (3):154-156

Du SJ, Li H, Bian Y, Zhong Y (2008) Heat-shock protein 90α1 is required for organized myofibril assembly in skeletal muscles of zebrafish embryos. Proceedings of the National Academy of Sciences of the United States of America 105 (2):554-559

Easter SS, Nicola GN (1996) The development of vision in the zebrafish (*Danio rerio*). Developmental Biology 180 (2):646-663

Ekker SC (2008) Zinc finger-based knockout punches for zebrafish genes. Zebrafish 5 (2):121-123

Ekker SC, Larson JD (2001) Morphant technology in model developmental systems. Genesis 30 (3):89-93

Emanuel JR (1991) Simple and efficient system for synthesis of non-radioactive nucleic acid hybridization probes using PCR. Nucleic Acids Research 19 (10):2790

Engelhardt A, Wohlke A, Distl O (2007) Evaluation of canine heat-shock transcription factor 4 as a candidate for primary cataracts in English Cocker Spaniels and wire-haired Kromfohrlanders. Journal of Animal Breeding and Genetics 124 (4):242-245

Enoki Y, Mukoda Y, Furutani C, Sakurai H (2010) DNA-binding and transcriptional activities of human HSF4 containing mutations that associate with congenital and age-related cataracts. Biochimica et Biophysica Acta-Molecular Basis of Disease 1802 (9):749-753

Enoki Y, Sakurai H (2011) Diversity in DNA recognition by heat shock transcription factors (HSFs) from model organisms. FEBS Letters 585 (9):1293-1298

Etard C, Roostalu U, Strahle U (2008) Shuttling of the chaperones Unc45b and Hsp90a between the A band and the Z line of the myofibril. Journal of Cell Biology 180 (6):1163-1175

Evans TG (2006) Expression, function and regulation of heat shock protein 70 (*hsp70*) gene during normal zebrafish (*Danio rerio*) embryogenesis. PhD Dissertation, University of Saskatchewan

Evans TG, Belak Z, Ovsenek N, Krone PH (2007) Heat shock factor 1 is required for constitutive Hsp70 expression and normal lens development in embryonic zebrafish. Comparative Biochemistry and Physiology A-Molecular & Integrative Physiology 146:131-140

Evans TG, Krone PH (2005) Heat shock proteins: molecular chaperones as critical players in normal eukaryotic development. Trends in Developmental Biology 1:95-105

Evans TG, Yamamoto Y, Jeffery WR, Krone PH (2005) Zebrafish *Hsp70* is required for embryonic lens formation. Cell Stress & Chaperones 10:66-78

Ferlin A, Moro E, Rossi A, Dallapiccola B, Foresta C (2003) The human Y chromosome's azoospermia factor b (AZFb) region: sequence, structure, and deletion analysis in infertile men. Journal of Medical Genetics 40 (1):18-24

Finckbeiner S, Ko PJ, Carrington B, Sood R, Gross K, Dolnick B, Sufrin J, Liu P (2011) Transient knockdown and overexpression reveal a developmental role for the zebrafish *enosf1b* gene. Cell & Bioscience 1:32

Fiorenza MT, Farkas T, Dissing M, Kolding D, Zimarino V (1995) Complex expression of murine heat shock transcription factors. Nucleic Acids Research 23 (3):467-474

Foley JE, Yeh JR, Maeder ML, Reyon D, Sander JD, Peterson RT, Joung JK (2009) Rapid mutation of endogenous zebrafish genes using zinc finger nucleases made by Oligomerized Pool ENgineering (OPEN). PLoS ONE 4 (2):e4348

Foresta C, Bettella A, Moro E, Roverato A, Merico M, Ferlin A (2001) Sertoli cell function in infertile patients with and without microdeletions of the azoospermia factors on the Y chromosome long arm. Journal of Clinical Endocrinology & Metabolism 86 (6):2414-2419

Frejtag W, Zhang Y, Dai R, Anderson MG, Mivechi NF (2001) Heat shock factor-4 (HSF-4a) represses basal transcription through interaction with TFIIF. Journal of Biological Chemistry 276 (18):14685-14694

Fujimoto M, Hayashida N, Katoh T, Oshima K, Shinkawa T, Prakasam R, Tan K, Inouye S, Takii R, Nakai A (2010) A novel mouse HSF3 has the potential to activate nonclassical heat-shock genes during heat shock. Molecular Biology of the Cell 21 (1):106-116

Fujimoto M, Izu H, Seki K, Fukuda K, Nishida T, Yamada S-i, Kato K, Yonemura S, Inouye S, Nakai A (2004) HSF4 is required for normal cell growth and differentiation during mouse lens development. EMBO Journal 23:4297-4306

Fujimoto M, Nakai A (2010) The heat shock factor family and adaptation to proteotoxic stress. FEBS Journal 277 (20):4112-4125

Fujimoto M, Oshima I, Shinkawa T, Wang BB, Inouye S, Hayashida N, Takii R, Nakai A (2008) Analysis of HSF4 binding regions reveals its necessity for gene regulation during development and heat shock response in mouse lenses. Journal of Biological Chemistry 283 (44):29961-29970

Gestri G, Link BA, Neuhauss SCF (2012) The visual system of zebrafish and its use to model human ocular diseases. Developmental Neurobiology 72 (3):302-327

Gidalevitz T, Prahlad V, Morimoto RI (2011) The stress of protein misfolding: From single cells to multicellular organisms. Cold Spring Harbor Perspectives in Biology 3:a009704

Goishi K, Shimizu A, Najarro G, Watanabe S, Rogers R, Zon LI, Klagsbrun M (2006) αA-crystallin expression prevents γ-crystallin insolubility and cataract formation in the zebrafish *cloche* mutant lens. Development 133 (13):2585-2593

Gordon S, Bharadwaj S, Hnatov A, Ali A, Ovsenek N (1997) Distinct stress-inducible and developmentally regulated heat shock transcription factors in *Xenopus* oocytes. Developmental Biology 181 (1):47-63

Green J, Paget MS (2004) Bacterial redox sensors. Nature Reviews Microbiology 2 (12):954-966

Greiling TMS, Clark JI (2008) The transparent lens and cornea in the mouse and zebra fish eye. Seminars in Cell & Developmental Biology 19 (2):94-99

Greiling TMS, Clark JI (2009) Early lens development in the zebrafish: A three-dimensional time-lapse analysis. Developmental Dynamics 238 (9):2254-2265

Greiling TMS, Clark JI (2012) Chapter one - New insights into the mechanism of lens development using zebra fish. In: Kwang WJ (ed) International review of cell and molecular biology, vol. 296. Academic Press, San Diego, California, pp 1-61

Greiling TMS, Houck SA, Clark JI (2009) The zebrafish lens proteome during development and aging. Molecular Vision 15:2313-2325

Grune T, Catalgol B, Licht A, Ermak G, Pickering AM, Ngo JK, Davies KJA (2011) HSP70 mediates dissociation and reassociation of the 26S proteasome during adaptation to oxidative stress. Free Radical Biology and Medicine 51 (7):1355-1364

Gupta J, Tikoo K (2012) Involvement of insulin-induced reversible chromatin remodeling in altering the expression of oxidative stress-responsive genes under hyperglycemia in 3T3-L1 preadipocytes. Gene 504 (2):181-191

Gupta T, Mullins MC (2010) Dissection of organs from the adult zebrafish. Journal of Visualized Experiments 37:e1717

Harrison CJ, Bohm AA, Nelson HC (1994) Crystal structure of the DNA binding domain of the heat shock transcription factor. Science 263 (5144):224-227

Hashikawa N, Yamamoto N, Sakurai H (2007) Different mechanisms are involved in the transcriptional activation by yeast heat shock transcription factor through two different types of heat shock elements. Journal of Biological Chemistry 282:10333-10340

Hawkins TA, Haramis AP, Etard C, Prodromou C, Vaughan CK, Ashworth R, Ray S, Behra M, Holder N, Talbot WS, Pearl LH, Strahle U, Wilson SW (2008) The ATPase-dependent chaperoning activity of Hsp90a regulates thick filament formation and integration during skeletal muscle myofibrillogenesis. Development 135 (6):1147-1156

He SY, Pirity MK, Wang WL, Wolf L, Chauhan BK, Cveklova K, Tamm ER, Ashery-Padan R, Metzger D, Nakai A, Chambon P, Zavadil J, Cvekl A (2010) Chromatin remodeling enzyme Brg1 is required for mouse lens fiber cell terminal differentiation and its denucleation. Epigenetics & Chromatin 3:21

Hofmann K, Baron M (2008) Boxshade version 3.21. http://sourceforge.net/projects/boxshade. Accessed 14 January 2014

Hong Y, Sarge KD (1999) Regulation of protein phosphatase 2A activity by heat shock transcription factor 2. Journal of Biological Chemistry 274 (19):12967-12970

Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren K, Matthews L, McLaren S, Sealy I, Caccamo M, Churcher C, Scott C, Barrett JC, Koch R, Rauch GJ, White S, Chow W et al. (2013) The zebrafish reference genome sequence and its relationship to the human genome. Nature 496 (7446):498-503

Hu Y, Mivechi NF (2006) Association and regulation of heat shock transcription factor 4b with both extracellular signal-regulated kinase mitogen-activated protein kinase and dual-specificity tyrosine phosphatase DUSP26. Molecular and Cellular Biology 26 (8):3282-3294

Hu YZ, Zhang J, Li S, Wang C, Chu L, Zhang Z, Ma Z, Wang M, Jiang Q, Liu G, Qi Y, Ma Y (2013) The transcription activity of heat shock factor 4b is regulated by FGF2. International Journal of Biochemistry & Cell Biology 45 (2):317-325

Imai F, Yoshizawa A, Fujimori-Tonou N, Kawakami K, Masai I (2010) The ubiquitin proteasome system is required for cell proliferation of the lens epithelium and for differentiation of lens fiber cells in zebrafish. Development 137 (19):3257-3268

Inouye S, Izu H, Takaki E, Suzuki H, Shirai M, Yokota Y, Ichikawa H, Fujimoto M, Nakai A (2004) Impaired IgG production in mice deficient for heat shock transcription factor 1. Journal of Biological Chemistry 279 (37):38701-38709

Izu H, Inouye S, Fujimoto M, Shiraishi K, Naito K, Nakai A (2004) Heat shock transcription factor 1 is involved in quality-control mechanisms in male germ cells. Biology of Reproduction 70 (1):18-24

Jing Z, Gangalum RK, Lee JZ, Mock D, Bhat SP (2013) Cell-type-dependent access of HSF1 and HSF4 to αB-crystallin promoter during heat shock. Cell Stress & Chaperones 18 (3):377-387

Johnson MH, Everitt BJ (1995) Essential reproduction, 4th edn. Wiley-Blackwell, San Francisco, California

Kallio M, Chang Y, Manuel M, Alastalo T-P, Rallu M, Gitton Y, Pirkkala L, Loones M-T, Paslaru L, Larney S, Hiard S, Morange M, Sistonen L, Mezger V (2002) Brain abnormalities, defective meiotic chromosome synapsis and female subfertility in HSF2 null mice. EMBO Journal 21 (11):2591-2601

Kane DA, Kimmel CB (1993) The zebrafish midblastula transition. Development 119 (2):447-456

Kanei-Ishii C, Tanikawa J, Nakai A, Morimoto RI, Ishii S (1997) Activation of heat shock transcription factor 3 by c-myb in the absence of cellular stress. Science 277 (5323):246-248

Kawazoe Y, Tanabe M, Sasai N, Nagata K, Nakai A (1999) HSF3 is a major heat shock responsive factor during chicken embryonic development. European Journal of Biochemistry 265 (2):688-697

Ke T, Wang QK, Ji B, Wang X, Liu P, Zhang X, Tang Z, Ren X, Liu M (2006) Novel *HSF4* mutation causes congenital total white cataract in a Chinese family. American Journal of Ophthalmology 142 (2):298-303

Kichine E, Roze V, Di Cristofaro J, Taulier D, Navarro A, Streichemberger E, Decarpentrie F, Metzler-Guillemain C, Levy N, Chiaroni J, Paquis-Flucklinger V, Fellmann F, Mitchell MJ (2012) *HSFY* genes and the *P4* palindrome in the *AZFb* interval of the human Y chromosome are not required for spermatocyte maturation. Human Reproduction 27 (2):615-624

Kim SA, Yoon JH, Ahn SG (2012) Heat shock factor 4a (HSF4a) represses HSF2 expression and HSF2-mediated transcriptional activity. Journal of Cellular Physiology 227 (1):1-6

Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF (1995) Stages of embryonic development of the zebrafish. Developmental Dynamics 203:253-310

Kinoshita K, Shinka T, Sato Y, Kurahashi H, Kowa H, Chen G, Umeno M, Toida K, Kiyokage E, Nakano T, Ito S, Nakahori Y (2006) Expression analysis of a mouse orthologue of HSFY, a candidate for the azoospermic factor on the human Y chromosome. Journal of Medical Investigation 53:117-122

Knoll-Gellida A, Andre M, Gattegno T, Forgue J, Admon A, Babin PJ (2006) Molecular phenotype of zebrafish ovarian follicle by serial analysis of gene expression and proteomic profiling, and comparison with the transcriptomes of other animals. BMC Genomics 7:46

Konat GW (1996) Generation of high efficiency ssDNA hybridization probes by linear polymerase chain reaction (LPCR). Scanning Microscopy Supplement 10:57-60

Kroeger PE, Morimoto RI (1994) Selection of new HSF1 and HSF2 DNA-binding sites reveals difference in trimer cooperativity. Molecular and Cellular Biology 14 (11):7592-7603

Krone PH, Evans TG, Blechinger SR (2003) Heat shock gene expression and function during zebrafish embryogenesis. Seminars in Cell & Developmental Biology 14:267-274

Krone PH, Lele Z, Sass JB (1997) Heat shock genes and the heat shock response in zebrafish embryos. Biochemistry and Cell Biology 75 (5):487-497

Krone PH, Sass JB, Lele Z (1997) Heat shock protein gene expression during embryonic development of the zebrafish. Cellular and Molecular Life Sciences 53 (1):122-129

Kryukov GV, Gladyshev VN (2000) Selenium metabolism in zebrafish: multiplicity of selenoprotein genes and expression of a protein containing 17 selenocysteine residues. Genes to Cells 5:1049-1060

Laemmli UK (1970) Cleavage of structural proteins during assembly of head of bacteriophage-T4. Nature 227 (5259):680-685

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) ClustalW and ClustalX version 2. Bioinformatics 23 (21):2947-2948

Lawn RM, Efstratiadis A, Oconnell C, Maniatis T (1980) The nucleotide-sequence of the human β-globin gene. Cell 21 (3):647-651

Leal MC, Cardoso ER, Nobrega RH, Batlouni SR, Bogerd J, Franca LR, Schulz RW (2009) Histological and stereological evaluation of zebrafish (*Danio rerio*) spermatogenesis with an emphasis on spermatogonial generations. Biology of Reproduction 81 (1):177-187

Lele Z, Engel S, Krone PH (1997) *hsp47* and *hsp70* gene expression is differentially regulated in a stress- and tissue-specific manner in zebrafish embryos. Developmental Genetics 21 (2):123-133

Lele Z, Hartson SD, Martin CC, Whitesell L, Matts RL, Krone PH (1999) Disruption of zebrafish somite development by pharmacologic inhibition of Hsp90. Developmental Biology 210 (1):56-70

Li CW, Zhou R, Ge W (2012) Differential regulation of gonadotropin receptors by bone morphogenetic proteins in the zebrafish ovary. General and Comparative Endocrinology 176 (3):420-425

Liew WC, Bartfai R, Lim Z, Sreenivasan R, Siegfried KR, Orban L (2012) Polygenic sex determination system in zebrafish. PLoS ONE 7 (4):e34397

Lindquist S (1986) The heat-shock response. Annual Review of Biochemistry 55:1151-1191

Lindquist S, Craig EA (1988) The heat-shock proteins. Annual Review of Genetics 22:631-677

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\triangle\triangle C}{}_{T}$ method. Methods 25 (4):402-408

Loison F, Debure L, Nizard P, Le Goff P, Michel D, Le Drean Y (2006) Up-regulation of the clusterin gene after proteotoxic stress: implication of HSF1-HSF2 heterocomplexes. Biochemical Journal 395:223-231

Lupas A, Vandyke M, Stock J (1991) Predicting coiled coils from protein sequences. Science 252 (5009):1162-1164

Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH et al. (2009) CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Research 37 (Database issue):D205-D210

Marshall OJ, Harley VR (2001) Identification of an interaction between SOX9 and HSP70. FEBS Letters 496 (2-3):75-80

Mathew ANU, Morimoto RI (1998) Role of the heat-shock response in the life and death of proteins. Annals of the New York Academy of Sciences 851 (1):99-111

Matsuoka Y, Kubota H, Adachi E, Nagai N, Marutani T, Hosokawa N, Nagata K (2004) Insufficient folding of type IV collagen and formation of abnormal basement membrane-like structure in embryoid bodies derived from Hsp47-null embryonic stem cells. Molecular Biology of the Cell 15 (10):4467-4475

McCurley AT, Callard GV (2008) Characterization of housekeeping genes in zebrafish: male-female differences and effects of tissue type, developmental stage and chemical treatment. BMC Molecular Biology 9:12

McMillan DR, Christians E, Forster M, Xiao X, Connell P, Plumier J-C, Zuo X, Richardson J, Morgan S, Benjamin IJ (2002) Heat shock transcription factor 2 is not essential for embryonic development, fertility, or adult cognitive and psychomotor function in mice. Molecular and Cellular Biology 22 (22):8005-8014

Mellersh CS, Pettitt L, Forman OP, Vaudin M, Barnett KC (2006) Identification of mutations in *HSF4* in dogs of three different breeds with hereditary cataracts. Veterinary Ophthalmology 9 (5):369-378

Meng X, Noyes MB, Zhu L, Lawson ND, Wolfe SA (2008) Targeted gene inactivation in zebrafish using engineered zinc finger nucleases. Nature Biotechnology 26 (6):695-701

Menke AL, Spitsbergen JM, Wolterbeek APM, Woutersen RA (2011) Normal anatomy and histology of the adult zebrafish. Toxicologic Pathology 39 (5):759-775

Metchat A, Åkerfelt M, Bierkamp C, Delsinne V, Sistonen L, Alexandre H, Christians ES (2009) Mammalian heat shock factor 1 is essential for oocyte meiosis and directly regulates Hsp90α expression. Journal of Biological Chemistry 284 (14):9521-9528

Min JN, Han MY, Lee SS, Kim KJ, Park YM (2000) Regulation of rat heat shock factor 2 expression during the early organogenic phase of embryogenesis. Biochimica et Biophysica Acta-Gene Structure and Expression 1494 (3):256-262

Min JN, Zhang Y, Moskophidis D, Mivechi NF (2004) Unique contribution of heat shock transcription factor 4 in ocular lens development and fiber cell differentiation. Genesis 40 (4):205-217

Morange M (2006) HSFs in development. In: Gaestel M (ed) Molecular chaperones in health and disease, Handbook of experimental pharmacology, vol. 172. Springer, New York, pp 153-169

Morano KA, Thiele DJ (1999) Heat shock factor function and regulation in response to cellular stress, growth, and differentiation signals. Gene Expression 7:271-282

Morimoto RI (1998) Regulation of the heat shock transcriptional response: cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. Genes & Development 12 (24):3788-3796

Morimoto RI (2011) The stress of misfolded proteins in aging and disease. FEBS Journal 278:23-24

Morrow G, Tanguay RM (2012) Small heat shock protein expression and functions during development. International Journal of Biochemistry & Cell Biology 44 (10):1613-1621

Mosser DD, Caron AW, Bourget L, Meriin AB, Sherman MY, Morimoto RI, Massie B (2000) The chaperone function of hsp70 is required for protection against stress-induced apoptosis. Molecular and Cellular Biology 20 (19):7146-7159

Mou LS, Xu JY, Li WY, Lei X, Wu YL, Xu GX, Kong XY, Xu GT (2010) Identification of vimentin as a novel target of HSF4 in lens development and cataract by proteomic analysis. Investigative Ophthalmology & Visual Science 51 (1):396-404

Muchowski PJ, Valdez MM, Clark JI (1999) αB-crystallin selectively targets intermediate filament proteins during thermal stress. Investigative Ophthalmology & Visual Science 40 (5):951-958

Mullins M (1995) Genetic methods: conventions for naming zebrafish genes. In: Westerfield M (ed) The zebrafish book: a guide for the laboratory use of zebrafish (*Brachydanio rerio*). University of Oregon Press, Eugene, pp 7.1-7.4

Mycko MP, Brosnan CF, Raine CS, Fendler W, Selmaj KW (2012) Transcriptional profiling of microdissected areas of active multiple sclerosis lesions reveals activation of heat shock protein genes. Journal of Neuroscience Research 90 (10):1941-1948

Mymrikov EV, Seit-Nebi AS, Gusev NB (2011) Large potentials of small heat shock proteins. Physiological Reviews 91 (4):1123-1159

Nagai N, Hosokawa M, Itohara S, Adachi E, Matsushita T, Hosokawa N, Nagata K (2000) Embryonic lethality of molecular chaperone Hsp47 knockout mice is associated with defects in collagen biosynthesis. Journal of Cell Biology 150 (6):1499-1506

Nakai A (2010) Novel aspects of heat shock factors. FEBS Journal 277 (20):4111-4111

Nakai A, Ishikawa T (2000) A nuclear localization signal is essential for stress-induced dimer-to-trimer transition of heat shock transcription factor 3. Journal of Biological Chemistry 275 (44):34665-34671

Nakai A, Ishikawa T (2001) Cell cycle transition under stress conditions controlled by vertebrate heat shock factors. EMBO Journal 20 (11):2885-2895

Nakai A, Morimoto RI (1993) Characterization of a novel chicken heat shock transcription factor, heat shock factor 3, suggests a new regulatory pathway. Molecular and Cellular Biology 13 (4):1983-1997

Nakai A, Suzuki M, Tanabe M (2000) Arrest of spermatogenesis in mice expressing an active heat shock transcription factor 1. EMBO Journal 19 (7):1545-1554

Nakai A, Tanabe M, Kawazoe Y, Inazawa J, Morimoto RI, Nagata K (1997) HSF4, a new member of the human heat shock factor family which lacks properties of a transcriptional activator. Molecular and Cellular Biology 17 (1):469-481

Nasevicius A, Ekker SC (2000) Effective targeted gene 'knockdown' in zebrafish. Nature Genetics 26 (2):216-220

Ohara O, Dorit RL, Gilbert W (1989) One-sided polymerase chain reaction: the amplification of cDNA. Proceedings of the National Academy of Sciences of the United States of America 86 (15):5673-5677

Ojima N, Yamashita M (2004) Cloning and characterization of two distinct isoforms of rainbow trout heat shock factor 1 - Evidence for heterotrimer formation. European Journal of Biochemistry 271 (4):703-712

Östling P, Björk JK, Roos-Mattjus P, Mezger V, Sistonen L (2007) Heat shock factor 2 (HSF2) contributes to inducible expression of *hsp* genes through interplay with HSF1. Journal of Biological Chemistry 282 (10):7077-7086

Ovsenek N, Heikkila JJ (1990) DNA sequence-specific binding activity of the heat-shock transcription factor is heat-inducible before the midblastula transition of early *Xenopus* development. Development 110:427-433

Paniagua R, Fraile B, Saez FJ (1990) Effects of photoperiod and temperature on testicular function in amphibians. Histology and Histopathology 5 (3):365-378

Parry DAD (1982) Coiled-coils in alpha-helix-containing proteins - analysis of the residue types within the heptad repeat and the use of these data in the prediction of coiled-coils in other proteins. Bioscience Reports 2 (12):1017-1024

Pfaffl MW (2006) Relative quantification. In: Dorak MT (ed) Real-time PCR. Taylor & Francis, New York, pp 63-82

Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST (c)) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Research 30 (9):10

Pirkkala L, Nykänen P, Sistonen L (2001) Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. FASEB Journal 15:1118-1131

Ponting CP, Schultz J, Milpetz F, Bork P (1999) SMART: identification and annotation of domains from signalling and extracellular protein sequences. Nucleic Acids Research 27 (1):229-232

Posner M, Hawke M, LaCava C, Prince CJ, Bellanco NR, Corbin RW (2008) A proteome map of the zebrafish (*Danio rerio*) lens reveals similarities between zebrafish and mammalian crystallin expression. Molecular Vision 14:806-814

Queitsch C, Sangster TA, Lindquist S (2002) Hsp90 as a capacitor of phenotypic variation. Nature 417 (6889):618-624

Råbergh CMI, Airaksinen S, Soitamo A, Bjorklund HV, Johansson T, Nikinmaa M, Sistonen L (2000) Tissue-specific expression of zebrafish (*Danio rerio*) heat shock factor 1 mRNAs in response to heat stress. Journal of Experimental Biology 203 (12):1817-1824

Rabindran SK, Haroun RI, Clos J, Wisniewski J, Wu C (1993) Regulation of heat shock factor trimer formation: role of a conserved leucine zipper. Science 259 (5092):230-234

Rabindran SK, Wisniewski J, Li L, Li GC, Wu C (1994) Interaction between heat shock factor and hsp70 is insufficient to suppress induction of DNA-binding activity in vivo. Molecular and Cellular Biology 14 (10): 6552-6560

Rallu M, Loones M, Lallemand Y, Morimoto R, Morange M, Mezger V (1997) Function and regulation of heat shock factor 2 during mouse embryogenesis. Proceedings of the National Academy of Sciences of the United States of America 94 (6):2392-2397

Rasmussen R (2001) Quantification on the LightCycler. In: Meuer S, Wittwer C, Nakagawara K (eds) Rapid cycle real-time PCR. Springer-Verlag, Berlin, pp 21-34

Ritossa F (1962) A new puffing pattern induced by temperature shock and DNP in Drosophila. Experientia 18:571-573

Rupik W, Jasik K, Bembenek J, Widlak W (2011) The expression patterns of heat shock genes and proteins and their role during vertebrate's development. Comparative Biochemistry and Physiology A-Molecular & Integrative Physiology 159 (4):349-366

Rutherford SL, Lindquist S (1998) Hsp90 as a capacitor for morphological evolution. Nature 396 (6709):336-342

Saint-Dizier M, Chopineau M, Dupont J, Combarnous Y (2004) Expression of the full-length and alternatively spliced equine luteinizing hormone/chorionic gonadotropin receptor mRNAs in the primary corpus luteum and fetal gonads during pregnancy. Reproduction 128 (2):219-228

Saito Y, Yamagishi N, Hatayama T (2009) Nuclear localization mechanism of Hsp105 and its possible function in mammalian cells. Journal of Biochemistry 145 (2):185-191

Sakurai H, Enoki Y (2010) Novel aspects of heat shock factors: DNA recognition, chromatin modulation and gene expression. FEBS Journal 277 (20):4140-4149

Sakurai H, Takemori Y (2007) Interaction between heat shock transcription factors (HSFs) and divergent binding sequences: Different binding specificities of yeast HSFs and human HSF1. Journal of Biological Chemistry 282 (18):13334-13341

Sambrook J, Russel DW (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor

Sandqvist A, Björk JK, Åkerfelt M, Chitikova Z, Grichine A, Vourch C, Jolly C, Salminen TA, Nymalm Y, Sistonen L (2009) Heterotrimerization of heat-shock factors 1 and 2 provides a transcriptional switch in response to distinct stimuli. Molecular Biology of the Cell 20 (5):1340-1347

Sarge KD, Murphy SP, Morimoto RI (1993) Activation of heat shock gene transcription by heat shock factor 1 involves oligomerization, acquisition of DNA-binding activity, and nuclear localizationand can occur in the absence of stress. Molecular and Cellular Biology 13 (3): 1392-1407

Shchedrina VA, Novoselov SV, Malinouski MY, Gladyshev VN (2007) Identification and characterization of a selenoprotein family containing a diselenide bond in a redox motif. Proceedings of the National Academy of Sciences of the United States of America 104 (35):13919-13924

Sheldon LA, Kingston RE (1993) Hydrophobic coiled-coil domains regulate the subcellular-localization of human heat-shock factor-II. Genes & Development 7 (8):1549-1558

Shi X, Cui B, Wang Z, Weng L, Xu Z, Ma J, Xu G, Kong X, Hu L (2009) Removal of *Hsf4* leads to cataract development in mice through down-regulation of γS-crystallin and *Bfsp* expression. BMC Molecular Biology 10:10

Shi Y, Mosser DD, Morimoto RI (1998) Molecular chaperones as HSF1-specific transcriptional repressors. Genes & Development 12 (5):654-666

Shinka T, Sato Y, Chen G, Naroda T, Kinoshita K, Unemi Y, Tsuji K, Toida K, Iwamoto T, Nakahori Y (2004) Molecular characterization of heat shock-like factor encoded on the human Y chromosome, and implications for male infertility. Biology of Reproduction 71 (1):297-306

Shinkawa T, Tan K, Fujimoto M, Hayashida N, Yamamoto K, Takaki E, Takii R, Prakasam R, Inouye S, Mezger V, Nakai A (2011) Heat shock factor 2 is required for maintaining proteostasis against febrile-range thermal stress and polyglutamine aggregation. Molecular Biology of the Cell 22 (19):3571-3583

Sistonen L, Sarge KD, Morimoto RI (1994) Human heat shock factors 1 and 2 are differentially activated and can synergistically induce hsp70 gene transcription. Molecular and Cellular Biology 14 (3):2087-2099

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, Chinwalla A, Delehaunty A, Delehaunty K, Du H, Fewell G, Fulton L, Fulton R, Graves T, Hou SF, Latrielle P et al. (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423 (6942):825-837

Skandhan KP, Rajahariprasad A (2007) The process of spermatogenesis liberates significant heat and the scrotum has a role in body thermoregulation. Medical Hypotheses 68 (2):303-307

Skjaerven KH, Olsvik PA, Finn RN, Holen E, Hamre K (2011) Ontogenetic expression of maternal and zygotic genes in Atlantic cod embryos under ambient and thermally stressed conditions. Comparative Biochemistry and Physiology A-Molecular & Integrative Physiology 159 (2):196-205

Smaoui N, Beltaief O, BenHamed S, M'Rad R, Maazoul F, Ouertani A, Chaabouni H, Hejtmancik JF (2004) A homozygous splice mutation in the *HSF4* gene is associated with an autosomal recessive congenital cataract. Investigative Ophthalmology & Visual Science 45 (8):2716-2721

Sollars V, Lu X, Xiao L, Wang X, Garfinkel MD, Ruden DM (2003) Evidence for an epigenetic mechanism by which Hsp90 acts as a capacitor for morphological evolution. Nature Genetics 33:70-74

Somasundaram T, Bhat SP (2004) Developmentally dictated expression of heat shock factors: Exclusive expression of HSF4 in the postnatal lens and its specific interaction with αB-crystallin heat shock promoter. Journal of Biological Chemistry 279 (43):44497-44503

Sprague J, Bayraktaroglu L, Clements D, Conlin T, Fashena D, Frazer K, Haendel M, Howe D, Mani P, Ramachandran S, Schaper K, Segerdell E, Song P, Sprunger B, Taylor S, Slyke CV, Westerfield M (2006) The Zebrafish Information Network: the zebrafish model organism database. Nucleic Acids Research 34:D581-D585

Stahl PJ, Mielnik AN, Barbieri CE, Schlegel PN, Paduch DA (2012) Deletion or underexpression of the Y-chromosome genes *CDY2* and *HSFY* is associated with maturation arrest in American men with nonobstructive azoospermia. Asian Journal of Andrology 14 (5):676-682

Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F et al. (2002) Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. Proceedings of the National Academy of Sciences of the United States of America 99 (26):16899-16903

Summerton J (1999) Morpholino antisense oligomers: the case for an RNase-H independent structural type. Biochimica et Biophysica Acta-Gene Structure and Expression 1489 (1):141-158

Tadros W, Lipshitz HD (2009) The maternal-to-zygotic transition: a play in two acts. Development 136 (18):3033-3042

Takaki E, Fujimoto M, Sugahara K, Nakahari T, Yonemura S, Tanaka Y, Hayashida N, Inouye S, Takemoto T, Yamashita H, Nakai A (2006) Maintenance of olfactory neurogenesis requires HSF1, a major heat shock transcription factor in mice. Journal of Biological Chemistry 281 (8):4931-4937

Takemori Y, Enoki Y, Yamamoto N, Fukai Y, Adachi K, Sakurai H (2009) Mutational analysis of human heat-shock transcription factor 1 reveals a regulatory role for oligomerization in DNA-binding specificity. Biochemical Journal 424 (2):253-261

Takemori Y, Enoki Y, Yamamoto N, Fukai Y, Adachi K, Sakurai H (2009) Mutational analysis of human heat-shock transcription factor 1 reveals a regulatory role for oligomerization in DNA-binding specificity. Biochemical Journal 424 (2):253-261

Tanabe M, Kawazoe Y, Takeda S, Morimoto RI, Nagata K, Nakai A (1998) Disruption of the *HSF3* gene results in the severe reduction of heat shock gene expression and loss of thermotolerance. EMBO Journal 17 (6):1750-1758

Tanabe M, Nakai A, Kawazoe Y, Nagata K (1997) Different thresholds in the responses of two heat shock transcription factors, HSF1 and HSF3. Journal of Biological Chemistry 272 (24):15389-15395

Tanabe M, Sasai N, Nagata K, Liu X-D, Liu PCC, Thiele DJ, Nakai A (1999) The mammalian *HSF4* gene generates both an activator and a repressor of heat shock genes by alternative splicing. Journal of Biological Chemistry 274 (39):27845-27856

Tang RY, Dodd A, Lai D, McNabb WC, Love DR (2007) Validation of zebrafish (*Danio rerio*) reference genes for quantitative real-time RT-PCR normalization. Acta Biochimica et Biophysica Sinica 39 (5):384-390

Tanikawa J, Ichikawa-Iwata E, Kanei-Ishii C, Nakai A, Matsuzawa S-I, Reed JC, Ishii S (2000) p53 suppresses the c-myb-induced activation of heat shock transcription factor 3. Journal of Biological Chemistry 275 (20):15578-15585

Tessari A, Salata E, Ferlin A, Bartoloni L, Slongo ML, Foresta C (2004) Characterization of *HSFY*, a novel *AZFb* gene on the Y chromosome with a possible role in human spermatogenesis. Molecular Human Reproduction 10 (4):253-258

Thisse C, Thisse B (2008) High-resolution *in situ* hybridization to whole-mount zebrafish embryos. Nature Protocols 3 (1):59-69

Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets - procedure and some applications. Proceedings of the National Academy of Sciences of the United States of America 76 (9):4350-4354

Traut W, Winking H (2001) Meiotic chromosomes and stages of sex chromosome evolution in fish: zebrafish, platyfish and guppy. Chromosome Research 9 (8):659-672

Trichilis A, Wroblewski J (1997) Expression of p53 and hsp70 in relation to apoptosis during Meckel's cartilage development in the mouse. Anatomy and Embryology 196 (2):107-113

Tu N, Hu Y, Mivechi NF (2006) Heat shock transcription factor (Hsf)-4b recruits Brg1 during the G1 phase of the cell cycle and regulates the expression of heat shock proteins. Journal of Cellular Biochemistry 98 (6):1528-1542

van der Straten A, Rommel C, Dickson B, Hafen E (1997) The heat shock protein 83 (Hsp83) is required for Raf-mediated signalling in *Drosophila*. EMBO Journal 16 (8):1961-1969

Vihervaara A, Sergelius C, Vasara J, Blom MAH, Elsing AN, Roos-Mattjus P, Sistonen L Transcriptional response to stress in the dynamic chromatin environment of cycling and mitotic cells. Proceedings of the National Academy of Sciences of the United States of America 110 (36):E3388-E3397

Voellmy R (2004) On mechanisms that control heat shock transcription factor activity in metazoan cells. Cell Stress & Chaperones 9 (2):122-133

von Koskull-Doring P, Scharf KD, Nover L (2007) The diversity of plant heat stress transcription factors. Trends in Plant Science 12 (10):452-457

Voss AK, Thomas T, Petrou P, Anastassiadis K, Scholer H, Gruss P (2000) Taube nuss is a novel gene essential for the survival of pluripotent cells of early mouse embryos. Development 127 (24):5449-5461

Waites GMH (1991) Thermoregulation of the scrotum and testis: Studies in animals and significance for man. In: Zorgniotti AW (ed) Temperature and environmental effects on the testis. Plenum Press, New York, pp 9-18

Walsh D, Li K, Wass J, Dolnikov A, Zeng F, Li Z, Edwards M (1993) Heat-shock gene-expression and cell-cycle changes during mammalian embryonic development. Developmental Genetics 14 (2):127-136

Walsh D, Li Z, Wu Y, Nagata K (1997) Heat shock and the role of the HSPs during neural plate induction in early mammalian CNS and brain development. Cellular and Molecular Life Sciences 53:198-211

Wang GH, Huang H, Dai R, Lee KY, Lin S, Mivechi NF (2001) Suppression of heat shock transcription factor HSF1 in zebrafish causes heat-induced apoptosis. Genesis 30 (3):195-197

Wang GH, Zhang J, Moskophidis D, Mivechi NF (2003) Targeted disruption of the heat shock transcription factor *hsf*-2 gene results in increased embryonic lethality, neuronal defects, and reduced spermatogenesis. Genesis 36 (1):48-61

Wang GH, Ying ZK, Jin XJ, Tu NX, Zhang Y, Phillips M, Moskophidis D, Mivechi NF (2004) Essential requirement for both hsf1 and hsf2 transcriptional activity in spermatogenesis and male fertility. Genesis 38 (2):66-80

Westerfield M (1995) The zebrafish book: a guide for the laboratory use of zebrafish (*Brachydanio rerio*). University of Oregon Press, Eugene

Wilkerson DC, Skaggs HS, Sarge KD (2007) HSF2 binds to the Hsp90, Hsp27, and c-Fos promoters constitutively and modulates their expression. Cell Stress & Chaperones 12 (3):283-290

Wride MA (2011) Lens fibre cell differentiation and organelle loss: many paths lead to clarity. Philosophical Transactions of the Royal Society B-Biological Sciences 366 (1568):1219-1233

Xiao XZ, Zuo XX, Davis AA, McMillan DR, Curry BB, Richardson JA, Benjamin IJ (1999) HSF1 is required for extra-embryonic development, postnatal growth and protection during inflammatory responses in mice. EMBO Journal 18 (21):5943-5952

Xing H, Wilkerson DC, Mayhew CN, Lubert EJ, Skaggs HS, Goodson ML, Hong Y, Park-Sarge O-K, Sarge KD (2005) Mechanism of *hsp70i* gene bookmarking. Science 307 (5708):421-423

Xu YM, Huang DY, Chiu JF, Lau ATY (2012) Post-translational modification of human heat shock factors and their functions: A recent update by proteomic approach. Journal of Proteome Research 11 (5):2625-2634

Yamamoto N, Takemori Y, Sakurai M, Sugiyama K, Sakurai H (2009) Differential recognition of heat shock elements by members of the heat shock transcription factor family. FEBS Journal 276 (7):1962-1974

Yan LJ, Rajasekaran NS, Sathyanarayanan S, Benjamin IJ (2005) Mouse HSF1 disruption perturbs redox state and increases mitochondrial oxidative stress in kidney. Antioxidants & Redox Signaling 7 (3-4):465-471

Yan LJ, Christians ES, Liu L, Xiao X, Sohal RS, Benjamin IJ (2002) Mouse heat shock transcription factor 1 deficiency alters cardiac redox homeostasis and increases mitochondrial oxidative damage. EMBO Journal 21 (19):5164-5172

Yeh FL, Hsu LY, Lin BA, Chen CF, Li IC, Tsai SH, Hsu T (2006) Cloning of zebrafish (*Danio rerio*) heat shock factor 2 (HSF2) and similar patterns of HSF2 and HSF1 mRNA expression in brain tissues. Biochimie 88 (12):1983-1988

Yeh FL, Hsu T (2000) Detection of a spontaneous high expression of heat shock protein 70 in developing zebrafish (*Danio rerio*). Bioscience, Biotechnology, and Biochemistry 64 (3):592-595

Yeh FL, Hsu T (2002) Differential regulation of spontaneous and heat-induced HSP 70 expression in developing zebrafish (*Danio rerio*). Journal of Experimental Zoology 293 (4):349-359

Yue L, Karr TL, Nathan DF, Swift H, Srinivasan S, Lindquist S (1999) Genetic analysis of viable Hsp90 alleles reveals a critical role in *Drosophila* spermatogenesis. Genetics 151 (3):1065-1079

Zhang Y, Frejtag W, Dai R, Mivechi NF (2001) Heat shock factor-4 (HSF-4a) is a repressor of HSF-1 mediated transcription. Journal of Cellular Biochemistry 82 (4):692-703

Zheng W, Wang Z, Collins JE, Andrews RM, Stemple D, Gong Z (2011) Comparative transcriptome analyses indicate molecular homology of zebrafish swimbladder and mammalian lung. PLoS ONE 6 (8):e24019

Zhou R, Tsang AH, Lau SW, Ge W (2011) Pituitary adenylate cyclase-activating polypeptide (PACAP) and its receptors in the zebrafish ovary: evidence for potentially dual roles of PACAP in controlling final oocyte maturation. Biology of Reproduction 85 (3):615-625

9.0 APPENDIX

Table A1. List of species and NCBI accession numbers used for heat shock factor protein phylogenic and protein domain analysis.

Species	**Common Name (Symbol)**	**HSF1**	**HSF2**	**HSF3**	**HSF4**
Danio rerio	Zebrafish (z)	NP_571675.1	NP_571942.1		NP_001013335.1
Homo sapien	Human (h)	NP_005517.1	NP_004497.1		NP_001035757.1
Mus musculus	Mouse (m)	NP_032322.1	NP_032323.3	BAI50338.1	NP_036069.1
Canus lupis	Dog (d)	XP_532354	XP_533482.2		NP_001041586.1
Equus caballus	Horse (e)	XP_001505050.2	XP_001503083.2		XP_001497009.1
Xenopus laevis	Frog (x)	NP_001090266.1	NP_001089021.1		
Xenopus tropicalis	Frog (xt)			NP_001039053.1	
Gallus gallus	Chicken (c)	P38529.1	NP_001161236.1	XP_420166.2	NP_001165845.1
Taeniopygia guttata	Zebrafinch (f)		XP_002189589.1		XP_002188828
Oncorhynchus mykiss	Trout (t)	NP_001118220.1	NP_001117849.1		
Saccharomyces cerevisiae	Yeast (sc)	NP_011442.1			
Drosophila melanogaster	Fruitfly (dr)	NP_476575.1			

Table A2. Summary of primers used for qPCR. The table shows the name given to the primer, primer sequence, expected amplicon size, the starting nucleotide position on the mRNA sequence where the forward primer binds, and the range of annealing temperatures for which the primers produced a specific amplicon.

Gene (Accession #)	Sequence (5'-3')	Amplicon size (bp)	Primer Start Position in mRNA (bp)	Range of Annealing temperatures (^{o}C)
Elfα (NM_131263.1)	CAAGGAAGTCAGCGCATACA TCTTCCATCCCTTGAACCAG	134	587	56-63
βactin (NM_131031)	CGAGCAGGAGATGGGAACC CAACGGAAACGCTCATTGC	102	722	59-67
hsp70 (NM_131397.2)	ATCAACGAGCCCACGGCTGC ACATGCGGTTCGCAGCCTCC	285	1189	57-67
hsf1 a+b (NM_131600.1)	TGTGGACACGCCCCTTTCGC GCAGGCGACGCTCAAGCACT	84	1225	56-69
hsf1b (NM_131600.1)	GGCTTCTCCACCTCATCTCT CTGTCGATGGACTCCAGATG	77	1225	53-63
hsf2 (NM_131867)	GGATGAGTCCCTGGAGATGA GATAGAAGAGCGTGGCTTCG	95	1436	53-63
hsf4 (both trascripts)	GGTCCAGGTGTTGCGGAGCC AGGGGTGGAGCAGCCGTCAT	237	569	60-69
hsf4b (Ensembl hsf4 201)	GGTCCAGGTGTTGCGGAGCC AGCCAGCGGTGGGGTACAGC	886	569	53-63
hsf4a (NM_001013317)	GCCAAGGAAGTTCTGCCCAA GCTAAAAGTGGTCTCGCCCC	968	306	58-67
hsf4a splice-MO (both transcripts)	GGTGTTGCGGAGCCAACAGG GTGACACCACTTCCCTCCACA	90	575	ND
fgf1 (NM_200760.1)	GCCGTAGGTAAGCAGCTCGCT GTCGCGCGGGCTGGACTTTA	134	5	56-65