IDENTIFICATION AND CHARACTERIZATION OF A TRANSCRIPTION FACTOR

IDENTIFICATION AND CHARACTERIZATION OF A ZINC FINGER
TRANSCRIPTION FACTOR THAT INTERACTS WITH THE CELL CYCLE
REGULATOR HOST CELL FACTOR-1

By

HELEN HOI LAM WONG, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

In Partial Fulfilment of the Requirements

For the Degree

Master of Science

McMaster University

© Copyright by Helen Hoi Lam Wong, March 2002

MASTER OF SCIENCE (2002)   McMaster University
(Biochemistry)   Hamilton, Ontario

Title: Identification and Characterization of a Zinc Finger Transcription Factor That Interacts with the Cell Cycle Regulator Host Cell Factor-1

Author: Helen Hoi Lam Wong, B.Sc.
                   (University of Waterloo)

Supervisor: Dr. John P. Capone

Number of Pages: xii, 109

ABSTRACT

Host cell factor-1 (HCF-1) is an evolutionarily conserved protein first discovered through its interaction with the herpes simplex viral transactivator VP16. Both the amino-terminal VP16 interaction domain and the adjacent, less studied basic domain of HCF-1 are essential in cell cycle regulation. However, the mechanism(s) of this regulation is unknown. Our objective was to provide insight into the importance of the basic domain by studying proteins that interact with this region. Using the yeast two-hybrid system with the HCF-1 basic region as bait, we identified a novel protein named HCF-1 interacting zinc finger protein (HIZ).

The purpose of this research was to characterize HIZ and determine if its interaction with HCF-1 could reveal a novel cellular target for HCF-1. A putative HIZ full-length sequence was determined and its primary structure was studied in detail. HIZ contains 16 C_2H_2 zinc fingers in its amino terminal. HIZ also contains a novel glutamine-rich motif adjacent to a potent autonomous transactivation domain. The presence of a DNA-binding domain structure and a strong, functional transactivation domain indicate that HIZ is a novel transcriptional factor. Northern blot analysis showed tissue specific expression of HIZ with highest levels in the testis, skeletal muscle, liver and pancreas, suggesting possible roles in spermatogenesis, differentiation and metabolism.

The HIZ-HCF-1 interaction was verified *in vitro* using glutathione-s-transferase (GST) pull-down assays and the minimal HIZ domain for HCF-1 interaction was mapped to the same region critical for transactivation. Recent studies have identified the transcription factors Miz-1, GABP, LZIP, Zhangfei and PGC-1β as proteins that interact

with the two domains of HCF-1 important in cell cycle control. Also, these studies have shown that interaction with HCF-1 is important in regulation of their transactivation potential. Thus, the effect of HCF-1 on HIZ activation was studied using transient transfection assays. Similar to its effect on the known cell cycle inhibitor, Miz-1, our studies showed that cotransfection with HCF-1 significantly inhibited HIZ transactivation.

The expression pattern of HIZ in a matched tumour/normal tissue array showed that HIZ is expressed at a significantly lower level in tumours of the lung and uterus in comparison to normal tissues from the same individual. This finding suggests a correlation between HIZ expression and tumour formation, possibly in conjunction with the known cell cycle regulator HCF-1.

DEDICATIONS

This thesis is dedicated to my Mom, Dad and Hubert for their unending love and

support.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

Chapter 3: Results

Chapter 4: Discussion

References

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CeHCF   *Caenorhabditis elegans* HCF homolog

EST   Expressed sequence tag

GAL4DBD   GAL4 DNA binding domain

GAL4AD   GAL4 DNA activation domain

HBM   HCF binding motif

HCF   Host cell factor

HCF_{KEL} HCF kelch repeats

HCF_{SAS} HCF self-association sequence

HCF_{PRO} repeats   HCF processing repeats

HCF_{VID} HCF VP16 interaction domain

HSV IE genes   Herpes simplex virus immediate-early genes

MCS   Multiple cloning site

NLS   Nuclear localization sequence

ORF   Open reading frame

VIC   VP16-induced complex

ZF   Zinc finger

CHAPTER ONE

INTRODUCTION

1.1   Transcription Overview

Eukaryotic gene expression is controlled by multiple mechanisms, one of which is transcription initiation. Transcriptional control is an extraordinarily complex process that is tightly regulated at many levels. The presence of DNA-binding proteins that mediate transcription has been known and extensively studied for over two decades (Crepin *et al.*, 1975). However, the number of co-activators and co-repressors found and the layers of control they create were unexpected.

The presence of three distinct RNA polymerases (Roeder and Rutter, 1969) was the first indication that eukaryotes employ a more complex system of transcriptional control than prokaryotes. Following the discovery of RNA polymerase II (RNAPII), the enzyme responsible for the transcription of mRNA genes, *cis*-acting elements began to be mapped and *trans*-acting factors isolated. *Cis*-acting elements are the specific DNA sequences responsible for directing RNAPII to initiate transcription of protein-encoding genes. These elements include ubiquitous core promoter elements such as the TATA box, responsible for recruiting the RNAPII holoenzyme for basal transcription, and more diverse, gene-specific enhancer, suppressor and silencer elements that interact with DNA-binding transcription factors to modulate the level of transcription in response to cell stimuli (reviewed in Lemon and Tjian, 2000) (see figure 1).   RNAPII
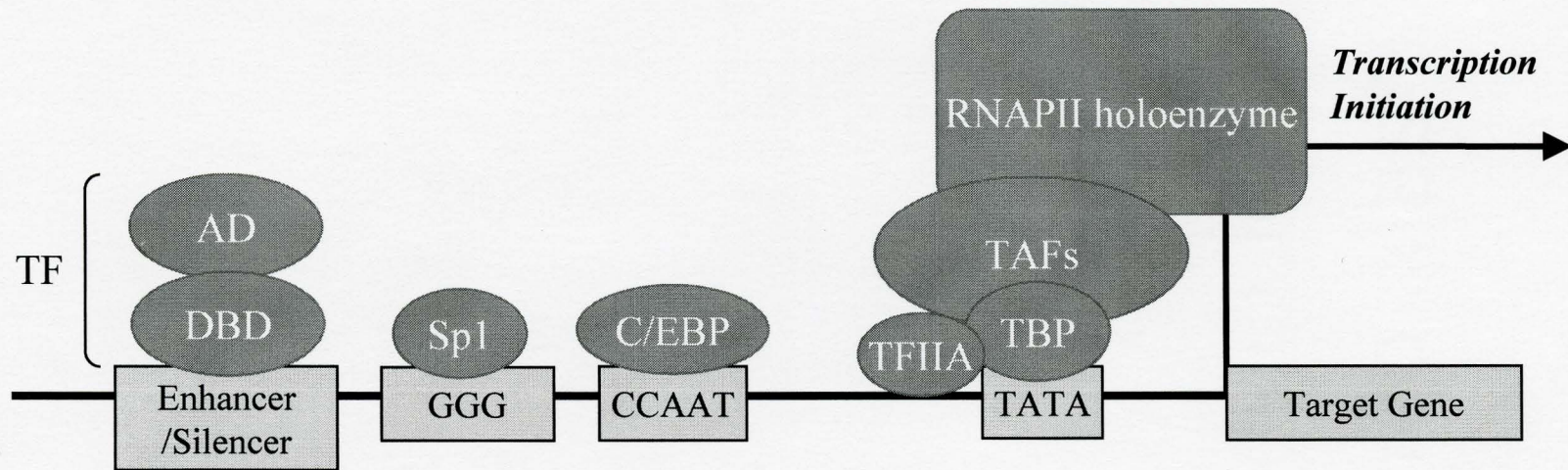
Figure 1. Schematic Representation of Components Involved in Transcription Initiation. The TATA box is found in most promoters and recruits the RNAPII core transcription complex. Various DNA binding transcription factors mediate the level of transcription initiation by binding to more diverse and gene-specific DNA elements and to co-regulators. Abbreviations used: GGG = guanine-rich binding sites of Sp1; TF = transcription factor, AD = activation domain; DBD = DNA binding domain; C/EBP = CCAAT/enhancer binding protein; TBP = TATA binding protein; TAFs = TBP associated factors; RNAPII holoenzyme = RNA polymerase II holoenzyme. The RNAPII holoenzyme includes RNA polymerase II, TFIIB, TFIIE, TFIIF, TFIIH, Srbs and Swi/snf. Note: TAFs and TBP are collectively known as TFIID. Figure adapted from Ogbourne and Antalis, 1998 and Merika *et al.,* 1998.

cannot bind promoter elements on it own, but rely on *trans*-acting factors. A functional RNAPII requires interaction with a number of general transcription factors to form a complex known as the RNAPII holoenzyme (Chen *et al.*, 1994; Hansen and Tjian, 1995). This complex includes transcription factor IIB (TFIIB), TFIIE, TFIIF, TFIIH, some Srbs (suppressors of RNA polymerase B), and the SWI/SNF (switch/sucrose nonfermenting) complex. The RNAPII holoenzyme is recruited to DNA by TFIID, which binds the TATA box and TFIIA (Reviewed by Greenblat, 1997) to form the core/basal transcription machinery.

TFIID was the first eukaryotic general transcription factor to be cloned (Pugh and Tjian, 1990). The recombinant TFIID could initiate basal transcription but unlike a crude TFIID fraction, it failed to support stimulated, higher levels of transcription in response to enhancer-binding regulators (Pugh and Tjian, 1990). This observation led to the discovery of several TATA-binding protein (TBP)-associated factors (TAFs), which are not needed for basal transcription but mediate activator responsiveness (Dynlacht *et al.*, 1991; Tanese *et al.*, 1991). This discovery was the first evidence of a new class of factors called co-activators.

Since then, a large number of co-activators and co-repressors have been identified that are necessary for mediating DNA-binding transcription factors and the activity of the core transcription machinery. In particular, the herpes simplex virus VP16 protein has been a benchmark for elucidating the mechanisms of transcriptional activators because of its remarkably strong activation domain (reviewed in Flint and Shenk, 1997, Hori and Carey, 1994).

In addition to factors that are inherent to the core machinery, DNA-binding transcription factors and those that interact with DNA-binding factors, there exists additional co-factors that are responsible for chromatin remodelling (reviewed in Ogbourne and Antalis, 1998; Lemon and Tjian, 2000). These chromatin-modifying factors include histone acetylases such as GCN5 (reviewed in Wolffe and Pruss, 1996), deacetylases including HDAC-1 (Kadosh and Struhl, 1997; Zhang et al., 1998) and transcriptional coactivators such as cAMP-responsive element-binding protein (CBP)/p300 and CBP-associated factor (p/CAF), which contain intrinsic histone acetylase activity (reviewed in Aranda and Pascual, 2001). Biochemical evidence shows that hyperacetylated histones correlate to transcriptionally active domains (Hebbes et al., 1994), while hypoacetylated histones accumulate within transcriptionally silenced domains (Braunstein et al., 1993).

In summary, biochemical and genetic studies have shown and continue to show that there exists a remarkably complex control mechanism that uses extensive interplay between activators, repressors, chromatin and the basal transcription machinery for transcription initiation. As an example of the fine specificity of this system, the herpes simplex viral (HSV) protein VP16 can discriminate between a single residue difference between the cellular transcription factors Oct-1 and Oct-2 to preferentially bind Oct-1 for initiation of HSV immediate-early genes (HSV IE genes) (Lai et al., 1992). As well, a single mutation at residue 134 in HCF-1 can abolish VP16 binding and stop initiation of the HSV IE genes and subsequent lytic infection (Goto et al., 1997).

1.2 Zinc Finger Transcription Factors

DNA-binding transcription factors can be divided into different families based on common DNA binding structures. One type of DNA-binding domain is the zinc finger motif. Proteins containing this domain include the ubiquitous transcription factor Sp1 as well as many proteins involved in development and differentiation such as WT-1 (critical in the formation of kidney and gonads), Krox 20 (critical in gene expression in developing hindbrain) and the *Drosophila* Kruppel protein (responsible for specifying abdominal cells) (reviewed by Wolfe *et al.*, 2000).

Transcription factors containing this domain, called C_2H_2 zinc finger transcription factors, are very abundant and evolutionarily well conserved. This structure was first identified as a repeating motif in the *Xenopus* 5S rRNA transcription factor TFIIIA (Brown *et al.*, 1985; Miller *et al.*, 1985). They represent approximately 0.7% of genes in *Saccharomyces cerevisiae* and *Caenorhabditis elegans* and have also been found in plants and fungi (Bohm *et al.*, 1997; Clarke and Berg, 1998; Takatsuji, 1998). Recent evidence suggests that zinc fingers can also function as RNA and protein binding domains (Shastry, 1996; Mackay and Crossley, 1998). More studies on the function of this structure may provide insight into why there are a large number of zinc finger proteins encoded in the human genome.

C_2H_2 zinc finger motifs share the consensus sequence: (F/Y)-X-C-X_{2-5}-C-X_3-(F/Y)-X_5-ψ-X_2-H-X_{3-5}-H (where ψ is any large, hydrophobic residue) which serves to coordinate a single zinc ion per motif using the conserved cysteines and histidines (see figure 2). Zinc finger proteins tend to have more than one DNA-binding finger that are joined by linkers. Linkers often consist of five residues between the last conserved
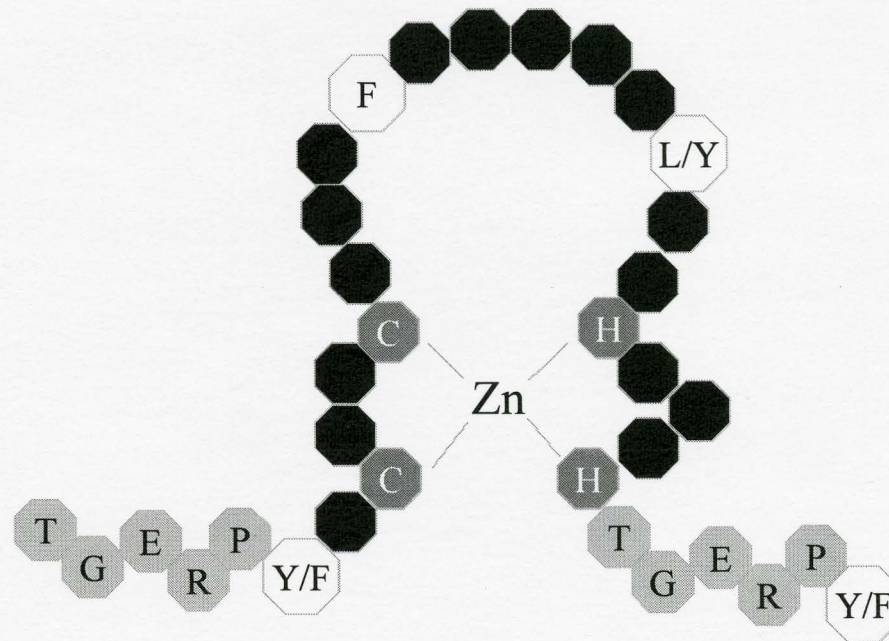
Figure 2. Representation of a C_2H_2-type zinc finger structure and two linker sequences. Two cysteines and two histidines coordinate a zinc ion, forming a finger-like structure capable of binding DNA. The only other conserved residues for this structure are three large, hydrophobic residues (white; residues labelled are those found in HIZ) that are structurally important. The number of residues in the middle region between the second C and first H may vary between 10-12 residues. Linker sequences (light gray) usually have the sequence TGEKP, but in HIZ and its chicken homolog cKr2, this linker sequence is TGERP.

histidine of one finger and the conserved Y/F residue of the following finger, with approximately half of the linkers with the sequence TGEKP (Wolfe *et al.*, 2000). Mutagenesis studies have shown that linkers are critical for high affinity DNA binding (Choo and Klug, 1993; Ryan and Darby, 1998) and in proteins with only one or two zinc finger structures encoded, additional elements are usually required for specific DNA recognition (Fairall *et al.*, 1992; Fairall *et al.*, 1993; Bowers *et al.*, 1999).

Zinc fingers fold in the presence of zinc (Foster *et al.*, 1997) and the crystal structure of several zinc finger domains have been shown to form two-stranded antiparallel β-sheets and an α-helix (Lee *et al.*, 1989; Parraga *et al.*, 1988; Pavletich and Pabo, 1991). The two cysteines at one end of the β-sheet and two histidines in the carboxy-terminal of the α-helix coordinate the zinc ion (Elrod-Erickson *et al.*, 1996; Pavletich and Pabo, 1991). As well, three conserved hydrophobic residues surround the zinc binding site to provide structural stability (Michael *et al.*, 1992).

1.3 Host Cell Factor-1

1.2.1 HCF-1 and HSV Infection

Host cell factor-1 (HCF-1) is an endogenous cellular protein first identified by its critical role in HSV infection. Gene expression during HSV lytic infection is a very well studied cascade initiated by the viral protein VP16 (also known as Vmw65 or αTIF) (reviewed in O'Hare, 1993).

VP16 is a 490 residue viral protein with two autonomous functional domains (see figure 3). The amino-terminus (aa 1-388) contains a short tetrapeptide motif called the HCF binding motif (HBM): (D/E)HXY (where X is any amino acid) that associates with

HCF-1 independent of the carboxy-terminus. VP16 first forms a stable heteromeric dimer with HCF-1 and then interacts with the cellular DNA-binding factor Oct-1 to form the trimeric VP16-induced complex (VIC) (see figure 4) (Gerster and Roeder, 1988; Katan *et al.*, 1990; Werstuck and Capone, 1993; reviewed in Herr, 1998). The VIC binds TAATGARAT *cis*-elements (where R is any purine), which occur in ALL immediate-early (IE) promoters, and activates HSV-1 IE gene transcription using the very potent activation domain of VP16 (Sadowski *et al*, 1988; Triezenberg *et al*, 1988; Herr, 1998; O'Hare, 1993). The VP16 activation domain is within the carboxy-terminus and is enriched with acidic residues (residues 410-490) (Sadowski *et al.*, 1988; Cousens *et al.*, 1989).

Extensive site-directed mutagenesis within the amino-terminal of VP16 (responsible for VIC assembly) showed that distinct residues on VP16 are critical for binding HCF-1 and DNA-bound Oct-1 (Lai and Herr, 1997). This study also showed that VP16 is capable of interacting with DNA and the mutations that hindered this binding also hindered Oct-1 binding to DNA (Lai and Herr, 1997), implying a role for VP16 in stabilizing Oct-1 binding to TAATGARAT elements.

Oct-1 is a widely expressed, 766-residue transcription factor that recognizes DNA elements with the octamer motif ATGCAAAT through its bipartite POU domain (Herr and Cleary, 1995). The POU domain of Oct-1 consists of an amino-terminal POU-specific (POU$_S$) domain and a carboxy-terminal POU-homeo (POU$_H$) domain that are joined by a flexible linker (reviewed in Herr and Cleary, 1995). This domain is flanked on each side by an amino-terminal and a carboxy-terminal activation domain (Tanaka and
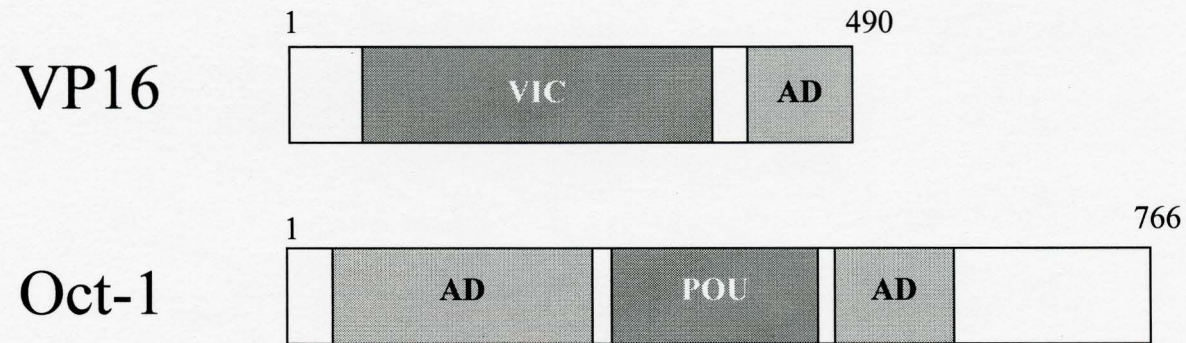
Figure 3. VP16 and Oct-1 domain structures. The amino-terminal region of VP16 and the POU domain of Oct-1 are critical for VIC assembly. VP16 contains a carboxy-terminal acidic activation domain (AD) and Oct-1 contains two ADs: an amino-terminal AD enriched with glutamine residues and a carboxy-terminal AD enriched in serine and threonine residues. HSV gene transcription requires only the VP16 activation domain.
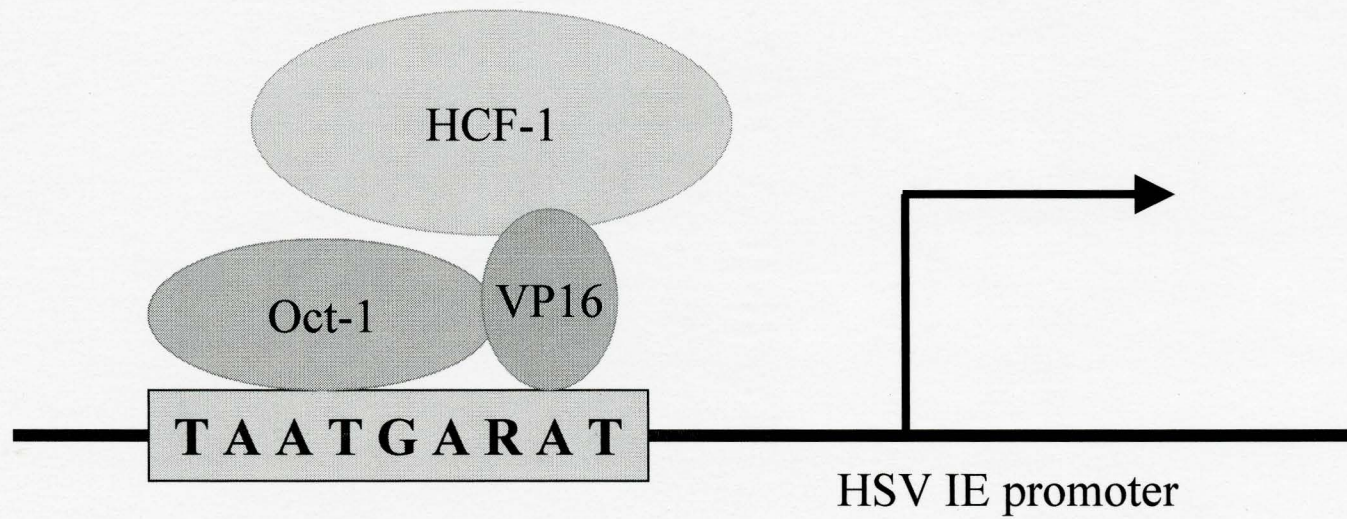
Figure 4. Schematic Representation of the VP16 induced complex (VIC). Formation of the complex initiates transcription of HSV IE genes.

Herr, 1990; Tanaka *et al.*, 1992) but only the POU domain is important for VIC assembly (Kristie *et al.*, 1989; Stern *et al.*, 1989) (see figure 3).

Oct-1 has a cellular role of activating transcription from small nuclear (sn)RNA-type promoters that possess a proximal sequence element (PSE) instead of a TATA box, binding the snRNA activating protein complex (SNAPc) and not TFIID. The activation domains of Oct-1 associate with SNAP$_c$ to activate snRNA genes but not mRNA genes (Tanaka *et al*, 1988). *In vitro,* Oct-1 can bind TAATGARAT elements at low affinity but cannot activate transcription from either snRNA-type or mRNA-type promoters with this sequence, likely because the binding is not very efficient (Cleary *et al.*, 1993). However, when VP16 binds Oct-1, it can stabilize Oct-1 and the strong activation domain of VP16 can activate transcription of HSV IE mRNA genes (Cleary *et al.*, 1993). Oct-2 is a transcription factor containing a POU domain with the same DNA-binding specificity as Oct-1 but activates mRNA-type promoters more efficiently (Tanaka *et al.*, 1988). VP16 recognizes the POU$_H$ domain on Oct-1 but because of a single residue difference (glutamic acid in Oct-1, alanine in Oct-2), VP16 is not able to associate with Oct-2 and thus cannot alter the DNA-binding specificity of Oct-2 to bind TAATGARAT elements (Lai *et al.*, 1992).

In vitro, VP16 and Oct-1 form a stable dimer, but *in vivo*, the interaction between the amino-termini of HCF-1 and VP16 is crucial for HSV IE gene transcription (reviewed in Herr, 1998). The resulting VIC can activate both snRNA-type promoters using the Oct-1 activation domain and mRNA-type promoters using the VP16 activation domain (Cleary *et al.*, 1993).

The mechanism by which HCF-1 stabilizes VP16 interaction with the POU_H domain of Oct-1 remains unknown, but HCF-1 is known to be a determining factor of whether the HSV IE genes are activated or remain latent in the cell.

1.3.2 HCF-1 Domain Structure

The HCF-1 mRNA is very large, encoded by 26 exons spread across a region spanning approximately 24kb in the chromosomal region Xq28 (Frattini *et al*, 1994). HCF-1 is translated as a large 2035-residue protein of approximately 300kD and consists of multiple distinct domains for various functions (see figure 5). Amino acids 20-380 at the amino-terminus are necessary for binding VP16 and this region is thus commonly referred to as the VP16-interaction domain (HCF_{VID}). Within this domain are six repeats of approximately 50-65 amino acids referred to as the HCF_{KEL} repeats, which resemble repeats found in the *Drosophila* Kelch protein (Xue and Cooley, 1993). Each repeat forms a β-sheet to form a six-bladed β-propeller structure (Bork and Doolittle, 1994) sufficient for VIC assembly both *in vivo* and *in vitro* (La Boissiere *et al*, 1997; Wilson *et al.*, 1997).

Adjacent to HCF_{VID} is a region with a high concentration of basic amino acids, followed by a domain located near its center that contains six 26-amino acid repeats, called the HCF_{PRO} repeats (Wilson *et al.*, 1993; Kristie *et al.*, 1995). Post-translation, HCF-1 is proteolytically processed at one of the six repeats to form a collection of amino-
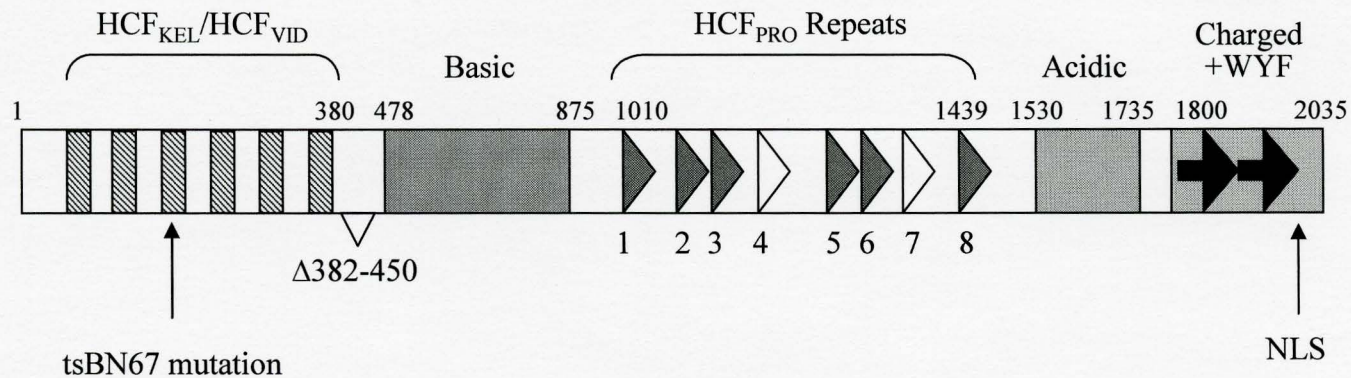
Figure 5. HCF-1 domain structure. Structural features of HCF-1 are as follows from the amino to carboxy terminal: hatched rectangles representing the kelch repeats of the VP16 interaction domain; the basic domain; triangles representing the HCF processing (HCF$_{PRO}$) repeats (open triangles represent the two less conserved, non-functioning repeats); the acidic domain; and at the carboxy-terminus are two fibronectin type 3 repeats (black arrows) and a nuclear localization sequence (NLS) within residues 2015-2031. A splice variant of HCF-1 (HCF-1$_{\Delta382-450}$) exists that lack residues 382-450 and lacks the ability to self-associate after processing. A cell cycle deficient HCF-1 in the tsBN67 cell line is caused by a proline-to-serine at residue 134. Figure adapted from Wysocka *et al.,* 2000.

and carboxy-terminal fragments of varying lengths, ranging in size from 110-150kD that usually remain strongly, but non-covalently associated, even in the presence of 3M urea (Wilson et al., 1993, Kristie and Sharp, 1993). The function of this unusual processing may be a mechanism that regulates the various functions of HCF-1 (Vogel and Kristie, 2000b), but no direct evidence of this has been shown. These repeats, so far unique to HCF-1, are found in the murine homolog (Kristie, 1997) but do not occur in the *Caenorhabditis elegans* homolog (CeHCF) (Liu et al., 1998), suggesting that mammalian HCF-1 evolved in more complex organisms to accommodate additional cellular roles.

Separately synthesized amino- (HCF-1_N) and carboxy-subunits (HCF-1_C) readily associate *in vitro* and studies by Wilson et al., showed that this association was attributed to two matched pairs of HCF-1 amino- and carboxy-self association elements: HCF-1_{SAS1} which consists of HCF-1_{SAS1N} and HCF-1_{SAS1C} partners, and HCF-1_{SAS2} which consists of HCF-1_{SAS2N} and HCF-1_{SAS2C} partners (Wilson et al., 2000). HCF-1_{SAS1N} has been mapped to residues 360-402 and associates with HCF-1_{SAS1C} at residues 1812-2002, a region that contains a pair of fibronectin type 3 (Fn3) repeats (Wilson et al., 2000). The second set of self-association domains lie within residues 491-755 of the basic region (HCF-1_{SAS2N}) and residues 1436-1756 of the acidic region (HCF-1_{SAS2N}) (Wilson et al., 2000). This study also showed that association of HCF-1_N subunits with HCF-1_C subunits are required for nuclear localization (Wilson et al., 2000).

Carboxy-terminal to the HCF$_{PRO}$ repeats is a region enriched with acidic amino acids and both the amino- and carboxy-termini are enriched with charged and large, hydrophobic residues (i.e. W, Y, F). Within the carboxy-terminus, at residues 2015-2031 is a nuclear import signal that may be required for nuclear import of VP16 (La Boissiere

et al., 1999). Deletion of these residues was shown to abrogate nuclear accumulation of both HCF-1 and VP16 (La Boissiere *et al.*, 1999) and nuclear localization of HCF-1 has been shown to correlate directly with reactivation of HSV lytic replication (Kristie *et al.*, 1999).

1.3.3 HCF-1 in Cell Cycle Control

HCF-1 is highly expressed in fetal tissues suggesting a role in cell proliferation and development (Wilson *et al.*, 1995). Although the role of HCF in HSV transcription had been well characterized, its cellular role remained very poorly understood until it was found that the proliferation defect of the temperature-sensitive BHK-21-derived tsBN67 cell line was due solely to a mutation in HCF-1 (Goto *et al.*, 1997). The change of a single proline-to-serine substitution at HCF residue 134 in the third kelch repeat was responsible for the cell cycle arrest at the G_0/G_1 stage and also disrupted VP16 interaction (Goto *et al.*, 1997). However, the mechanisms involved in this cell cycle regulation are still unclear.

Since the HCF mutation is located in the VP16 interaction domain, VP16 likely exploits HCF-1 to gauge the cell cycle status of the infected cell by binding to the domain of HCF-1 required for cell cycle proliferation. However, the VP16 interaction domain (residues 1-380) is not sufficient to rescue the tsBN67 cell proliferation defect. The basic domain (residues 478-875) is also needed to overcome the cell proliferation defect (Wilson *et al.*, 1997) suggesting that more than one domain of HCF-1 is required for cell cycle progression, only one of which HSV targets (see figure 6).

HCF-1 is a very well conserved protein, with homologs in both insect and nematode extracts capable of forming the VIC (Kristie *et al.* 1989). Interestingly, a mutation in the full-length *C. elegans* HCF-1 (P145S) corresponds to the human HCF-1 P134S cell cycle deficient mutation and disrupts VIC formation (Liu *et al*, 1999).

Sequencing of CeHCF showed a high degree of similarity to human HCF-1 at the amino- and carboxy-terminus (Liu *et al*, 1998). In addition to the HCF$_{VID}$, the *C. elegans* homolog has the self-association domains SAS1N and SAS1C, as well as the carboxy terminal NLS. CeHCF lacks the basic domain, which is necessary in HCF-1 for cell cycle rescue of tsBN67 cells. Therefore, it is not surprising that CeHCF is not sufficient for rescuing the cell cycle defect although capable of forming the VIC (Lee and Herr, 2001).

1.3.4 HCF-1 Binding Partners

Several novel interactions between HCF and various proteins have been identified with the yeast two-hybrid system, and represent possible cellular targets of HCF-1 function (see figure 7).

The basic leucine zipper proteins LZIP (also known as Luman) and Zhangfei were discovered by virtue of their interactions with the HCF-1 VP16 interaction domain (Freiman and Herr, 1997; Lu *et al.*, 1997; Lu and Misra, 2000). LZIP and Zhangfei use the same tetrapeptide HBM (D/EHXY) as VP16 for HCF-1 interaction and both contain an acidic activation domain (Luciano and Wilson, 2000; Lu and Misra, 1997). Similar to VP16, the binding of LZIP and Zhangfei to the cell cycle deficient HCF mutant (P134S) is impaired.
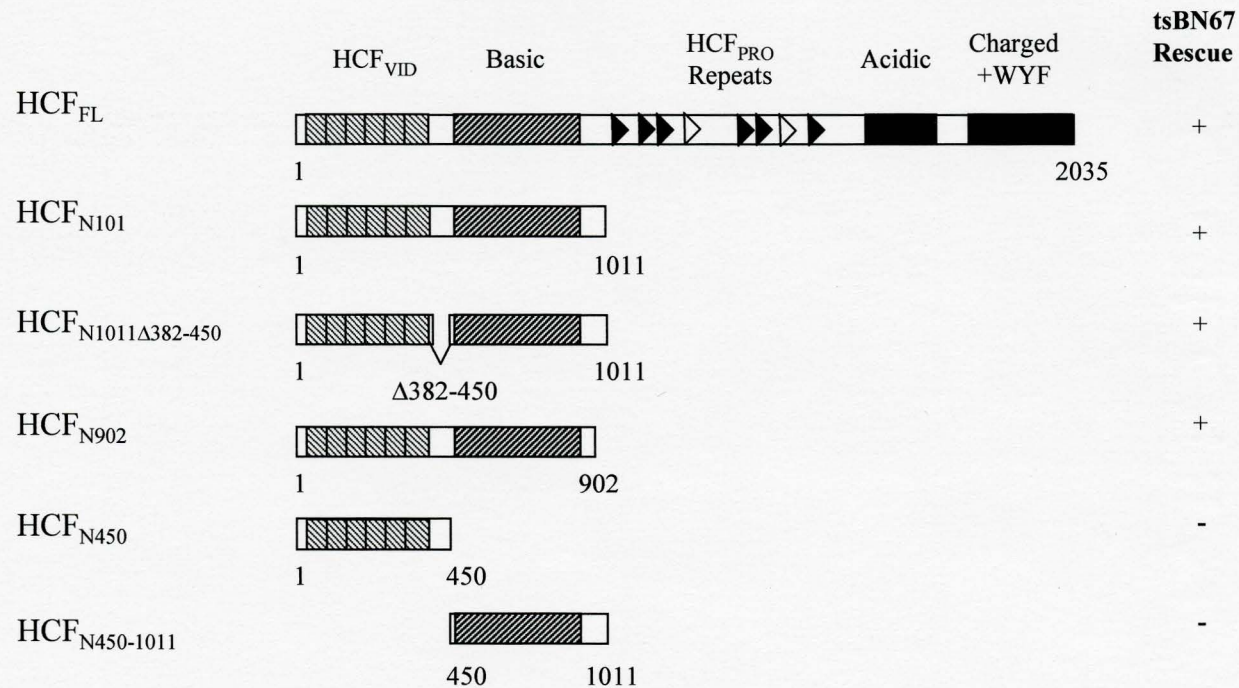
Figure 6. The basic domain of HCF-1 is involved in cell cycle regulation. HCF-1 fragments were constructed by Wilson *et al.* to investigate the regions necessary to rescue the temperature sensitive cell cycle defect in tsBN67 cells. Both the HCF$_{VID}$ and basic regions were shown to be required for the rescue. Residues 382-450 represent a region that can be removed by alternative splicing and were shown to be unnecessary for the rescue. Adapted from Wilson *et al.*, 1997.
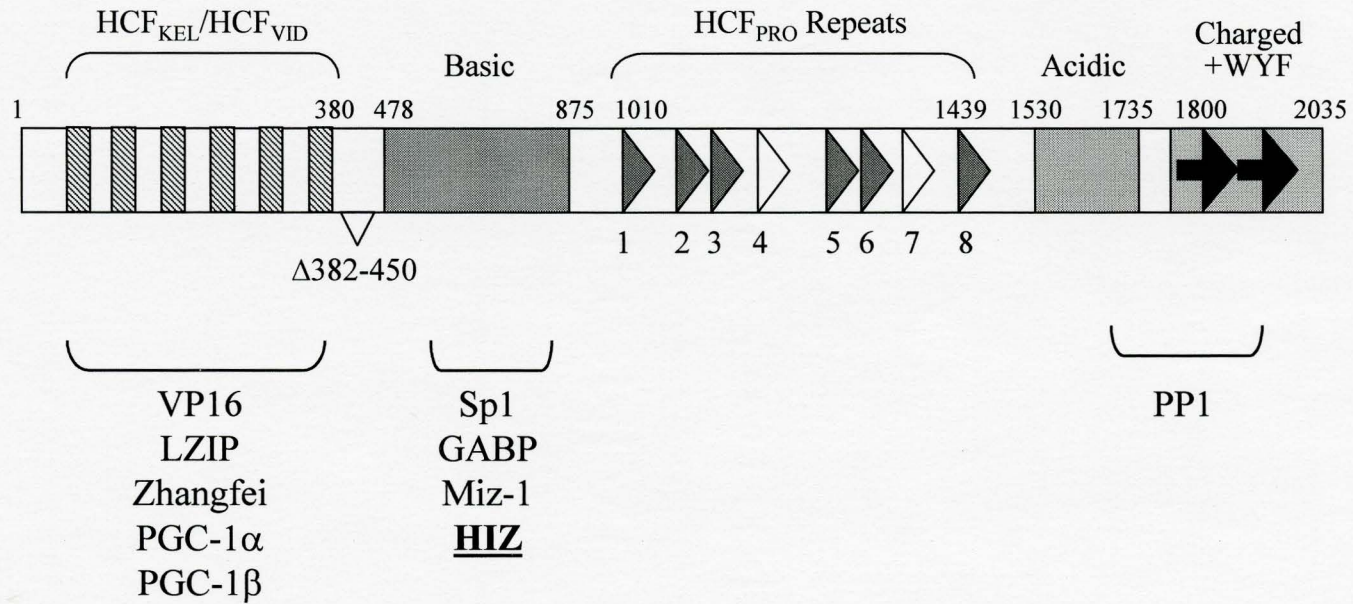
Figure 7. Summary of HCF-1 Binding Proteins. LZIP, Zhangfei, PGC-1 α and PGC-1β both bind using the same HCF-1 binding motif (D/EHXY) used by VP16 in the amino terminal HCF_{KEL}/HCF_{VID} domain. Sp1, GABP, Miz-1 and HIZ all bind within the basic region and the protein phosphatase PP1 binds within the carboxy-terminus.

LZIP is a member of the cyclic AMP response element (CRE)-binding protein/activating transcription factor (ATF) family 1 protein of the basic leucine zipper superfamily. It contains a DNA binding domain related to members of the CREB/ATF family and is capable of activating promoters with CRE elements (Lu *et al.*, 1997). Its activation potential is inhibited by VP16 suggesting competition between the two proteins, possibly for HCF. It has a mouse homolog of the same name (Burbelo *et al*, 1994) that contains two conserved LxxLL motifs (Luciano and Wilson, 2000). LxxLL motifs are found in many transcriptional coactivators and mediate binding to nuclear hormone receptors (Heery *et al*, 1997). Both the HCF binding motif and the two LxxLL motifs are required for LZIP mediated transactivation (Luciano and Wilson, 2000). Mutation of the LxxLL motifs did not disrupt HCF-1 binding (Luciano and Wilson, 2000), suggesting that LZIP activation requires proteins additional to HCF-1.

Similar to LZIP, the activation potential of Zhangfei is impaired upon mutation of the HBM, implying that HCF-1 is involved in regulating its transactivation ability. However, contrary to LZIP, Zhangfei does not bind to consensus bZIP sites *in vitro*.

Recently, the transcriptional coactivators PGC-1α (peroxisome proliferator activated receptor γ (PPARγ) coactivator-1α) and PGC-1β were found to contain the HBM and both were shown to bind HCF-1. Furthermore, the addition of HCF-1 in transient transfection assays with GAL4-PGC-1β demonstrated that HCF-1 could increase the activation potential of PGC-1β from 10 to 30-fold (Lin *et al.*, 2002). Both PGC-1α and PGC-1β interact with nuclear hormone receptors and co-activators such as PPARα and PPARγ, hepatic nuclear hormone receptor 4α (HNR4α), nuclear respiratory factor-1 and -2 (NRF-1 and NRF-2) and glucocorticoid receptor (GR) (Puigserver *et al.*,

1998; Wu *et al.*, 1999; Lin *et al.*, 2002). Evidence suggests that PGC-1α and PGC-1β may play a role in regulatory pathways in the liver that are stimulated during fasting (Yoon *et al.*, 2001).

Three known transcription factors: GABP, Sp1 and Miz-1 associate with the basic domain of HCF-1 (Vogel and Kristie, 2000a; Gunther *et al.*, 2000; David Piluso, personal communication). There are *cis*-regulatory elements on the promoter of HSV IE genes that bind GABP and Sp1, suggesting their involvement in viral gene expression (Jones and Tjian, 1985; LaMarco and McKnight, 1989; Triezenberg *et al.*, 1988). Like LZIP and Zhangfei, GABP, Sp1 and Miz-1 all contain both a DNA-binding domain and an activation domain.

GABP is a heterodimeric transcription factor consisting of an α-subunit that contains the characteristic ETS domain which interacts with the DNA sequence (C/A)GGS(A/T)(G/C), and a β-subunit for transcriptional activation (Shaeffer *et al*, 1998). GABP and Sp1 have been shown to activate transcription in a cooperative manner (Rosemarin, *et al.*, 1998; Vassias *et al.*; Nuchprayoon *et al.*, 1999) and thus could be regulated by binding to a common protein, such as HCF-1. The activation domain of GABP has been mapped to residues in the domain responsible for binding HCF-1 (Vogel and Kristie, 2000a). HCF-1 binding efficiency correlated directly with the critical residues for transactivation (Vogel and Kristie, 2000a), alluding to a role for HCF-1 in regulation of its transactivation potential.

Miz-1 is an 803 amino acid transcription factor initially discovered through its interaction with the proto-oncogene c-*myc* and was therefore named myc-interacting zinc finger protein 1. Miz-1 is responsible for cell cycle arrest at G_1 by activating

transcription of the cell cycle inhibitor p15^{INK4b} (Staller *et al.*, 2001). HCF$_{750-902}$ is sufficient to bind and inhibit Miz-1 transactivation of the p15^{INK4b} gene (David Piluso, personal communication), implicating a novel role for HCF-1 in co-repression of transcription. This evidence also provides a possible mechanism by which HCF-1 can be involved in cell cycle control. HCF-1 may have a role similar to c-myc in repressing p15^{INK4b} transcription, thus stimulating cell proliferation.

The protein phosphatase PP1 binds the carboxy-terminal of HCF-1 within residues 1604-1956 (Ajuh *et al.*, 2000), suggesting phosphorylation as a possible mechanism for regulation.

1.3.5 HCF-1 Related Proteins

Using a database search with the amino terminal, a new HCF protein, called HCF-2 was identified. HCF-2 is remarkably similar to the *C. elegans* HCF in size and organization and can also support VIC assembly (Johnson *et al.*, 1999). However, both VP16 and LZIP preferentially interact with HCF-1. HCF-2 binds VP16 very weakly and fails to bind LZIP at all, likely because of differences in the fifth and sixth kelch repeats (HCF$_{KEL}$5 and HCF$_{KEL}$6) in the HCF$_{VIC}$ domain (Johnson *et al.*, 1999). HCF-2 is expressed in the spleen, thymus, prostate, ovary, small intestine, colon and peripheral blood leukocytes at low levels but is highly expressed in testis tissue, in contrast with HCF-1, which is expressed at low levels in all of the tissues listed (Johnson *et al.*, 1999). Like CeHCF, HCF-2 lacks the basic domain and cannot rescue the cell proliferation defect. Co-expression with HCF-1 showed that HCF-2 can act as a potent inhibitor of HCF-1 cell cycle rescue of the defect (Johnson *et al.*, 1999) and thus, HCF-2 could

possibly serve as a regulator of HCF-1 cell cycle function. Also similar to CeHCF, HCF-2 is not proteolytically cleaved but contains the HCF-1$_{SASIN}$ and HCF-1$_{SASIC}$ self-association elements (Wilson *et al.*, 2000). This suggests that the self-association elements may function not only for self-association, but have additional roles, possibly to interact with other proteins.

Recently, through a database search for proteins containing the ß-propeller domain of HCF-1, a novel protein named HCF-like kelch repeat protein (HCLP-1) was discovered. It is a 406-amino acid protein expressed in a variety of human tissues and cancer cell lines (Zhou *et al.*, 2001). HCLP is not able to bind VP16 but binds LZIP. Unlike HCF-1, which is necessary for LZIP transactivation, HCLP represses LZIP activation of reporter gene (Zhou *et al.*, 2001).

1.3.6 HCF Summary

HCF-1 is a large, modular protein that is evolutionarily well conserved with homologs found in *C. elegans* and mice. The amino terminal HCF$_{VIC}$ and the adjacent basic regions are both involved in cell proliferation but the mechanisms by which these two domains regulate the cell cycle is unknown. To elucidate the mechanisms, the cellular targets of HCF-1 must be determined. Recent studies of HCF-1 interacting proteins have identified a role for HCF-1 in regulation of transactivation potential. HCF-1 likely plays a role in promoting cell proliferation by repressing those factors that inhibit while enhancing those that promote the cell cycle.

1.4 Objectives of the Research

The amino-terminal VP16 interaction domain of HCF-1 has been studied in great detail but little is known about the function of the adjacent basic domain except that it has a role in rescuing the cell proliferation defect in tsBN67 cells. To investigate the role of the basic domain of HCF-1, a yeast two-hybrid assay was performed using this domain as bait to screen a HeLa cDNA library for novel interactions with HCF-1. Through this system, a novel protein was identified to bind HCF-1. This novel protein was not found in the Genbank database and was named HIZ (HCF-1-interacting zinc finger protein).

The studies herein were designed to:

a) Characterize the expression of HIZ in various human tissues,

b) Characterize a putative full length nucleotide and amino acid sequence for HIZ,

c) Define the region in HIZ critical for transactivation,

d) Define the region in HIZ critical for binding HCF-1, and

e) Determine the effect of HCF-1 on HIZ activation.

The data presented will support the conclusions that HIZ is an evolutionarily conserved zinc finger protein that is expressed in a variety of human tissue. In addition, HIZ is shown to contain a transactivation domain within a region that corresponds to the residues essential for HCF-binding. These data support previous evidence that HCF has an important function in regulating transcription initiation by binding to a variety of transcription factors.

In addition, to investigate the co-activator function of HCF-1, various regions of HCF-1 were tethered to the GAL4 DNA binding domain (GAL4DBD). Through these

studies, we discovered the presence of an autonomous transactivation domain within the acidic domain of HCF-1 at residues 1573-1726.

CHAPTER TWO

MATERIALS AND METHODS

2.1 Materials

2.1.1 Chemicals and Reagents

The following is a list of chemicals and reagents used in this research and the corresponding suppliers:

Ampicillin	Sigma Chemical Company
Anti-gal4DNA-BD mouse monoclonal antibody	Santa Cruz Biotechnology
BioRad protein assay	BioRad Laboratories
Bovine serum albumin	Pharmacia Biotech
Chloramphenicol	Sigma Chemical Company
Complete Mini (EDTA free) protease inhibitor cocktail tablets	Roche Diagnostics
Deoxynucleotide triphosphates (dNTPs)	Pharmacia Biotech
Dimethylsulphoxide (DMSO)	Caledon Laboratories
ECL reagent	Amersham Life Sciences
ExpressHyb solution	Clontech
Fetal Bovine Serum	Sigma Chemical Company
Glutathione Sepharose 4B beads	Pharmacia Biotech

Hybond™-C pure nitrocellulose membrane	Amersham Life Sciences
IPTG	Biochemica and Synthetica
L-glutamine	Life Technologies
Luciferin	Biosynth
Molecular weight standards:	
1kb Plus DNA ladder	Life Technologies
Benchmark™ prestained protein ladder	Life Technologies
Benchmark™ protein ladder	Life Technologies
Penicillin	Life Technologies
Phenylmethylsulphonylfluoride (PMSF)	Boehringer Mannheim
5X reporter lysis buffer	Promega
Salmon sperm DNA	Sigma Chemical Company
Sephadex G-50	Pharmacia Biotech
Sheep anti-mouse polyclonal antibody, horseradish peroxidase	Amersham Life Sciences
Triton X-100	BDH Chemicals
X-gal (5-bromo-4-chloro-3-indoyl-β-D-galactoside)	Life Technologies

2.1.2 Enzymes

The following lists the enzymes used in the this research and the corresponding suppliers.

Advantaq	Clontech
All restriction endonucleases	New England Biolabs

Calf intestinal alkaline phosphatase (CIP)	New England Biolabs
DeepVent DNA polymerase	New England Biolabs
HotStarTaq polymerase	Qiagen
Klenow DNA polymerase	New England Biolabs
Lysozyme	Sigma Chemical Company
Pfu polymerase	Strategene
RNAguard ribonuclease inhibitor	Pharmacia Biotech
RNase A	Pharmacia Biotech
SP6 RNA polymerase	Promega
T4 DNA ligase	New England Biolabs
Vent DNA polymerase	New England Biolabs

2.1.3 Radiochemicals

L-[^{14}C]-Leucine	Amersham Biosciences
[α-^{32}P]-dATP	NEN Life Science Products
[^{14}C]-labeled molecular weight markers	Life Technologies

2.1.4 Bacterial Strains and Growth Conditions

The *Escherichia coli* DH5α bacterial strain was used as a host for all plasmids constructed with the exception of PGEX-4T-1-HCF$_{750-902}$ which required the Bl21 bacterial strain. All bacteria were grown at 37°C in 2YT media (1.6% bactotryptone, 1% bacto-yeast and 0.5% NaCl) supplemented with 100μg/ml ampicillin to select for bacteria containing plasmid.

2.1.5 Yeast Strains and Growth Conditions

The Y190 yeast strain was used in this research. Yeast were grown at 30°C in YPD media (1% bacto-yeast extract, 2% bacto-peptone and 2% glucose).

2.1.6 Mammalian Cell Lines and Culture

The following cell lines were used for transient transfection assays:

COS-1 cell line-- African green monkey kidney fibroblast cells transformed with the

SV40 T antigen; obtained from the American Type Culture Collection (ATCC).

BHK21 cell line-- Baby Hamster Kidney fibroblast cells; obtained from ATCC

HeLa cell line—Human epithelioid cervical carcinoma cells; obtained from ATCC

The COS-1, BHK-21 and HeLa cell lines were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% v/v fetal bovine serum, 1% v/v L-glutamine and 1% v/v penicillin-streptomycin and grown at 37°C with 5%CO_2.

2.1.7 Oligonucleotides

All oligonucleotides were synthesized by the Central Facility of the Institute for Molecular Biology (MOBIX), McMaster University.

Oligonucleotides listed in table 1 were synthesized to attempt to PCR amplify full length HIZ. Oligonucleotides listed in table 2 and 3 were used during this project for PCR amplification of HCF-1 and HIZ fragments respectively for deletion analysis.

Table 1. Oligonucleotides used during project for PCR amplification of full length HIZ.

Synthesis#	Sequence	Purpose
AB25769	GCGAATTCGCCATGGCCCCTCCCCCAGGTGT GATGGTCGGC	Forward primer for PCR amplification based on putative HIZ sequence (-15 and including start codon)
AB25770	GGCCATGGCACACGGGCGAGCGGCCCTAC	Forward primer for PCR amplification based on putative HIZ sequence (5' end begins at nt 241 of putative sequence, which is also nt 12 of chicken cKr2 gene)
AB25771	GCCCATGGGCCGACCAGCCACCGTCCCCT	Forward primer for PCR amplification based on putative HIZ sequence (5' end begins at Pro/Ala-rich region I: nt 786)
AB25772	GCCCATGGCCCCAGGAACCGCTGGCGCCT	Forward primer for PCR amplification based on putative HIZ sequence (5' end begins at Pro/Ala-rich region II: nt 1274)
AB25773	GCGGATCCGGCGGGGGCGGCGCCGTGCGT	Reverse primer for PCR amplification based on putative HIZ sequence (5' end begins at Pro-rich region I: nt 684)
AB25774	GCGGATGGCCCACTCATTTCCTGTGGCCC	Reverse primer for PCR amplification based on HBZ sequence of 2-hybrid screen (5' end begins at start of EST AI912315)

Table 2. Oligonucleotides used during project for PCR amplification of HCF fragments.

Synthesis#	Sequence	Purpose
AB20060	CGCGGATCCGGCAGCTGCTTCTGCAGCAGC	Reverse primer HCF-1 aa1631 with BamHI site
AB20061	CGCGAATTCCAGGCCGCAGCCACGGAGGAA	Forward primer HCF-1 aa1632 with EcoRI and NcoI sites
AB20062	CGCGGATCCCTGCTGCTGGGCGTGGGCCTC	Reverse primer HCF-1 aa1717 with BamHI site
AB20617	CGCGAATTCCAGCCACCCCCACCCACACAG	Forward primer HCF-1 aa1580 with EcoRI site and NcoI sites
AB20618	CGCGGATCCCATGACGGCCTGCTGCGC	Reverse primer HCF-1 aa1656 with BamHI site
AB20874	CGCGGATCCTGTCAGCACAATGGGTATGGT	Reverse primer HCF-1 aa1695 with BamHI site
AB20875	CGCGGATCCCCCCAGCTCCGCCTGAGTCAC	Reverse primer HCF-1 aa1676 with BamHI site
AB20876	CGCGAATTCCAAGAGCTAATGGCCGAGGCC	Forward primer HCF-1 aa1598 with EcoRI site and NcoI sites
AB20877	CGCGAATTCGGGCTCACCCCCGAGGAGCTG	Forward primer HCF-1 aa1615 with EcoRI site and NcoI sites

Table 3. Oligonucleotides used during project for PCR amplification of HIZ fragments.

Synthesis#	Sequence	Purpose
AB22502	GCGAATTCGCCATGGCCCCACCGCCGTCAGG CCCAGCC	Forward primer HIZ aa997 with EcoRI and NcoI sites
AB22503	GCGGATCCTCAAAACGTGTGCACCAGCTG	Reverse primer HIZ aa1055 with BamHI site
AB22504 AB24022	GAATTCGCCATGGCCGGGGGAGCTGTGGCTA CTGAG	Forward primer HIZ aa876 with EcoRI and NcoI sites
AB22505 AB24023	GCGGATTCGGCCAGGAGCTGCAGCGTGGC	Reverse primer HIZ aa996 with BamHI site
AB23493	GCGGATCCAGTGCTGGCCTCGCCATCCCC	Reverse primer HIZ aa912 with BamHI site
AB23494	GCGGATCCGCCCTCGTCCGTCTGGAGTGT	Reverse primer HIZ aa928 with BamHI site
AB23495	GCGGATCCAAGTGTTTCTACCTCCTGTAC	Reverse primer HIZ aa953 with BamHI site
AB23496	GCGGATCCGCGGATGATGAGGAGTTTCTG	Reverse primer HIZ aa974 with BamHI site
AB23488	GCGAATTCGCCATGGCCGGTGTGGTCCAGGA TGTCCTC	Forward primer HIZ aa913 with EcoRI and NcoI sites
AB23489	GCGAATTCGCCATGGCCTTGCAGAGCGTGCT GGTGCTGAGC	Forward primer HIZ aa929 with EcoRI and NcoI sites
AB23491	GCGAATTCGCCATGGCCCCTCCTGGGCTGAC GGAGCCG	Forward primer HIZ aa954 with EcoRI and NcoI sites
AB23492	GCGAATTCGCCATGGCCAGCGCCCCAGCCAC TGAGCTG	Forward primer HIZ aa975 with EcoRI and NcoI sites

2.1.8 Plasmids

2.1.8.1 Vectors

PGBT9 (Clontech Laboratories). A bacterial and yeast expression vector containing the

GAL4 DNA binding domain (GAL4DBD; aa 1-147 of GAL4) upstream of the

multiple cloning site (MCS).

pGEX-4T-1 (Pharmacia Biotech): A glutathione-S-transferase (GST)-fusion vector

containing the GST gene adjacent to the MCS.

pSG424: A gift from Dr. John Hassell. A mammalian expression vector containing the

GAL4DBD upstream of the MCS as described in Sadowski and Ptashne, 1989.

pSPUTK: A gift from Dr. David Andrews. A transcription/translation vector containing

a MCS between SP6 and T7 promoters as described in Falcone and Andrews,

1991.

2.1.8.2 Plasmids Constructed by Others

pGEX-4T-1 HCF$_{750-902}$ Bacterial expression vector expressing GST-fused HCF$_{750-902}$. Constructed by David Piluso.

pLUC   Luciferase reporter gene tethered to a GAL4 responsive promoter (5 GAL4 binding sites upstream of a TATA box). A gift from Dr. John Hassell.

2.1.8.3 Plasmids Constructed During Project

GAL4DBD-HCF-1 derivatives constructed for activation domain mapping in yeast and for subcloning into mammalian expression vector pSG424:

pGBT9 HCF$_{1521-1622}$

pGBT9 HCF$_{1623-1708}$

pGBT9 HCF$_{1573-1647}$

pGBT9 HCF$_{1573-1708}$

pGBT9 HCF$_{1521-1641}$

pGBT9 HCF$_{1573-1726}$

pGBT9 HCF$_{1573-1668}$

pGBT9 HCF$_{1573-1686}$

pGBT9 HCF$_{1606-1708}$

pGBT9 HCF$_{1521-1726}$

GAL4DBD-HCF-1 fusions subcloned into a mammalian expression vector:

pSG424 HCF$_{1521\text{-}1622}$

pSG424 HCF$_{1623\text{-}1708}$

pSG424 HCF$_{1573\text{-}1647}$

pSG424 HCF$_{1573\text{-}1708}$

pSG424 HCF$_{1521\text{-}1641}$

pSG424 HCF$_{1573\text{-}1726}$

pSG424 HCF$_{1573\text{-}1668}$

pSG424 HCF$_{1573\text{-}1686}$

pSG424 HCF$_{1606\text{-}1708}$

pSG424 HCF$_{1521\text{-}1726}$

GAL4DBD-HIZ derivatives constructed for subcloning into a mammalian expression vector pSG424:

pGBT9 HIZ$_{929\text{-}974}$

pGBT9 HIZ$_{913\text{-}928}$

pGBT9 HIZ$_{913\text{-}953}$

pGBT9 HIZ$_{913\text{-}974}$

pGBT9 HIZ$_{913\text{-}996}$

pGBT9 HIZ$_{876\text{-}953}$

pGBT9 HIZ$_{876\text{-}974}$

pGBT9 HIZ$_{876\text{-}996}$

pGBT9 HIZ$_{876\text{-}1055}$

GAL4DBD-HCF-1 fusions subcloned into a mammalian expression vector:

pSG424 HIZ$_{929-974}$

pSG424 HIZ$_{913-928}$

pSG424 HIZ$_{913-953}$

pSG424 HIZ$_{913-974}$

pSG424 HIZ$_{913-996}$

pSG424 HIZ$_{876-953}$

pSG424 HIZ$_{876-974}$

pSG424 HIZ$_{876-996}$

pSG424 HIZ$_{876-1055}$

HIZ derivatives constructed for *in vitro* translation:

pSPUTK HIZ$_{913-996}$

pSPUTK HIZ$_{876-974}$

pSPUTK HIZ$_{913-974}$

pSPUTK HIZ$_{876-996}$

pSPUTK HIZ$_{876-1055}$

2.2 Methods

2.2.1 Plasmid Construction

All plasmids were constructed using standard procedures as described in Ausubel *et al.* (1998).

EcoRI restriction sites were introduced in all forward primers, while *NcoI* and *BamHI* sites were introduced in all reverse primers synthesized for cloning into the various vectors.

For yeast activation assays, fragments were amplified using PCR, restriction digested with *EcoRI* and *BamHI* and then tethered to the GAL4DBD by ligation into pGBT9 at the *EcoRI* and *BamHI* sites in the vector's MCS.

For mammalian activation assays, the fragments were first cloned into pGBT9 and then subloned into the mammalian expression vector pSG424 by restriction digest with *XhoI* and *BamHI*. (An *XhoI* site is found in the pGBT9 vector sequence 350bp from the *EcoRI* restriction site.)

For GST-binding assays, fragments were subcloned from either pGBT9 or pSG424 clones by cutting with *NcoI* and *BamHI*.

2.2.2 Small Scale Plasmid DNA purification

Small scale purification of plasmids was performed using the lysis by boiling method as outlined in Sambrook *et al.* (1989).

2.2.3 Large Scale Plasmid DNA purification

Large scale purification of plasmids was performed using DNA purification columns purchased from Qiagen and performed according to the manufacturer's maxiprep protocol.

2.2.4 DNA Quantification by Fluorometry

Fluorometry was used to quantify plasmid concentration of large scale purifications as outlined by the Hoefer Mini-Fluorometer (TKO 100) instruction manual. Calf thymus DNA of a known concentration was used as a standard.

2.2.5 Preparing Competent cells

Competent *E.Coli* DH5α bacterial cells were prepared according to the outline by Ausubel *et al.*, 1997.

2.2.6 Bacterial Transformation By Heat Shock

For bacterial transformation of plasmids, DNA (5-10ng) was added to competent *E.Coli* DH5α cells (100μL) and incubated on ice for 40 minutes. The cells were then heat shocked at 37°C for 40 seconds and further incubated on ice for 5 minutes. 2YT media (400μL) was added and the cells were incubated at 37°C in a shaker for 1 hour. 125μL of the 500μL mixture was plated onto a 2YT plate containing ampicillin (100μg/mL) and incubated overnight at 37°C.

2.2.7 Yeast Transformation (Lithium Acetate/PEG method)

Yeast transformation was performed using the lithium acetate method as outlined in Elble, 1992.

2.2.8 Qualitative Agarose Overlay Assay for Detection of β-galactosidase Activity

The agarose overlay assay was performed using a modified protocol described in Bohen and Yamamoto, 1993. An assay solution of 0.5% agarose, 0.5M $NaPO_4$ (pH 7.0), 0.1% SDS and 0.05% (v/v) of 2% X-gal was prepared and poured over plates of yeast transformants grown overnight. The plates were incubated at 37°C and assayed for colour change of the yeast to blue.

2.2.9 Quantitative β-galactosidase Liquid Assay

Transcriptional activation in yeast was measured quantitatively using the β-galactosidase liquid assay exactly as outlined by Ausubel et al., 1997. Yeast samples were measured twice in triplicate.

2.2.10 Overexpression and Purification of GST Fusion Protein

Overexpression and purification of the GST fusion protein GST-HCF$_{750-902}$ in the pGEX-4T-1 vector were carried out using a modified protocol provided by Pharmacia Biotech. A 50mL culture of Bl21 cells containing pGEX-4T-1 HCF$_{750-902}$ was grown overnight in 2YT supplemented with ampicillin (100µg/mL). The following morning, the 50mL culture was diluted in 500mL of 2YT supplemented with ampicillin (100µg/mL) and grown at 37°C for approximately 1.5 hours to an OD$_{600}$ of 0.7. At this point, the

500mL culture was induced with 0.1mM IPTG for 2 hours at 37°C and then chilled on ice for 10 minutes before centrifugation at 6000 rpm for 10 minutes. The supernatant was decanted and the bacterial pellet was washed with ice cold PBS, then resuspended in 10mL of NETN buffer (0.5% NP40, 1mM EDTA, 20mM Tris-HCl pH8.0, 100mM NaCl, 1 protease inhibitor cocktail tablet/40mL). The cells were lysed using a probe sonicator and centrifuged for 10 minutes at 10 000 rpm to remove cell debris. 1mL of the supernatant was incubated with 50μL of glutathione Sepharose 4B beads (bed volume of 25μL) at 4°C for 1.5 hours. Beads were then washed twice with 1X PBS buffer and proteins were eluted by boiling in 20μL of 6X SDS loading buffer and resolved on a 12% SDS-PAGE gel to detect the efficiency of overexpression and purification.

2.2.11 *In vitro* Transcription and Translation of HIZ

In vitro transcription and translation were performed using the Promega TNT coupled system as specified in the Promega protocol. A 50μL reaction consisting of 25μL of TNT rabbit reticulocyte lysate, 2μL TNT reaction buffer, 1μL TNT SP6 RNA polymerase, 1μL 1mM amino acid mixture minus leucine, 10μL of [^{14}C]-Leucine (300mCi/mmol) at 10μCi/mL, 1μL of RNAguard ribonuclease inhibitor and 1μg of DNA template (pSPUTK vector plus insert) was incubated at 30°C for 2 hours. 3μL of the reaction was the run on a 12% SDS-PAGE gel, dried and exposed to X-Omat AR film (Kodak) to determine transcription/translation efficiency.

2.2.12 *In vitro* Protein Binding Assay

Protein extracts containing the GST binding protein only and the GST-HCF$_{750-902}$ fusion were each incubated with 50μL of glutathione Sepharose 4B beads (25μL bed volume) for 1.5 hours at 4°C and then washed twice with PBS. The beads were then incubated with 10μL of [^{14}C]-Leucine labelled HIZ protein, and 200μL IPAB buffer (150mM KCl, 0.02mg/mL BSA, 0.1% Triton X-100, 0.1% NP40, 5mM MgCl$_2$, 20mM Hepes pH7.9) plus protease inhibitors (12.5μL of 1 Complete^{TM} mini tablet dissolved in 1mL PBS) for 1 hour at 4°C. The beads were then washed 5 times with IPAB buffer. Bound proteins were eluted by adding 20μL of 6X SDS loading buffer and boiling for 5 minutes, and were resolved on a 12% SDS-PAGE gel, dried and exposed to X-Omat AR film (Kodak) or exposed to a phosphorscreen for quantitative analysis using phosphorimaging (STORM, Molecular Dynamics).

2.2.13 Polymerase Chain Reaction

To attempt to amplify the full length of HIZ, various primers were used (listed in table 1) in varying concentrations and combinations. The commercially bought Anchor Primer 1 (AP1) was used for amplification using Marathon Ready testis cDNA (Clontech) as indicated by the supplier.

For constructing derivatives to map the HCF and HBZ activation domains, a standard PCR amplification protocol similar to that outlined by Ausubel *et al.* was followed. A 100μL reaction mixture composed of 10X Thermopol buffer, 1mM of each dNTP, 1ng of DNA template, 1μM of each primer, 2U of Vent DNA polymerase and sterilized water was made and underwent 30 cycles as summarized below:

Step	Temperature (°C)	Time (min)
Denaturation	94	1
Annealing	Lowest primer T_m minus 2	1
Extension	72	1 per kb

T_m = melting temperature (in °C)
  = 31.5 + 0.41(%GC) + 16.60log [salt in reaction buffer] - 500/primer length

2.2.14 Transient Transfection Assay

Cells were seeded onto 6 well plates (approximately 2.5×10^5 cells/well) and grown overnight to 70% confluence for transient transfection assays. The following morning, prior to transfection, the 6 well plates were washed twice with PBS and the medium was replaced with 1mL of unsupplemented DMEM per well.

Cells were transfected according to the protocol given by GibcoBRL for use with Lipofectamine reagent. Briefly, 1µg of the pLUC (luciferase reporter) plasmid, 0.05 – 1.5µg of various constructs and 8µL of Lipofectamine reagent were incubated in 400µL of unsupplemented DMEM for 30 minutes. Following the incubation, 1.6mL of unsupplemented DMEM was added to the 400µL mixture and the resulting 2mL was split into 1mL per well for duplicates of each transfection. The cells were incubated for 5 hours before the media was replaced with culture medium (DMEM supplemented with 10% v/v fetal bovine serum, 1% v/v L-glutamine, 1% v/v penicillin-streptomycin). The media was also changed the following day and cells were harvested 40-44 hours post-transfection.

2.2.15 Luciferase Assay

The reporter gene used in transient transfection assays for transactivation potential was luciferase. For the luciferase assay, cell monolayers were washed twice with PBS, resuspended in 400μL of 1X reporter lysis buffer, and then lysed by freeze-thaw 40-44 hours post-transfection. Samples were scraped from each well using a cell lifter, collected in microcentrifuge tubes and centrifuged for 2 minutes at 14 000rpm. Supernatants were collected and assayed for luciferase activity using the Lumat LB9507 luminometer (PerkinElmer Life Sciences), which injects 100μL of assay buffer (470μM luciferin, 270 μM coenzyme A, 530μM ATP, 33.3mM DTT, 20mM Tricine, 1.07mM $(MgCO_3)_4Mg(OH)_2\cdot5H_2O$, 2.67mM $MgSO_4$, 0.1mM EDTA) to each lysate sample (20μL) and measured luciferase activity as relative light units (R.L.U.). To normalize luciferase activity for transfection efficiency, protein concentrations of lysates were determined using the Biorad/Bradford protein assay, performed as outlined by BioRad.

2.2.16 Western Immunoblotting

For preparation of cell lysates for Western immunoblotting, cell monolayers were washed twice with PBS and resuspended in 200μL of RIPA buffer (50mM TrisHCl pH7.2, 150mM NaCl, 0.1% SDS, 1% sodium deoxycholate, 1mM PMSF) 48 hours post-transfection. Cells were lysed by freeze-thawing followed by brief sonication. Samples were scraped from each well using a cell lifter and then collected in microcentrifuge tubes, then centrifuged for 2 minutes at 14 000rpm and the supernatant was collected. Protein concentration of the lysates was determined and 25μg of protein in each lysate

sample was resolved on a 12.5% SDS-PAGE gel concurrently with Benchmark^{TM} prestained markers. Proteins were transferred overnight at 23V to pure nitrocellulose membrane. The membranes were blocked the following morning using 3% skim milk solution in TBST buffer (20mM TrisHCl pH7.5, 150mM NaCl, 0.05% Tween20) for 1 hour, then washed using TBST buffer twice for 15 minutes each time. The membrane was probed with primary anti-gal4DNA-BD mouse monoclonal antibody diluted to 0.1µg/mL in 1% skim milk solution (in TBST buffer) and incubated for one hour at room temperature. The membrane was then washed with TBST twice for 15 minutes as previously done and incubated with secondary sheep anti-mouse polyclonal antibody coupled to horseradish peroxidase diluted 1:10 000 in TBST for 45 minutes. Following another wash as before, ECL reagent was prepared and applied to the membrane as specified by the manufacturer. The membrane was then exposed to X-Omat Blue XB-1 film (Kodak) for 30-60 minutes to detect GAL4DNA-BD protein fusions.

2.2.17 Multiple Tissue Northern Blots and Matched Tumour/Normal Array Hybridization

Multiple Tissue Northern (MTN) blots I (mRNA from the heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas) and II (mRNA from the spleen, thymus, prostate, testis, ovary, colon, small intestine and leukocyte) and a matched tumour/normal expression array were purchased from Clontech and probed with a radioactive portion of the HIZ carboxy-terminus.

The probes used were constructed using PCR amplification. Oligonucleotides AB22504 (forward primer) and AB22505 (reverse primer) were used to amplify HIZ$_{799-}$

919 for hybridization with the MTN blots and oligonucleotides AB22504 (forward primer) and AB22503 (reverse primer) were used to construct $HIZ_{799-978}$ hybridization with the matched tumour/normal expression array. The probes were radiolabelled using $[\alpha-^{32}P]$-dATP and purified using gel filtration on sephadex G-50.

The hybridization and subsequent washes were performed at high stringency as outlined by the manufacturer. The matched tumour/normal array was hybridized using ExpressHyb solution, according to the manufacturer's protocol.

CHAPTER THREE

RESULTS

3.1 HCF-1 Basic Domain Interacts with HIZ Using the Yeast Two-Hybrid System

In addition to the HCF_{VIC}, the basic domain of HCF has been implicated in cell cycle control because it was required for cell cycle rescue in tsBN67 cells (Goto *et al.,* 1997). To identify a role for this domain, we attempted to detect novel interactions with this region. We used the yeast two-hybrid system to screen a HeLa cDNA library with a region encompassing the basic domain as bait ($HCF_{450-1439}$). A 732bp cDNA insert including part of a poly A tail was identified to encode a potential HCF-1-binding protein. The two-hybrid screen was performed by David Piluso. David has further mapped the minimal domain of HCF-1 required for maximum binding of HIZ to $HCF_{750-902}$ using the yeast two-hybrid assay (see table 4).

3.2 HCF-1 Interacting Zinc Finger Protein (HIZ)

3.2.1 HIZ Putative Protein Sequence

The sequence of the 732bp cDNA insert identified to interact with the HCF basic domain did not match any genes deposited in the Genbank database. However, it matched 100% with the contiguous sequence (contig) NT_011150 on chromosome 19.

Using the genomic region upstream from our cDNA and the gene finder programs: Genscan (http://genes.mit.edu/GENSCAN.html) (Burge and Karlin, 1997) and

Table 4. The Minimal Region of HCF-1 Required for Maximum HIZ Interaction is within Residues 750-902. Using the yeast two-hybrid assay, David Piluso demonstrated that regions smaller 750-902 were able to bind but with less efficiency. (-) represents no binding and (++++) represents maximum binding.

HCF-1 region	Relative Level of HIZ Binding
450-1439	+++
450-1011	+++
450-902	+++
750-902	++++
450-836	+++
750-836	+++
836-902	-/+
750-816	-/+
750-811	-
750-806	-
770-836	++
790-836	-

```
   1 - ATGGTCGGCTCCCACGCGGACATGGCGCCGGCCTCTACTGCGGAGGGGGCCGGGGAGAAG -  60
   1 - M  V  G  S  H  A  D  M  A  P  A  S  T  A  E  G  A  G  E  K  -  20
                                   *
  61 - CCAGGCCCTGCGGCCCCTGCCCCGGCGGCCCAGTACGAATGTGGGGAGTGTGGCAAGTCA - 120
  21 - P  G  P  A  A  P  A  P  A  A  Q  Y  E  C  G  E  C  G  K  S  -  40
                                                               *
 121 - TTCCGGTGGTCGTCCCGGCTCCTGCACCACCAGCGCACGCACACAGGCGAGCGGCCCTAC - 180
  41 - F  R  W  S  S  R  L  L  H  H  Q  R  T  H  T  G  E  R  P  Y  -  60
                *
 181 - AAGTGCCCAGACTGCCCCAAGGCTTTCAAAGGCTCCTCGGCCCTGCTCTACCACCAGCGA - 240
  61 - K  C  P  D  C  P  K  A  F  K  G  S  S  A  L  L  Y  H  Q  R  -  80

 241 - GGCCACACGGGCGAGCGGCCCTACCAGTGCCCCGACTGTCCCAAGGCCTTCAAGCGCTCC - 300
  81 - G  H  T  G  E  R  P  Y  Q  C  P  D  C  P  K  A  F  K  R  S  - 100

 301 - TCTCTGCTGCAGATCCACCGTAGCGTGCACACCGGCCTGCGGGCCTTCATCTGCGGCCAG - 360
 101 - S  L  L  Q  I  H  R  S  V  H  T  G  L  R  A  F  I  C  G  Q  - 120
           *
 361 - TGCGGCCTGGCCTTCAAGTGGTCGTCCCACTACCAGTACCACTTAAGGCAGCACACAGGC - 420
 121 - C  G  L  A  F  K  W  S  S  H  Y  Q  Y  H  L  R  Q  H  T  G  - 140

 421 - GAGCGCCCCTACCCGTGCCCGGACTGCCCCAAGGCCTTCAAGAACTCGTCCAGCCTGCGG - 480
 141 - E  R  P  Y  P  C  P  D  C  P  K  A  F  K  N  S  S  S  L  R  - 160
                                                            *
 481 - CGCCACCGCCACGTGCACACCGGCGAGCGGCCCTACACCTGTGGAGTCTGCGGGAAGAGC - 540
 161 - R  H  R  H  V  H  T  G  E  R  P  Y  T  C  G  V  C  G  K  S  - 180

 541 - TTCACGCAGAGCACCAACCTGCGGCAGCACCAGCGCGTGCACACGGGCGAGCGGCCCTTC - 600
 181 - F  T  Q  S  T  N  L  R  Q  H  Q  R  V  H  T  G  E  R  P  F  - 200

 601 - CGCTGCCCGCTCTGCCCCAAGACCTTCACCCACTCCTCCAACCTGCTGCTGCACCAGCGC - 660
 201 - R  C  P  L  C  P  K  T  F  T  H  S  S  N  L  L  L  H  Q  R  - 220

 661 - ACGCACGGCGCCGCCCCCGCCCCGGGTACCGCCTCCGCGGCCCCGCCCCCCCAGTCCCGG - 720
 221 - T  H  G  A  A  P  A  P  G  T  A  S  A  A  P  P  P  Q  S  R  - 240
                                                            *
 721 - GAGCCCGGCAAGGTCTTCGTGTGCGACGCCTACCTGCAGCGGCACCTCCAGCCCCACAGC - 780
 241 - E  P  G  K  V  F  V  C  D  A  Y  L  Q  R  H  L  Q  P  H  S  - 260

 781 - CCGCCCGCGCCTCCCGCCCCGCCGCCCCCGCCCCCGCCCGTGGTGCCTGAGCTCTTTTTG - 840
 261 - P  P  A  P  P  A  P  P  P  P  P  P  P  V  V  P  E  L  F  L  - 280

 841 - GCGGCGGCGGAGGCCACGGTGGAGCTGGTGTACCGCTGCGATGGCTGCGAGCAGGGATTC - 900
 281 - A  A  A  E  A  T  V  E  L  V  Y  R  C  D  G  C  E  Q  G  F  - 300

 901 - AGCAGCGAGGAGCTGCTCCTGGAGCACCAGCCGTGCCCCGGGCCCGATGCGGCGCCCCAG - 960
 301 - S  S  E  E  L  L  L  E  H  Q  P  C  P  G  P  D  A  A  P  Q  - 320
           *
 961 - CCCCAGGAGGCACCCGCCGAGGCGCCCAAGGCCGACCAGCCACCGTCCCCTCTGCCGCAG - 1020
 321 - P  Q  E  A  P  A  E  A  P  K  A  D  Q  P  P  S  P  L  P  Q  - 340
                                                   *
1021 - CCCCCTCCTCCCGCCGCCGCCCCCGCGCCTGGCTTTGCCTGTCTGCCCTGCGGCAAGTCC - 1080
 341 - P  P  P  P  A  A  A  P  A  P  G  F  A  C  L  P  C  G  K  S  - 360

1081 - TTCCGGACGGTGGCTGGGCTCTCCCGCCACCAGCACAGCCACGGGGCTGCCGGCGGGCAA - 1140
 361 - F  R  T  V  A  G  L  S  R  H  Q  H  S  H  G  A  A  G  G  Q  - 380
                                             *
```

```
1141 - GCGTTCCGCTGCGGCAGCTGCGACGGCTCCTTCCCGCAGCTGGCCAGCCTCCTGGCGCAT - 1200
 381 - A  F  R  C  G  S  C  D  G  S  F  P  Q  L  A  S  L  L  A  H  - 400

1201 - CAGCAGTGCCACGTGGAAGAGGCCGCGGCCGGGCGCCCGCCCCCGCAGGCTGAGGCTGCG - 1260
 401 - Q  Q  C  H  V  E  E  A  A  A  G  R  P  P  P  Q  A  E  A  A  - 420

1261 - GAGGTGACCTGCCCCCAGGAACCGCTGGCGCCTGCCGCCCCCGTCCCGCCGCCACCCCCG - 1320
 421 - E  V  T  C  P  Q  E  P  L  A  P  A  A  P  V  P  P  P  P  P  - 440

1321 - TCCGCCCCCGCTTCTGCGGAGCGGCCCTACAAATGTGCCGAGTGCGGCAAGTCCTTCAAG - 1380
 441 - S  A  P  A  S  A  E  R  P  Y  K  C  A  E  C  G  K  S  F  K  - 460

1381 - GGCTCCTCCGGGCTGCGCTACCACCTGCGGGACCACACGGGCGAGCGGCCCTACCAGTGT - 1440
 461 - G  S  S  G  L  R  Y  H  L  R  D  H  T  G  E  R  P  Y  Q  C  - 480

1441 - GGCGAGTGCGGCAAGGCCTTCAAGCGCTCCTCCCTGCTGGCCATCCACCAGCGGGTGCAC - 1500
 481 - G  E  C  G  K  A  F  K  R  S  S  L  L  A  I  H  Q  R  V  H  - 500
                                      *
1501 - ACGGGCCTGCGGGCCTTCACCTGTGGCCAGTGCGGCCTCACCTTCAAGTGGTCGTCCCAC - 1560
 501 - T  G  L  R  A  F  T  C  G  Q  C  G  L  T  F  K  W  S  S  H  - 520

1561 - TACCAGTACCACCTGCGGCTGCACTCTGGCGAGCGGCCCTACGCCTGCGGGGAGTGTGGC - 1620
 521 - Y  Q  Y  H  L  R  L  H  S  G  E  R  P  Y  A  C  G  E  C  G  - 540
                                      *
1621 - AAGGCCTTCCGCAACACGTCGTGCCTGCGTCGCCACCGCCACGTGCACACTGGCGAGAGG - 1680
 541 - K  A  F  R  N  T  S  C  L  R  R  H  R  H  V  H  T  G  E  R  - 560

1681 - CCCCACGCCTGCGGTGTCTGCGGCAAGAGCTTCGCGCAGACCTCCAACCTGCGGCAGCAC - 1740
 561 - P  H  A  C  G  V  C  G  K  S  F  A  Q  T  S  N  L  R  Q  H  - 580

1741 - CAGCGCGTGCACACGGGCGAGCGGCCCTTCCGCTGCCCGCTCTGCCCCAAGACCTTCACC - 1800
 581 - Q  R  V  H  T  G  E  R  P  F  R  C  P  L  C  P  K  T  F  T  - 600

1801 - CACTCCTCCAACCTGCTGCTGCACCAGCGCACGCACTCGGCGGAGCGCCCCTTCACCTGC - 1860
 601 - H  S  S  N  L  L  L  H  Q  R  T  H  S  A  E  R  P  F  T  C  - 620
                                      *
1861 - CCCATCTGCGGTCGCGGCTTCGTTATGGCCGCCTATCTGCAGCGGCACCTGAGGACGCAC - 1920
 621 - P  I  C  G  R  G  F  V  M  A  A  Y  L  Q  R  H  L  R  T  H  - 640

1921 - GCCCCGGCCAACACGCCTCCCAGCACCACAGCCCCTGCCGCCGGCCCCCAGCCCCCTGCT - 1980
 641 - A  P  A  N  T  P  P  S  T  T  A  P  A  A  G  P  Q  P  P  A  - 660

1981 - CCACTGGCTGCTGCGCGGGCCCCGCCAGCCACCCAAGATGTCCACGTCCTGCCCCACCTC - 2040
 661 - P  L  A  A  A  R  A  P  P  A  T  Q  D  V  H  V  L  P  H  L  - 680

2041 - CAGGCCACGCTCTCCCTCGAGGTGGCGGGGGGCACGGCCCAGGCCCCGAGCTTGGGGCCA - 2100
 681 - Q  A  T  L  S  L  E  V  A  G  G  T  A  Q  A  P  S  L  G  P  - 700

2101 - GCAGCGCCCAACTCTCAGACGTTCCTCCTGGTGCAAACTGCCCAGGGCCTCCAGCTGATC - 2160
 701 - A  A  P  N  S  Q  T  F  L  L  V  Q  T  A  Q  G  L  Q  L  I  - 720

2161 - CCCAGCAGCGTGCAGCCCCCTACACCTCCGCCCCCTCCCGCACCTCCCAAGCTCATCCTG - 2220
 721 - P  S  S  V  Q  P  P  T  P  P  P  P  P  A  P  P  K  L  I  L  - 740

2221 - CTGCCCTCCTCCAGTGCTGGGGCTGGGGGCGGCCGTGCAAGGCAGGGCCCGCGGGCAGTG - 2280
 741 - L  P  S  S  S  A  G  A  G  G  G  R  A  R  Q  G  P  R  A  V  - 760
```

```
2281 - GGCAAAGCGGGCCAGGGGGCGGGAGTGGTCTGGCTGCCAGGCCCTGGGGGTCTAGGGGTG - 2340
 761 - G  K  A  G  Q  G  A  G  V  V  W  L  P  G  P  G  G  L  G  V  - 780

2341 - CACGGAGCGGCCAGCGCTGGGGCCAGCGGGACAGGGCAGAGCCTCATCGTTCTGCAGAAT - 2400
 781 - Q  G  A  A  S  A  G  A  S  G  T  G  Q  S  L  I  V  L  Q  N  - 800

2401 - GTGGGGGGTGGGGAGGCAGGGCCACAGGAAATGAGTGGGGTGCAGCTCCAGCCCCTCCGA - 2460
 801 - V  G  G  G  E  A  G  P  Q  E  M  S  G  V  Q  L  Q  P  L  R  - 820

2461 - CCAGCCCCAGAAGTAACCACGGTCCAGCTCCAGCCAGCGCAGGAGGTGACCACAGTCCAG - 2520
 821 - P  A  P  E  V  T  T  V  Q  L  Q  P  A  Q  E  V  T  T  V  Q  - 840

2521 - CTCCAGCCAGCACAGGAAGTAACCACGGTCCAGCTCCAGCCAGCACAGGAGGTGACCACG - 2580
 841 - L  Q  P  A  Q  E  V  T  T  V  Q  L  Q  P  A  Q  E  V  T  T  - 860

2581 - GTCCAGCTCCAGCCCGTGGCCGGCCAGCTCTCCAATTCCAGTGGGGGAGCTGTGGCTACT - 2640
 861 - V  Q  L  Q  P  V  A  G  Q  L  S  N  S  S  G  A  V  A  T  - 880

2641 - GAGGCACCCAACCTGCTGGTTGTTCAGAGCGGGGCAGCTGAGGAGTTGCTCACTGGCCCG - 2700
 881 - E  A  P  N  L  L  V  V  Q  S  G  A  A  E  E  L  L  T  G  P  - 900

2701 - GGCCCCGGGGAGGCGGGGGATGGCGAGGCCAGCACTGGTGTGGTCCAGGATGTCCTCTTT - 2760
 901 - G  P  G  E  A  G  D  G  E  A  S  T  G  V  V  Q  D  V  L  F  - 920

2761 - GAGACACTCCAGACGGACGAGGGCTTGCAGAGCGTGCTGGTGCTGAGCGGGGCCGATGGC - 2820
 921 - E  T  L  Q  T  D  E  G  L  Q  S  V  L  V  L  S  G  A  D  G  - 940
                                                              *
2821 - GAACAGACTCGACTCTGCGTACAGGAGGTAGAAACACTTCCTCCTGGGCTGACGGAGCCG - 2880
 941 - E  Q  T  R  L  C  V  Q  E  V  E  T  L  P  P  G  L  T  E  P  - 960

2881 - CCTGCCACCGGCCCACCCGGACAGAAACTCCTCATCATCCGCAGCGCCCCAGCCACTGAG - 2940
 961 - P  A  T  G  P  P  G  Q  K  L  L  I  I  R  S  A  P  A  T  E  - 980

2941 - CTGCTGGACAGCAGCAACACTGGAGGAGGCACCGCCACGCTGCAGCTCCTGGCCCCACCG - 3000
 984 - L  L  D  S  S  N  T  G  G  G  T  A  T  L  Q  L  L  A  P  P  - 1000

3001 - CCGTCAGGCCCAGCCTCGGGCCCCGCGGGGCTCCCCGGGGCTCCAGCCTCCCAGATGGTG - 3060
1001 - P  S  G  P  A  S  G  P  A  G  L  P  G  A  P  A  S  Q  M  V  - 1020

3061 - CAAGTGGTCCCCGCAGGAGCTGGGCCTGGTGTTATGACCCCTCAGGGCCTGCCCTCCATC - 3120
1021 - Q  V  V  P  A  G  A  G  P  G  V  M  T  P  Q  G  L  P  S  I  - 1040

3121 - CAGATTGTCCAGACTCTACCCGCAGTCCAGCTGGTGCACACGTTTTGAGGAGAGGCAGTG - 3180
1041 - Q  I  V  Q  T  L  P  A  V  Q  L  V  H  T  F  stop - 1055

       ATTCCCCTCCCGCCCCGCACAGAGACCCCAACTCACTGCCAGCCGGGGCGGGGCAGGGTG - 3240
       CCGCAGGCTGGGCTTGCTAATAAAGACCCGAGTCTCCCCAAAAAAAAAAAAAAAAAAAAA - 3300
       AAA   - 3303
```

Figure 8. Putative Full-Length Amino Acid and Nucleotide Sequences of HIZ. Zinc fingers are underlined and in bold. The one fragmented zinc finger structure is in bold only. The glutamine-rich motif: EVTTVQLQPAQ in the carboxy-terminal is underlined. The transactivation domain is situated within residues 876-996. The * represents putative phosphorylation sites as determined by NetPhos 2.0.

Fgenesh++ (www.softberry.com/nucleo.html) (Lander *et al.*, 2001), a putative mRNA sequence for HIZ was deduced (see figure 8). These two programs predict genes based on homology to known proteins and consensus sequences in introns. The predicted HIZ sequence is strikingly similar in size and structure to a cloned chicken zinc finger protein, cKr2. This chicken transcription factor is a nuclear-localized 1173 amino acid protein with a time and tissue specific expression pattern that suggests a role in differentiating Schwann cells (Schutz and Niessing, 1992).

Like HIZ, cKr2 contains 16 zinc fingers divided into two clusters of 7 that are similar in sequence and separated by unlinked ZFs (see figures 9 and 10, table 5). The cKr2 sequence also encodes 16 ZFs: 6 linked ZFs (ZF 1-6) in block one, 7 linked ZFs in block two (ZF 10-16) and 3 in between. Schutz and Niessing suspected that their first cluster was missing a ZF that would be a counterpart to ZF 10 based on high homologies between ZFs 1 and 11, 2 and 12, 3 and 13 etc. Their amino acid sequence supports this because it is missing an initiation codon and begins with the linker sequence (TGERP). Within amino acids 291-311 in HIZ, there is also another putative zinc finger sequence. It contains all the conserved residues of a ZF except for the last histidine. It has the sequence: YRCDGCEQGFSSEELLLEHQPCP. Because this putative sequence is based on a draft human genome sequence, it is likely that there was a single nucleotide error in the sequencing changing the actual code 'cac' which codes for histidine (H) to 'ccc', the code for proline. Therefore, both HIZ and cKr2 likely have 17 ZF structures: cKr2 with an extra ZF structure in its first cluster of 6 for a total of 7, and HIZ with an extra ZF between the 2 ZF clusters for a total of independent 3 ZF structures in between the two clusters.

```
Identities = 120/143 (83%), Positives = 127/143 (87%)
```



```
HIZ:  82 HTGERPYQCPDCPKAFKRSSLLQIHRSVHTGLRAFICGQCGLAFKWSSHYQYHLRQHTGE 141
         HTGERPY+C +CPKAFKRSSLLQIH+SVHTGLRAF C  CGLAFKWSSHYQYHLRQHTGE
cKr2:  5 HTGERPYKCSECPKAFKRSSLLQIHQSVHTGLRAFKCALCGLAFKWSSHYQYHLRQHTGE 64

      142 RPYPCPDCPKAFKNSSSLRRHRHVHTGERPYTCGVCGKSFTQSTNLRQHQRVHTGERPFR 201
          RPY C  CPKAFKNSSSLRRHRH+HTGERPY C  CGK+FTQSTNLRQHQR HTGERP+
       65 RPYKCTSCPKAFKNSSSLRRHRHIHTGERPYVCSACGKAFTQSTNLRQHQRTHTGERPYA 124

      202 CPLCPKTFTHSSNLLLHQRTHGA 224
          C  C KTFTHSSNLLLHQRTH +
      125 CSHCSKTFTHSSNLLLHQRTHSS 147
```

Figure 9. HIZ Zinc Fingers 3-7 are Very Similar to cKr2 Zinc Fingers 1-5. A BLAST search identified a highly conserved region within residues 82-224 of HIZ. This region encodes zinc fingers 3-7 and is 83% similar in amino acid sequence to residues 5-147 in cKr2, which encodes zinc fingers 1-5 of cKr2.

```
Identities = 158/208 (75%), Positives = 175/208 (83%), Gaps = 2/208 (0%)
```
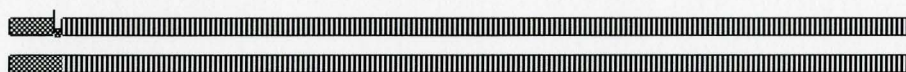


```
HIZ:  436  PPPPPSAPAS--AERPYKCAECGKSFKGSSGLRYHLRDHTGERPYQCGECGKAFKRSSLL 493
           PPP    P S  +ERPY+C ECGK+FKGSSGLRYH+RDHTGERPY+C EC KAFKRSSLL
cKr2: 296  PPPATPTPTSVPSERPYQCTECGKAFKGSSGLRYHMRDHTGERPYKCSECPKAFKRSSLL 355

      494  AIHQRVHTGLRAFTCGQCGLTFKWSSHYQYHLRLHSGERPYACGECGKAFRNTSCLRRHR 553
           AIHQRVHTGLRA+ C  CGLTFKWSSH QYHLRLH+GERPY C +C KAF+NTSCL  HR
      356  AIHQRVHTGLRAYKCPSCGLTFKWSSHLQYHLRLHTGERPYRCPDCPKAFKNTSCLGPHR 415

      554  HVHTGERPHACGVCGKSFAQTSNLRQHQRVHTGERPFRCPLCPKTFTHSSNLLLHQRTHS 613
           +HTGERPHAC +CGK+F QTSNLRQHQR HTGERP+ C  C KTFTHSSNL LHQRTHS
      416  QLHTGERPHACPICGKAFTQTSNLRQHQRTHTGERPYACSHCGKTFTHSSNLQLHQRTHS 475

      614  AERPFTCPICGRGFVMAAYLQRHLRTHA 641
           + RP  CP+C + FVMA+YLQRHLRTHA
      476  SARPHQCPLCPKAFVMASYLQRHLRTHA 503
```

Figure 10.   The Last Seven Zinc Fingers of HIZ and cKr2 are Highly Conserved.   A BLAST search identified a strong match between the last seven zinc fingers, which are encoded by residues 436-641 in HIZ and 295-503 in cKr2.

Table 5. A Summary of the 16 Zinc Finger Sequences and their Linker Sequences in HIZ.
ZFs 1-7 and 10-16 form two clusters. ZF's 8 and 9 are not linked to any other ZF. The residues
in bold represent the conserved residues in the consensus sequence. ZF = zinc finger.

ZF Number	Region (aa)	ZF Sequence	Linker Sequence
1	32-54	YECGECGKSFRWSSRLLHHQRTH	
			TGERP
2	60-82	YKCPDCPKAFKGSSALLYHQRGH	
			TGERP
3	88-110	YQCPDCPKAFKRSSLLQIHRSVH	
			TGLRA
4	116-138	FICGQCGLAFKWSSHYQYHLRQH	
			TGERP
5	144-166	YPCPDCPKAFKNSSSLRRHRHVH	
			TGERP
6	172-194	YTCGVCGKSFTQSRNLRQHQRVH	
			TGERP
7	200-222	FRCPLCPKTFTHSSNLLLHQRTH	
8	352-374	FACLPCGKSFRTVAGLSRHQHSH	
9	382-404	FRCGSCDGSFPQLASLLAHQQCH	
10	450-472	YKCAECGKSFKGSSGLRYHLRDH	
			TGERP
11	478-500	YQCGECGKAFKRSSLLAIHQRVH	
			TGLRA
12	506-528	FTCGQCGLTFKWSSHYQYHLRLH	
			SGERP
13	534-556	YACGECGKAFRNTSCLRRHRHVH	
			TGERP
14	562-584	HACGVCGKSFAQTSNLRQHQRVH	
			TGERP
15	590-612	FRCPLCPKTFTHSSNLLLHQRTH	
			SAERP
16	618-640	FTCPICGRGFVMAAYLQRHLRTH	

The linker sequence in cKr2 and HIZ both have the lysine (K) in the typical linker sequence TGEKPY (Schuh *et al.*, 1986) replaced by arginine (R). Both residues have similar properties (large, hydrophilic, basic and positively charged) but it is not known why these two proteins use one instead of the other.

In addition to its similarities to cKr2, the HIZ sequence around the putative ATG initiation codon (**GTGATGG**) conform to Kozak's consensus sequence for translation initiation, with a purine at position -3 (3 nucleotides from the A in the ATG codon) and a G at position +4 (Kozak, 1986). Also, stop codons appear in all three frames upstream of this putative start site.

It is important to note that there is an open reading frame shift between the two ZF clusters. Our putative sequence, as predicted by Fgenesh++, changes the ORF at nt851 after coding the glutamic acid, by skipping over nt852 and starts coding in a new frame at position 853.

As mentioned above, the 16 zinc fingers are in two clusters of seven and separated by two (potentially three) unlinked zinc fingers (see table 5). It is interesting to note that within one of the unlinked zinc fingers (ZF#9) is an LxxLL motif (LASLL) (see figure 11). The LxxLL motif has been demonstrated in a number of transcriptional coactivators as a nuclear hormone receptor binding motif (Heery *et al.*, 1997). This putative nuclear receptor binding motif has also been identified in the HCF-binding proteins LZIP, Miz-1 and PGC-1β (Luciano and Wilson, 2000; Lin *et al.*, 2002).

By studying the primary sequence of HIZ, we have identified distinctive regions in HIZ that are enriched in certain residues (see figure 12). After each group of ZFs are
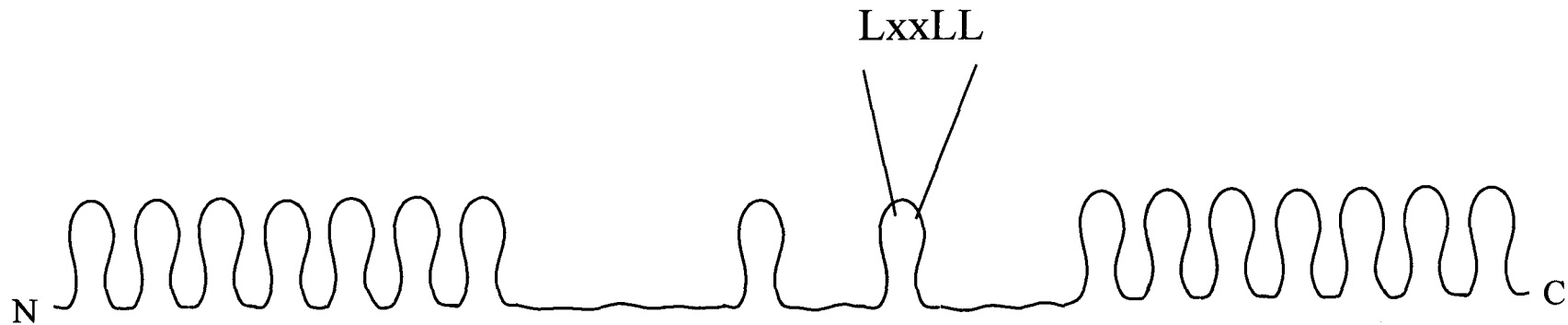
LxxLL

Figure 11. Schematic representation indicating the location of the LxxLL motif in HIZ. HIZ contains 16 zinc finger structures, 14 of which are divided into two clusters of 7 zinc fingers linked by the sequence TGERP. The remaining 2 zinc finger structures (numbers 8 and 9) occur between the two clusters and are independent. The LxxLL motif occurs within the ninth zinc finger.
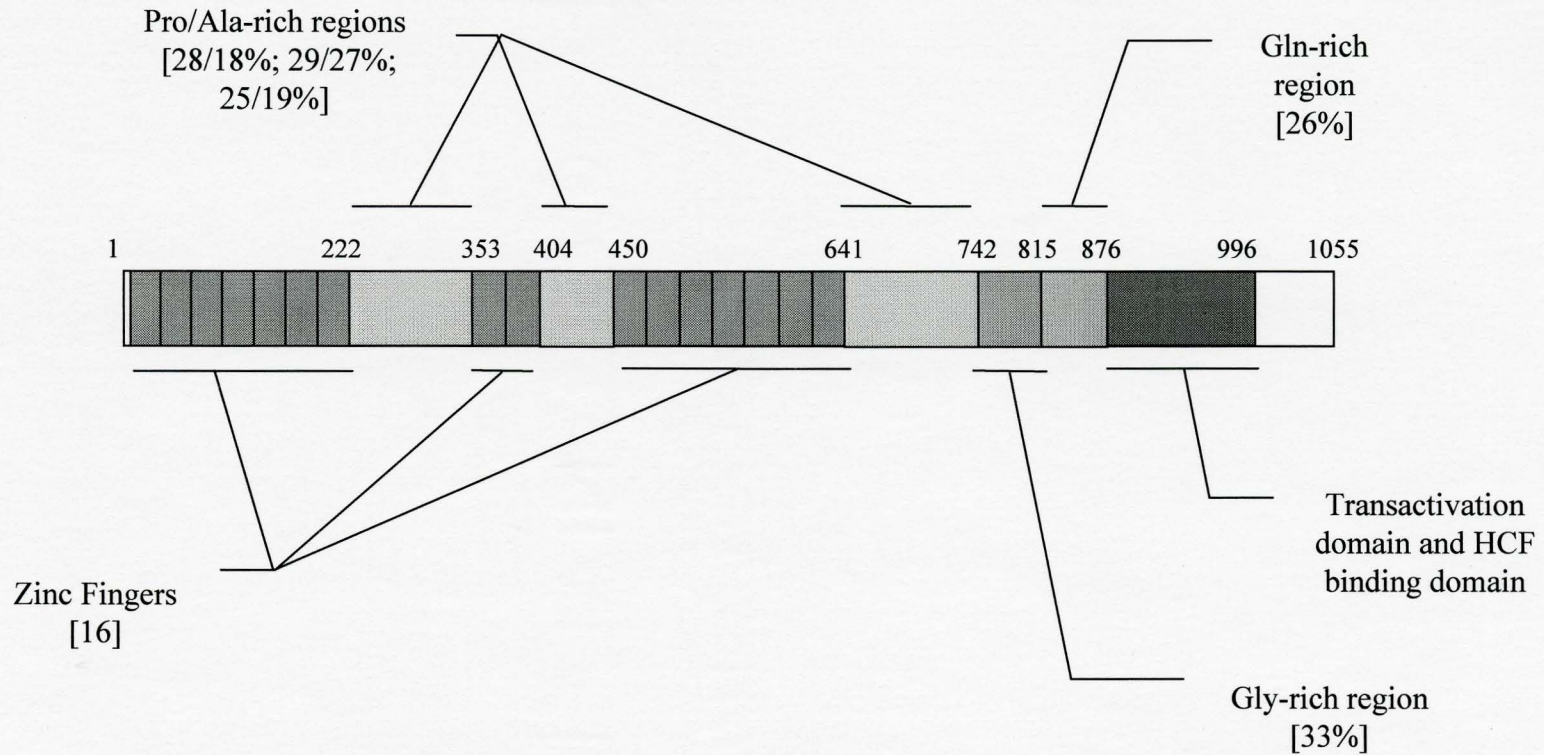
Figure 12. HIZ domain structure. The putative amino acid sequence shows distinct regions in HIZ. Sixteen C_2H_2-type zinc fingers span the amino terminus in two clusters separated by proline/alanine-rich regions and two unlinked zinc fingers, one of which contains an LxxLL motif (ZF#9). There is also a region enriched in glutamine which contains a novel motif repeated three and a half times. In the carboxy-terminus is a potent transactivation domain which corresponds to the region critical for binding HCF-1.

regions rich in proline and alanine. There is also a small region of less than 100 amino acids that is enriched in glycine residues (33%) and a glutamine-rich region (26%) adjacent to a transactivation domain that also contains the novel glutamine-rich motif.

Recently, an mRNA sequence was deposited into Genbank (accession no. AF367249.1) which corresponds to nt 2374-3000 of our putative HIZ sequence. It is expressed in the liver but this sequence is likely incomplete because it lacks a start site.

3.2.2 HIZ Contains Evolutionarily Conserved Regions

From a BLAT search on the human genome browser website http://genome.ucsc.edu (Lander *et al.*, 2001), constructed by the University of California, Santa Cruz (UCSC) Centre for Biomolecular Science and Engineering (CBSE), we discovered that our putative amino acid sequence was recognized as a putative gene in the human genome. On their website is a BLAT search program developed at UCSC, which is similar to the NCBI BLAST but uses the current assembly of human genome sequences as their target database instead of GenBank sequences. A BLAT search of our putative sequence of HIZ is shown in figure 13. This search shows the incomplete mRNA sequence deposited into Genbank in June, 2001 (accession no. AF367249). The nonhuman mRNA sequences from Genbank showing alignment to the HIZ sequence represent similar zinc finger regions in the chicken cKr2 protein and in various *Drosophila* proteins.

It is also interesting that there are conserved regions between the predicted HIZ sequence and the fish *Tetradon nigroviridis*. This is a fish that has been found to contain a genome 8 times smaller than the human genome because of its remarkably low content
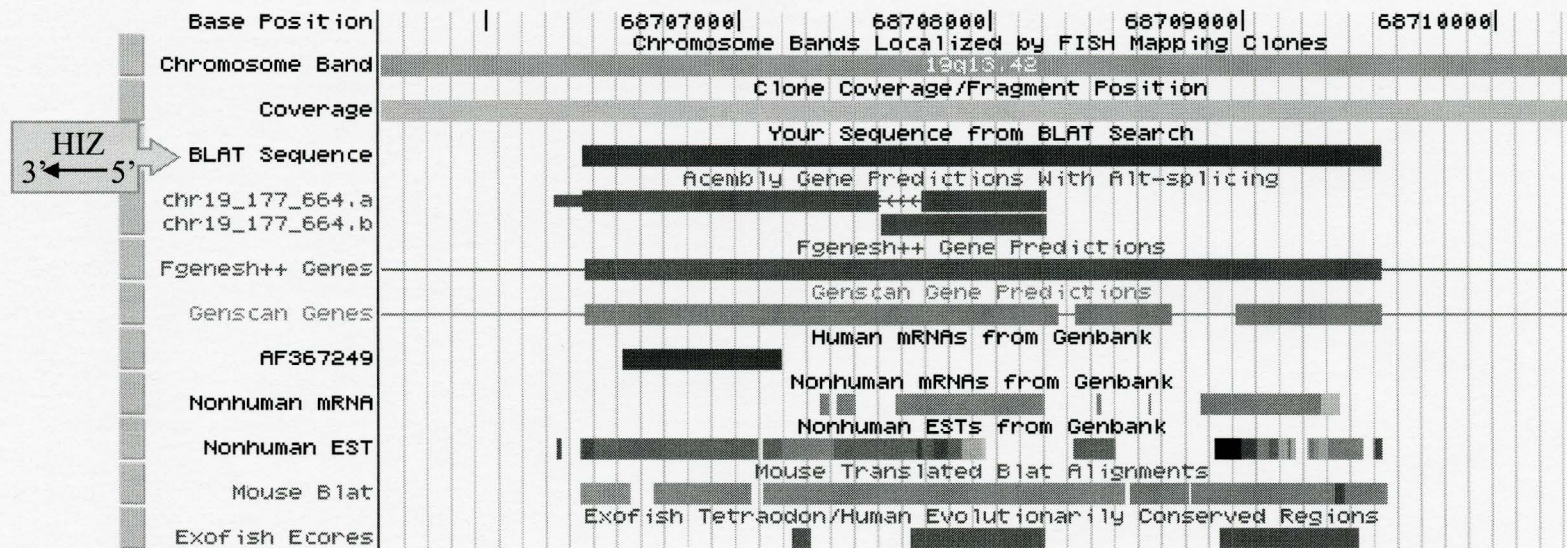
Figure 13. HIZ contains evolutionarily conserved regions from *Drosophila,* mouse, fish and chicken. A BLAT search (similar to NCBI's BLAST search) using the HIZ putative full length nucleotide sequence, recognized various regions of HIZ as significantly similar to proteins found in *Drosophila* and chicken, as well as regions in the mouse and *Tetradon nigroviridis* fish genome. This program also mapped the HIZ sequence to the chromosomal position 19q13.42.

of repetitive DNA sequences, yet contains a gene pool similar to human and mice. Therefore, its genome is currently being sequenced by Genoscope and the Whitehead Institute for Genomic Research. A program called Exofish has been developed by Genoscope (http://www.genoscope.cns.fr/externe/English/Site/site.html) to aid in the identification of novel human genes. The conserved sequences in *T. nigroviridis* span the two large clusters of zinc fingers in the HIZ sequence, the first cluster of which is missing from our present clone, but may play an important evolutionarily conserved role yet to be identified.

It is known that that there exists a conserved synteny between this region of the human chromosome 19 and the mouse chromosome 17. As seen in the BLAT search, there are mouse sequence alignments to the entire putative HIZ sequence. There has also been a mouse homolog of cKr2 identified in a library produced by the National Cancer Institute, Cancer Genome Anatomy Project (NCI CGAP) which produces a large numbers of expressed sequence tags (ESTs) in its efforts to characterize genes. ESTs serve the same purpose as sequence-tagged sites in being standard markers for genome mapping, but ESTs serve an additional and important role in representing an expressed gene (Adams *et al.*, 1991). An EST is a sequence from a cDNA clone and thus denotes an mRNA. This murine EST (accession no. AW320799) was from an infiltrating ductal carcinoma in the mammary tissue of a 5 month old female virgin mouse and was specifically noted by the submitter to be a putative homolog of cKr2 based on sequence alignment. However, only 365 base pairs of this EST were sequenced and no other information is available regarding this homolog. In addition, there are 5 non-human ESTs that match the HIZ sequence with 80-85% similarity over their sequenced regions

(usually approximately 300bp): BB648938 (mouse embryo); BF780985 (mouse cDNA kidney); W99790 (mouse embryo); BG276745 (mouse germinal B-cell) and BF524226 (adult rat). All of the listed ESTs are short sequences that are uncharacterized. However, it gives us valuable evidence of the importance of HIZ because of its significant similarity across species.

3.2.3 HIZ mRNA Expression

Investigating gene expression often provides insight on the possible functions of a novel protein. In addition to detecting the abundance of its mRNA transcript in different tissues, northern blot analysis allows for determination of the mRNA transcript size as well as its possible splice variants.

To detect the expression of HIZ in various human tissues, two multiple-tissue northern (MTN) blots (Clontech) were probed with a $[^{32}P]$-radiolabelled fragment of HIZ ($HIZ_{798-919}$). The HIZ mRNA was predominantly detected at 4.2kb, with additional variants located at 1.4kb in skeletal muscle and an extra band at 3.5kb in both skeletal muscle and faintly in the pancreas. The 4.2kb mRNA was highly expressed in the heart, liver, skeletal muscle, pancreas and testis and expressed at a much lower level in the placenta, ovary, brain, lung and kidney. No bands were detected in spleen, thymus, prostate, small intestine, colon and peripheral blood leukocytes (see figure 14).

The putative HIZ sequence matched 13 ESTs which gave additional expression information. With the expanding collection of ESTs, the National Centre for Biotechnology Information (NCBI) has developed a program called 'AceView' (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly /index.html) which shows the
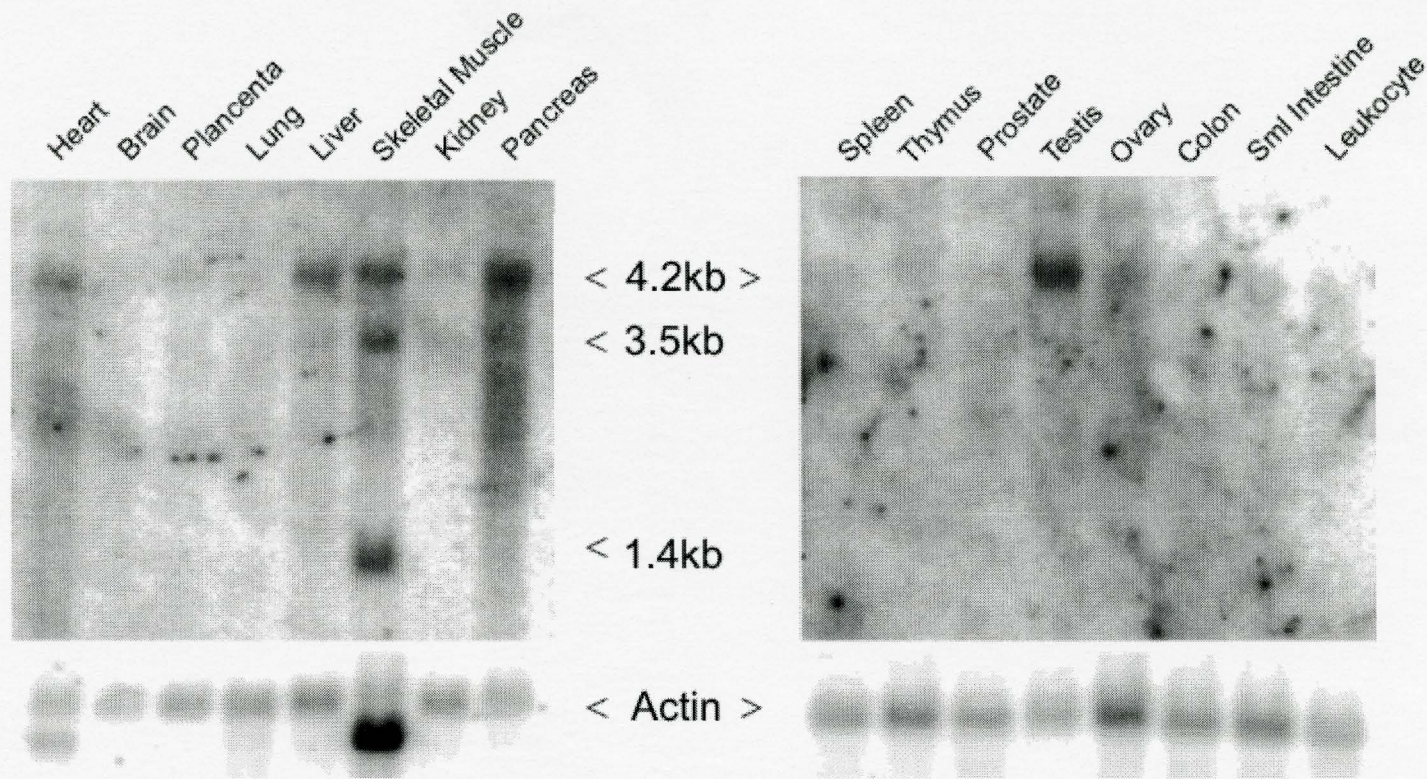
Figure 14. HIZ mRNA tissue distribution. Northern blots using a fragment of HIZ as a probe show that it is not ubiquitously expressed. A 4.2kb fragment was strongly expressed in the testis, liver, skeletal muscle and pancreas. A less pronounced band was found in the heart, placenta and lung. Strong bands at 3.5kb and 1.4kb were also found in the skeletal muscle. Membranes were stripped and reprobed with β-actin to ensure equal loads of mRNA per gel. Abbreviations: Sml Intestine = small intestine; Leukocyte = peripheral blood leukocytes

Table 6. A Summary of the ESTs Matching the HIZ sequence. A BLAST search of the NCBI EST database (dbEST) was performed to find ESTs matching the HIZ putative nucleotide sequence. The following is a list of the accension numbers of the ESTs, where they match to HIZ, and the tissue they were expressed in. Note: the sequences of each EST are not representative of their full sequence, but only the small fragment that was sequenced by the supplier. The numbers in bold represent the end of the HIZ protein.

EST #	Corresponding nt region on HIZ	Organ	Tissue Type
BF724583	1324-1717	Eye	Iris
BI761477	1638-2270	Pooled colon, kidney, stomach	
BG163321.1	1656-1984	Kidney	Hypernephroma
BG681196.1	2161-3063	Skin	squamous cell carcinoma
BE891963.1	2197-2577	Skin	melanotic melanoma
BE207234.1	2329-2743	Lung	small cell carcinoma
BE266530.1	2348-2912	Lung	small cell carcinoma
BG913938.1	2420-2597	Brain	anaplastic oligodendroglioma with 1p/19q loss
AA062959.1	2796-**3168**	Heart	Fetal
AI912315.1	2814-**3168**	Uterus	Well-differentiated endometrial carcinoma
AW615682.1	2993-**3168**	Lung	small cell carcinoma
AA827434.1	3028-**3168**	Kidney	tumour
D60780.1/ D60343.1	3080-**3168**	Brain	Fetal
AI597633.1	3081-**3168**	Brain	anaplastic oligodendroglioma
AA035617.1	3139-**3168**	Heart	Fetal
BG290052	*Splice variant*: 1739-1824 or 569-654 and 1987-2246	Bladder	transitional cell papilloma

alignment of mRNAs and ESTs to the human genome to predict genes (see table 6). Human ESTs corresponding to the HIZ sequence were found in libraries constructed from the eye, kidney, colon, skin, lung, brain, heart, uterus and bladder.

The HIZ gene is labelled by AceView as chr_177_664.a and chr_177_664.b (see figure 13). This program deduced that there were two splice variants because of EST no. BG290052. This EST is missing a region on the genomic sequence. However, another EST (no. EG290052) contains the sequence that is spliced out of BG290052. Therefore, both spliced and non-spliced variants are likely expressed.

A possible role for HIZ in myogenesis was suggested when a high level of expression in skeletal muscle was seen on the MTN blot. However, a northern blot using RNA isolated from differentiating C2C12 (mouse myoblast) cells showed that there were no differences in HIZ expression during myoblast differentiation (see figure 15).

3.2.4 HIZ Contains a Potent Transactivation Domain Within Residues 876-996

Recent studies of the HCF-1 basic domain binding partners GABP and Miz-1 have shown that HCF-1 interacts with regions critical for transactivation to regulate their activation potential (Vogel and Kristie, 2000a, David Piluso, personal communication).

By tethering the fragment of HIZ isolated from the yeast two-hybrid assay to GAL4DBD, we discovered that HIZ contains a strong activation domain. Various truncations were constructed to find the minimal activation domain (see figure 16). The smallest region that retained the high transactivation potential was $HIZ_{876-996}$ in both BHK-21 and HeLa cell lines (see figures 17 and 18). In BHK-21 cells, this region was

62

Days Post Confluence   0 1 1 2 2 3 3

HIZ

myogenin
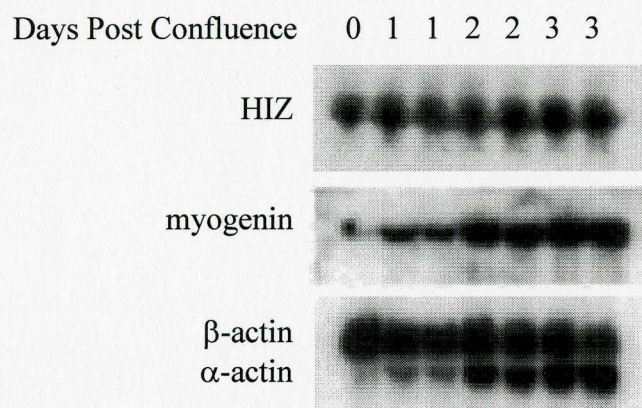
β-actin
α-actin

Figure 15. HIZ Expression is Not Altered During Myogenic Differentiation.
HIZ expression did not increase during myoblast differentiation in comparison to
the myogenic differentiation markers myogenin and α-actin (myogenin and actin
panels provided by Mark van Delft).

(Amino Acids Indicated)
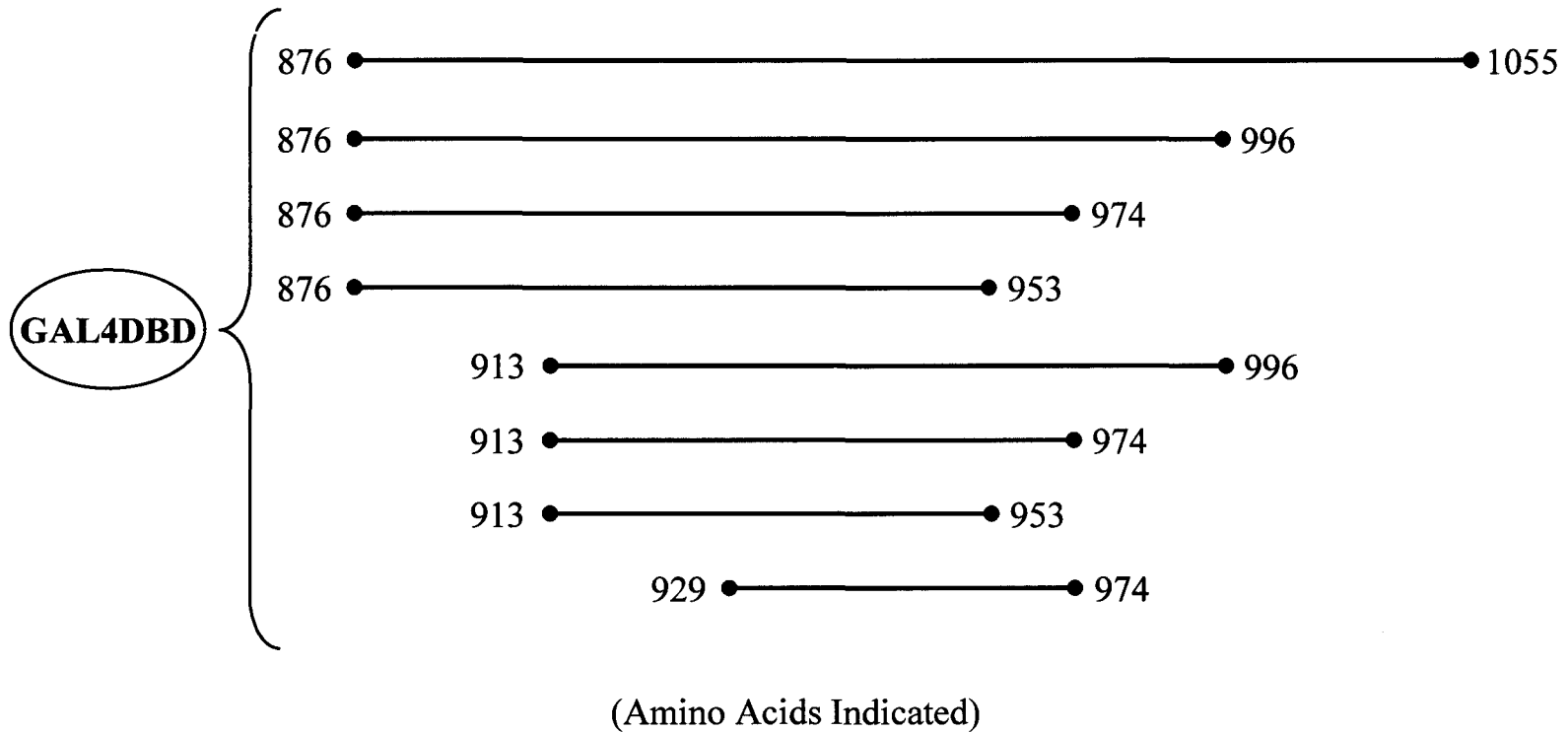
Figure 16. Schematic Representation of HIZ Fragments Constructed and Tested for Transactivation Potential. All constructs were tethered to the Gal4 DNA binding domain by cloning into the pSG424 vector and used in transient transfection assays in COS-1, BHK-21 and HeLa cell lines to test for their ability to activate a luciferase reporter gene.
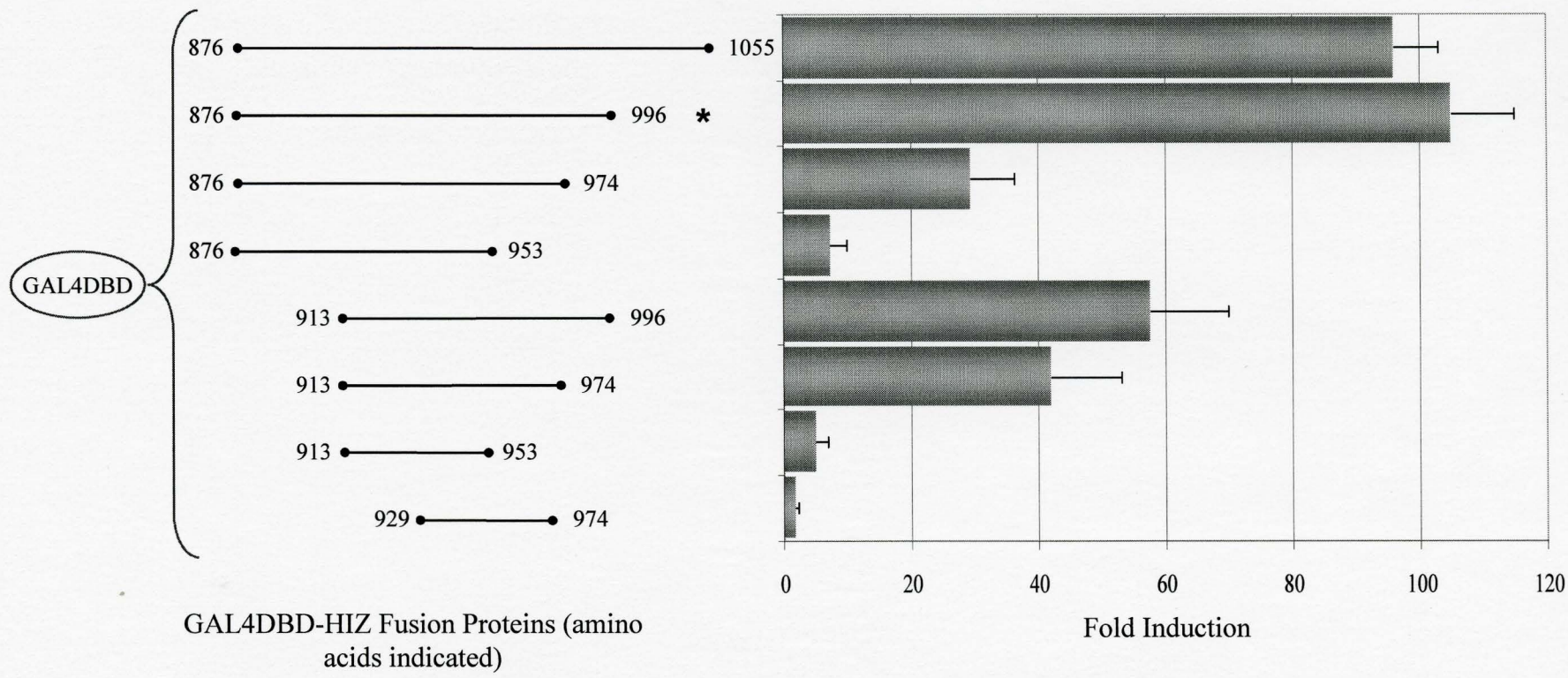
Figure 17. HIZ Contains an Activation Domain Within Residues 876-996. Various GAL4DBD-HIZ fusion proteins were expressed in BHK-21 cells and their transactivation potentials were assayed as relative light units normalized to protein (RLU/ug protein) compared to the background GAL4DBD vector alone (fold induction). The GAL4DBD was fused to the amino-termini of all derivatives. Results are an average of three independent experiments performed in duplicate ±S.D. The * indicates the region with the maximum activation potential.

Figure 18. The HIZ Activation Domain is Active in HeLa Cells. Various GAL4DBD-HIZ fusion proteins were expressed in HeLa cells and their transactivation potentials were assayed as relative light units normalized to protein (RLU/ug protein) compared to the background GAL4DBD vector alone (fold induction). The GAL4DBD was fused to the amino-termini of all derivatives. Results are an average of three independent experiments performed in duplicate ±S.D. The * indicates the region with the maximum activation potential.

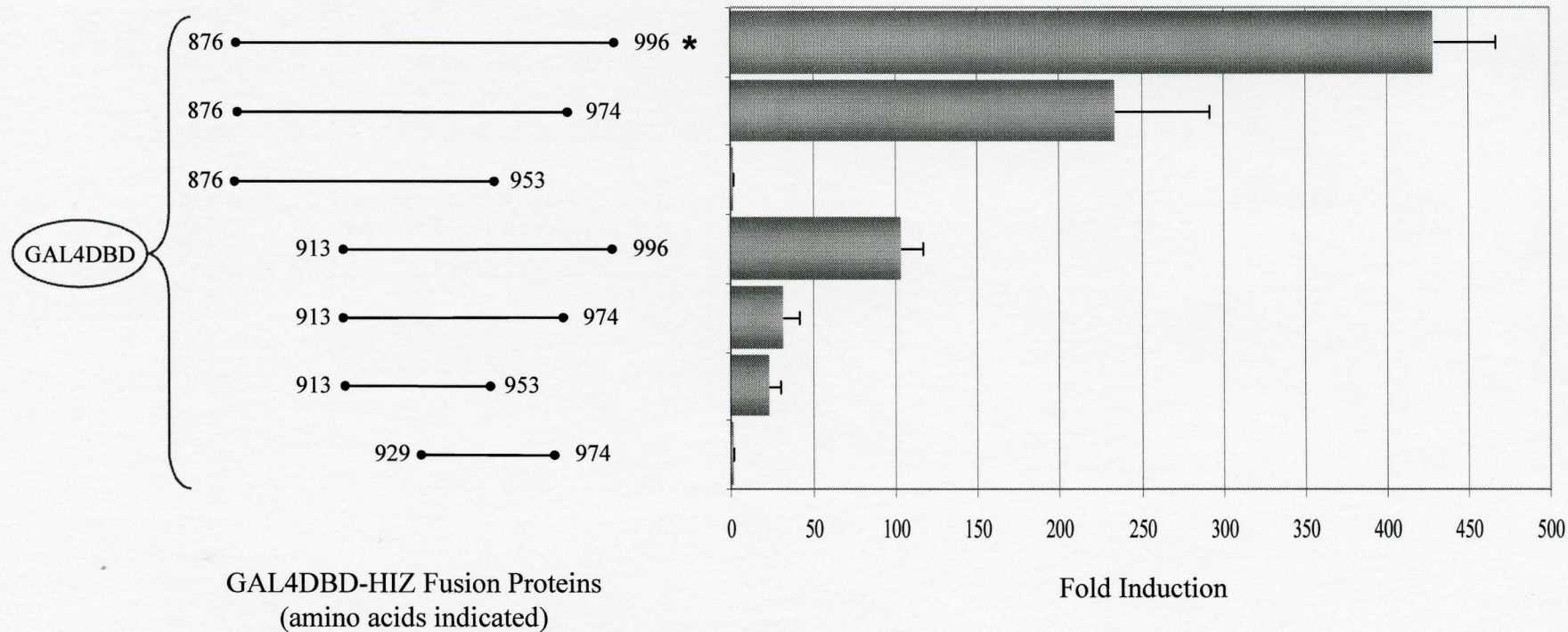able to activate the reporter gene approximately 100-fold over the GAL4DBD expression vector alone, and in HeLa cells, approximately 400-fold. Some activity was retained in smaller regions but the results varied between cell lines. Interestingly, a deletion at the amino-terminal from 876 to 913 resulted in half the maximum activation in BHK-21 cells but reduced the activation potential to approximately a quarter in HeLa cells. Likewise, deletions at the carboxy-terminal from 996 to 974 result in a decrease to a quarter of the full potential in BHK-21 cells but only decreased the full potential by half in HeLa cells. Deleting both ends ($HIZ_{836-897}$) did not decrease the activation significantly compared to deleting the carboxy-terminal in BHK-21 cells, but drastically reduced the activation potential in HeLa cells to approximately 1% of the full potential.

The HIZ activation domain has a very high transactivation potential – approximately 1/7 activity of the potent VP16 activation domain in both HeLa and BHK-21 cell lines (see figure 19).

We chose to continue our experiments in HeLa cells because HIZ expression in this cell line led to higher levels of transcription (approximately four-fold) than in BHK-21 cells. Also, BHK-21 is a hamster cell line and may therefore be lacking inherent components present in human cells.

3.2.5 HIZ Binds Specifically to $HCF_{750-902}$

The same series of HIZ regions tested for transactivation (see figure 16) were translated *in vitro* to test its binding to $HCF_{750-902}$. This GST *in vitro* binding assay showed that $HIZ_{876-974}$, $HIZ_{876-996}$, $HIZ_{913-996}$ showed the highest levels of binding. Quantitative data
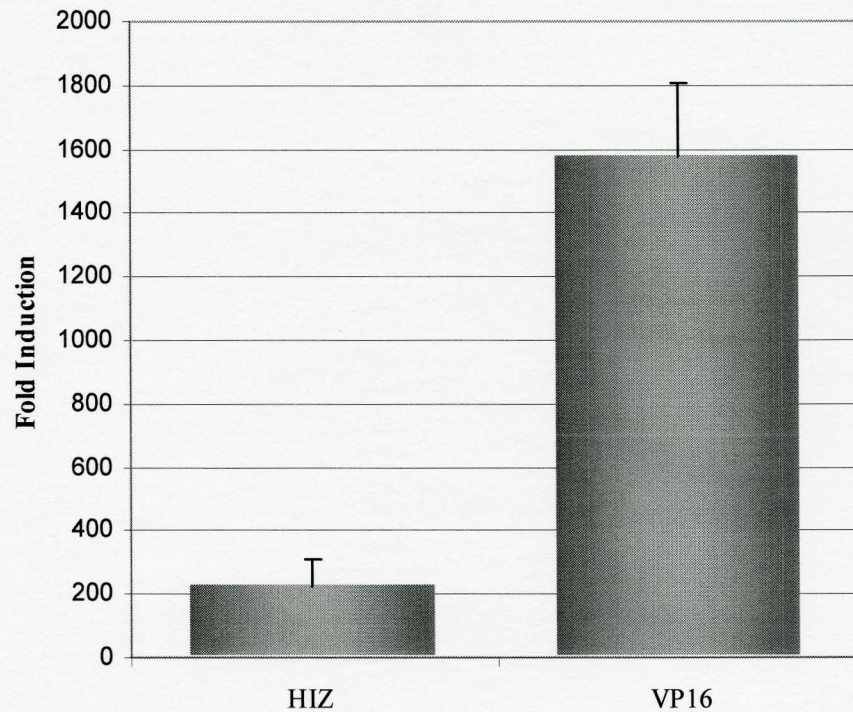
Figure 19. HIZ Activation Potential is 1/7 of the Potent VP16 Activation Domain. HeLa cells were transfected with the activation domains of HIZ (residues 876-996) and VP16 (residues 410-490). The HIZ activation domain was able to induce transcription of the reporter gene over 200 times empty vector alone. This level of activation was demonstrated to be comparable to VP16. The same result was found in BHK-21 cells. These results reflect the average of two independent experiments done in duplicate ±S.D.

obtained using phosphorimaging indicated that the highest level of binding was seen by HIZ$_{876-996}$. Phosphorimaging showed that HIZ$_{876-953}$ also bound with the same affinity but because the 10% load of the translation did not appear as a single clear band, quantification was difficult and may not be as accurate.

There appears to be a correspondence between the critical regions for binding and those for activation (see figure 21). HIZ$_{876-974}$, HIZ$_{876-996}$, HIZ$_{913-996}$ retained transactivation activity and also retained its binding potential to HCF-1. These three were able to bind HCF-1 at an efficiency of 7-14% of the load. These regions were able to bind HCF-1 specifically and not GST-alone.

The smaller fragments showed binding efficiencies in the range of 0-4% (see figures 20 & 21). Thus, HCF-1 interacts with HIZ at the residues responsible for activation (see figure 18).

3.2.6 Full Length HCF-1 and HCF-1$_{750-902}$ Repress HIZ Transactivation Potential

Recent studies have shown that HCF-1 can function as a regulator of transactivation potential. HCF-1 binding has been implicated in regulating the transactivation potential of LZIP, Zhangfei, GABP and Miz-1. Thus, we were interested in the effect of HCF-1 on the HIZ transactivation domain.

Transient transfection assays were done in HeLa cells. Full length HCF-1 and the minimal region of HCF-1 required for maximum binding (HCF$_{750-902}$) were co-transfected in separate experiments at increasing levels (0.5 to 1.5µg) with 0.05µg of the HIZ activation domain (residues 876-996) tethered to a GAL4DBD. This resulted in a decrease of the HIZ activation potential to approximately 50% when 10-15 times HCF-1
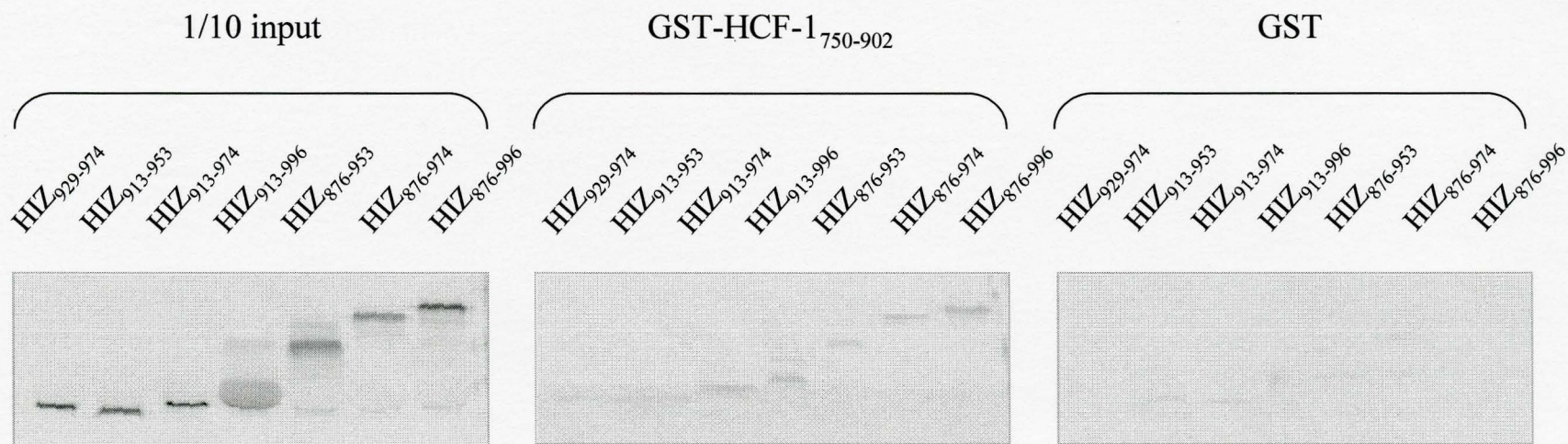
Figure 20. HIZ$_{913-974}$ Interacts with HCF-1$_{750-902}$ Specifically. A series of HIZ deletions were tested for their ability to bind HCF-1$_{750-902}$. HIZ$_{876-996}$, HIZ$_{876-974}$ and HIZ$_{913-996}$ were able to bind HCF-1 significantly (7-14% of total load). Smaller derivatives bound with lower affinity or not at all (0-4%). Binding affinity for each derivative was determined quantitatively using phosphorimaging (Storm, Molecular Dynamics) and shown in Figure 21.

	% HCF-1 binding		Activation Potential (HeLa)
* 876 •————————————• 996	12.8	10.2	++++
876 •————————————• 974	6.6	8.5	++
876 •————————• 953	4.4	3.2	-
913 •————————————• 996	13.7	9.2	++
913 •————————• 974	0.2	2.7	+
913 •——————• 953	2.2	1.0	+
929 •————• 974	2.0	0.7	-

Figure 21. Quantitative Results of HCF-1-HIZ GST-binding Assay Using Phosphorimaging. Critical residues for transactivation potential correspond to those required for HCF-1 binding as determined by phosphorimaging. GST-binding assay results represent two independent experiments. The activation data reflect the results presented in figure 18 ('++++' indicates maximum activation potential; '-' indicates very minimal or no activation potential).
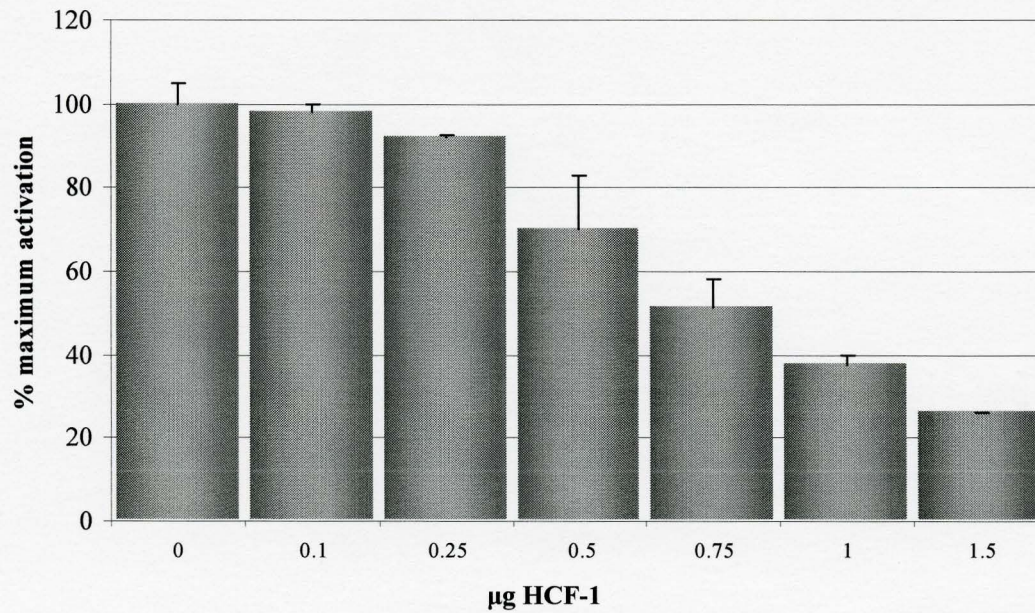
Figure 22. Full Length HCF-1 Represses HIZ Activation Potential. HeLa cells were co-transfected with 0.05µg of GAL4-HIZ$_{876-996}$ and increasing amounts of full length HCF-1 (0.01µg to 1.5µg). The HIZ activation potential alone was set to 100%. HCF-1 is shown to repress HIZ activation to approximately 25%. These results reflect the average of two independent experiments performed in duplicate ±S.D. Amount of plasmid was kept constant using the empty expression plasmid for HCF-1 (pCDNA3.1).
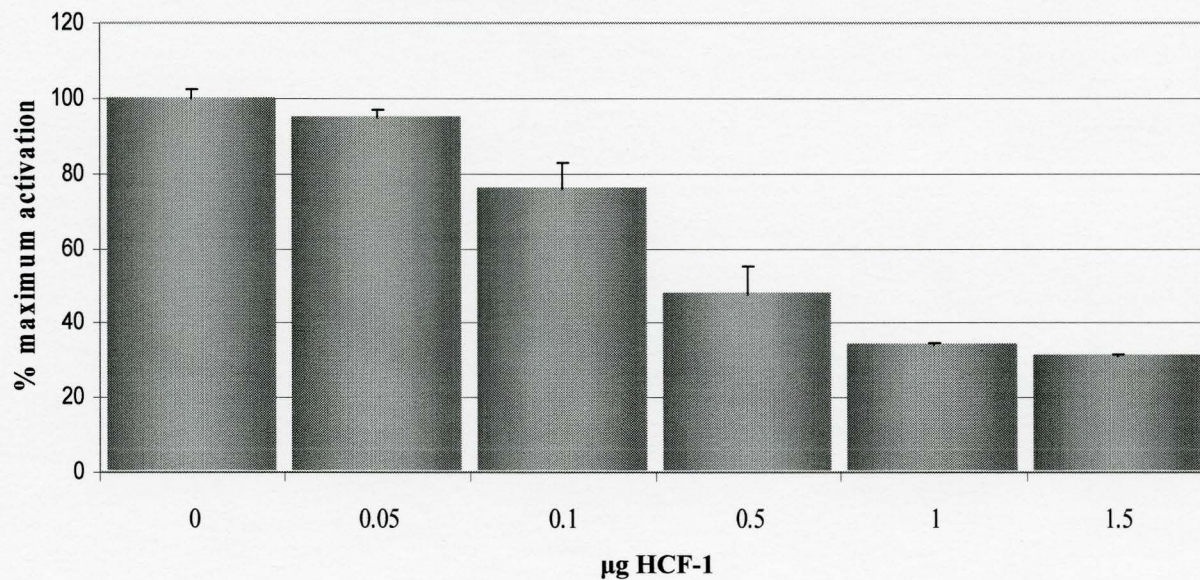
Figure 23. HCF-1$_{750-902}$ is also Capable of Repressing HIZ Transactivation Potential. HeLa cells were co-transfected with 0.05μg of GAL4-HIZ$_{876-996}$ and increasing amounts of HCF-1$_{750-902}$ (0.01μg to 1.5μg). The HIZ activation potential alone was set to 100%. HCF-1$_{750-902}$ is shown to repress HIZ activation to approximately 30%. These results reflect the average of two independent experiments performed in duplicate ±S.D. Amount of plasmid was kept constant using the empty expression plasmid for HCF-1$_{750-902}$ (pCGN).

plasmid was added (0.5-0.75μg of HCF-1 vs. 0.05μg HIZ). HIZ transactivation was further repressed to approximately 25% of its maximum activation potential when 1.5μg of HCF-1 was co-transfected (30 times the amount of HIZ plasmid transfected) (see figures 22 & 23). To test if the HCF-1 repression was specific for HIZ, we repeated the experiment using GAL4-VP16 instead of GAL4-HIZ (see figure 24).

Interestingly, HCF-1 did not repress the VP16 activation domain as it did HIZ, but instead acted to increase the activation potential of VP16, in the same manner seen with co-transfection with other HCF-1 binding partners such as PGC-1β. The VP16 activation potential doubled when 0.5μg of plasmid with HCF-1 was co-transfected (10 times the amount of VP16 plasmid transfected) and increased to approximately 175% of the level with GAL4-VP16 alone when 1.5μg of HCF-1 plasmid was co-transfected.

3.2.7 HIZ may have a Role in Suppressing Cell Proliferation

Zinc fingers are believed to play critical roles in transcriptional regulation, leading us to hybridize a HIZ probe (HIZ$_{798-978}$) to a matched tumour/normal expression array prepared from 68 human tumours paired with normal tissue from the same individual. The results showed that HIZ is expressed at an equal level in normal and tumour tissues in kidney, breast, prostate, ovary, cervix, colon, stomach, rectum and small intestine. However, HIZ is expressed at a significantly higher level in normal tissues of the uterine and lung compared to their matched tumour samples (see figure 25). This finding suggests a role for HIZ in suppressing cell proliferation, which can possibly be negatively regulated by the cell cycle promoter HCF-1 as our results above demonstrate that HCF-1 represses the transactivation potential of HIZ.
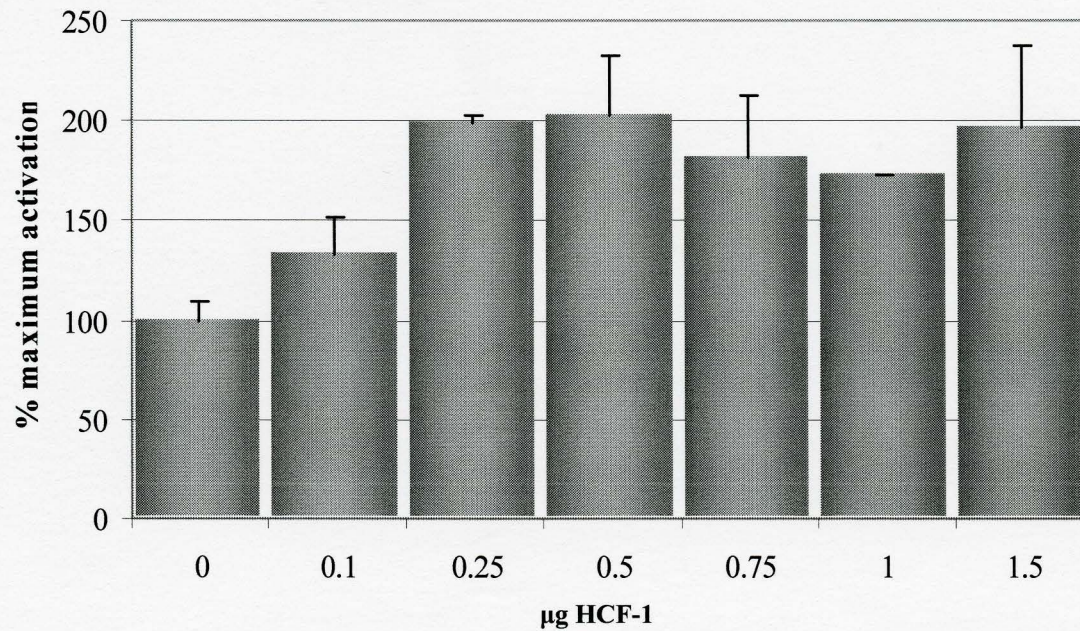
Figure 24. Full Length HCF-1 Increases VP16 Activation Potential. HeLa cells were co-transfected with 0.05µg of GAL4-VP16$_{AD}$ and increasing amounts of full length HCF-1 (0.01µg to 1.5µg). The VP16 activation potential alone was set to 100%. HCF-1 is shown to increase the activation potential of VP16. These results reflect the average of two independent experiments done in duplicate ±S.D. Amount of plasmid was kept constant using the empty expression plasmid for HCF-1 (pCDNA).
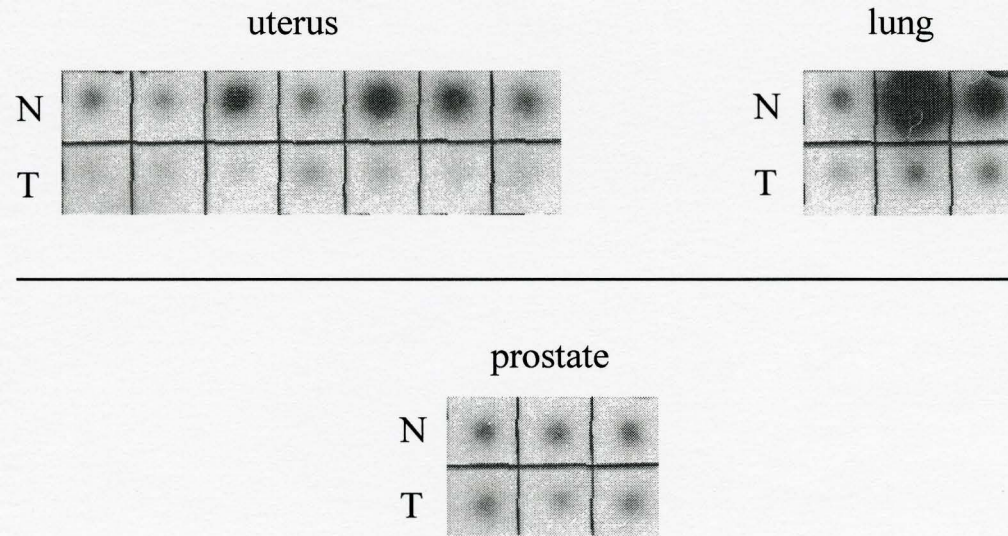
Figure 25. HIZ Expression is Lower in Uterine and Lung Tumour Tissues in Comparison to Matched Normal Tissues. Hybridization of radiolabelled HIZ to a matched tumour/normal tissue array that contained matched samples from 68 different individuals with various tumours showed that HIZ was expressed at a lower level in all 7 uterine samples and all 3 lung samples on the array. HIZ was found to be expressed at equal levels in matched samples from kidney, breast, ovary, cervix, colon, stomach, rectum, small intestine and prostate (as shown above).

3.2.8 Obtaining the Full Length Clone of HIZ

The HIZ sequence found in the HeLa library screened for interaction with HCF-1 was only 244 amino acids-long. Using a BLAST search of the EST database (http://www.ncbi.nlm.nih.gov/BLAST/), we ordered several ESTs that corresponded to the HIZ sequence that we had. The longest clone found (EST accession no. AI912315) codes from residue 445 of the putative HIZ sequence to the stop codon. At the 5' end, it reaches the second cluster of zinc fingers (ZF 10-17).

We attempted to clone the full length HIZ using PCR techniques. As template, we used several cDNA libraries. First, we tried HepG2 (liver) and HeLa (cervical) cDNA libraries that were reverse transcribed using the SMART RACE cDNA amplification kit (Clontech). We then purchased testis Marathon-Ready cDNA (Clontech), which are "libraries" of dsDNA with partially double-stranded adaptors ligated to each end. This cDNA has been cited in many recent papers for 5' RACE. However, from all of the above libraries we were not able to amplify a full length clone. The cDNA was present in all three libraries though, because we were able to amplify an internal fragment at the 3' end without difficulty.

Aware of the high GC-content at the 5'-end, we tried adding DMSO as well as DMSO substitutes: Q-solution (Qiagen) and GC-melt (Clontech) to PCR reactions with increasing annealing temperatures to overcome possible secondary structures in the cDNA. In addition to performing PCR reactions with various annealing temperatures, magnesium and different primers were used at different concentrations. Times and temperatures of the annealing and elongation periods were altered, but no correct PCR products were amplified. Different polymerases in different combinations have also been

used including Clontech's polymerase mix (AdvanTaq with TaqStart Antibody and proofreading polymerase for long distance and GC-rich PCR amplification), HotStarTaq (Qiagen) and Proofstart (Qiagen), as well as the standard Taq, Vent, Deep Vent and Pfu polymerases. Reamplification of PCR reactions with nested primers was also done in the event that the mRNA is in low abundance. However, the full length clone of HIZ was not isolated with any of the above techniques.

3.3 HCF-1 Contains an Autonomous Transactivation Domain

Although HCF-1 is shown to repress HIZ activation in a manner similar to Miz-1, HCF-1 has been found to play a role in synergistic transactivation of LZIP, Zhangfei, GABP and PGC-1β. Co-activators often contain an activation domain as demonstrated by the well-studied coactivator VP16. Thus, we were interested in whether the coactivator role of HCF-1 was also attributed by the presence of an activation domain. By constructing a series of GAL4DBD fusions to various regions of HCF, a novel acidic transactivation domain was found in HCF-1.

3.3.1 Yeast Transactivation Assay

A total of ten fusion proteins were constructed using various fragments of the HCF acidic region cloned into the yeast GAL4DBD expression plasmid pGBT9 (see figure 26). These clones were transformed into yeast strain Y190 and tested for their ability to activate the GAL4-responsive reporter gene, β-galactosidase.

GAL4DBD   <0.1

GAL4DBD +

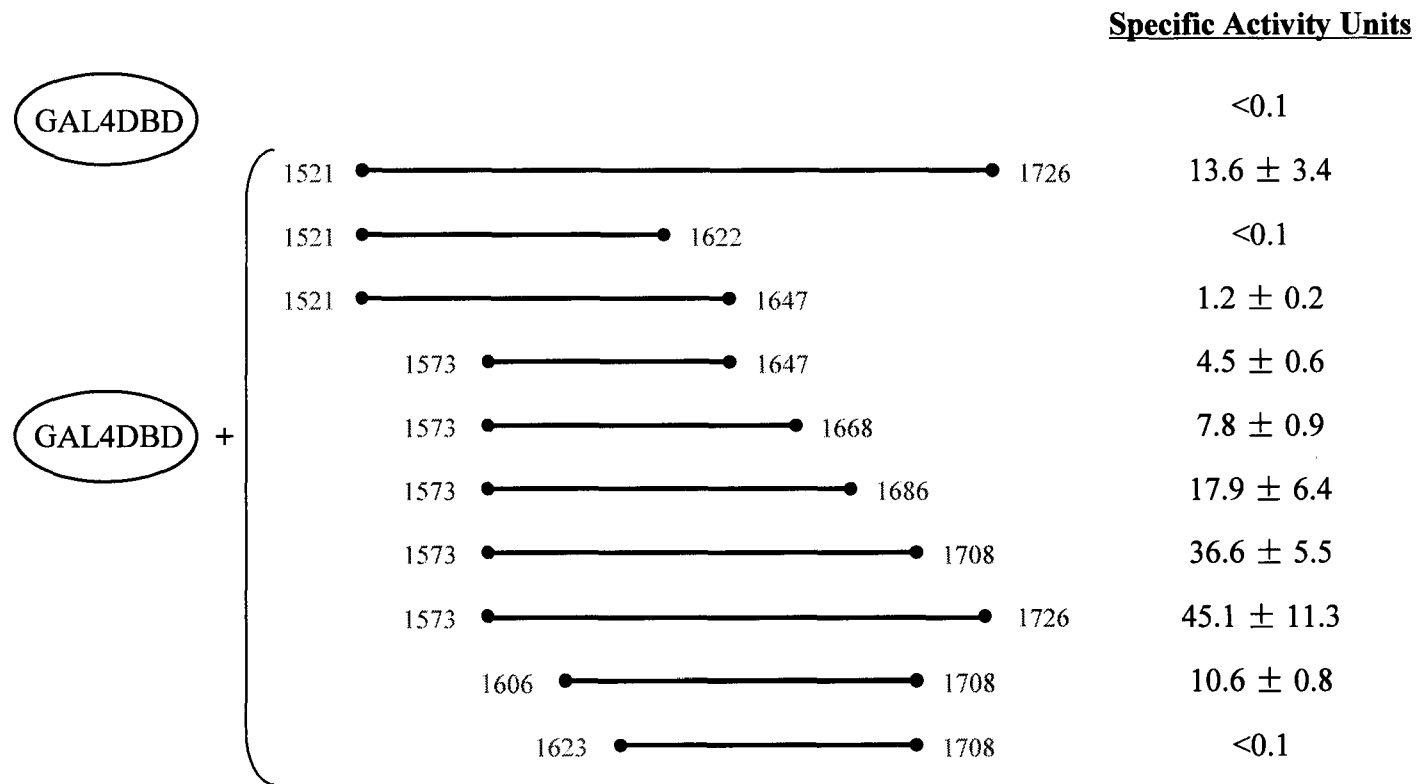|		Specific Activity Units
1521 •————————————————• 1726		13.6 ± 3.4
1521 •——————————• 1622		<0.1
1521 •——————————• 1647		1.2 ± 0.2
1573 •——————• 1647		4.5 ± 0.6
1573 •——————————• 1668		7.8 ± 0.9
1573 •——————————• 1686		17.9 ± 6.4
1573 •——————————————• 1708		36.6 ± 5.5
1573 •————————————————• 1726		45.1 ± 11.3
1606 •——————————• 1708		10.6 ± 0.8
1623 •——————————• 1708		<0.1

Figure 26. **Schematic Representation of HCF-1 Fragments Constructed and Tested for Autonomous Transcription Activation in yeast.** All constructs were tethered to the Gal4 DNA binding domain and its ability to activate transcription of the *LacZ* reporter gene was measured using a quantitative *LacZ* assay.

The largest region tested was HCF-1$_{1521-1726}$ which had a specific activity of 13.6±3.4 when compared to the control GAL4DBD alone. However, the highest activity was localized to HCF$_{1573-1726}$ suggesting a possible repressive domain between residues 1521-1572. HCF$_{1573-1726}$ gave a specific activity of 45.1±11.3 and progressive deletions from both carboxy- and amino- terminals corresponded to decreasing specific activity (see figure 26).

3.3.2 Mammalian Transactivation Assay

The ten GAL4DBD-HCF fusion proteins tested in yeast were subcloned into the mammalian expression vector pSG424 and then co-transfected into COS-1 cells with a plasmid containing the luciferase reporter gene downstream from a GAL4 responsive promoter (pLUC) (see figure 27). These transfections were repeated in BHK-21 (see figure 28) and HeLa cell lines (see figure 29) and showed similar trends, suggesting that this autonomous transactivation potential is an intrinsic property of the region and not cell-type specific.

This assay reflected most of the trends seen in yeast except that in mammalian cells, HCF$_{1523-1708}$ showed some activity while HCF$_{1521-1647}$ was inactive. Again, the longest clone HCF$_{1521-1726}$ showed a lower activation potential than HCF$_{1573-1726}$ (20-fold compared to 38-fold in COS-1 cells). To detect the protein expression of the fusion proteins, a Western blot was performed using cell lysates from COS-1 cells. This assay showed that although the level of expression varied between each HCF fragment, higher levels of expression did not correlate to activation. The fragments showing the highest levels of expression activated poorly (see figure 30).
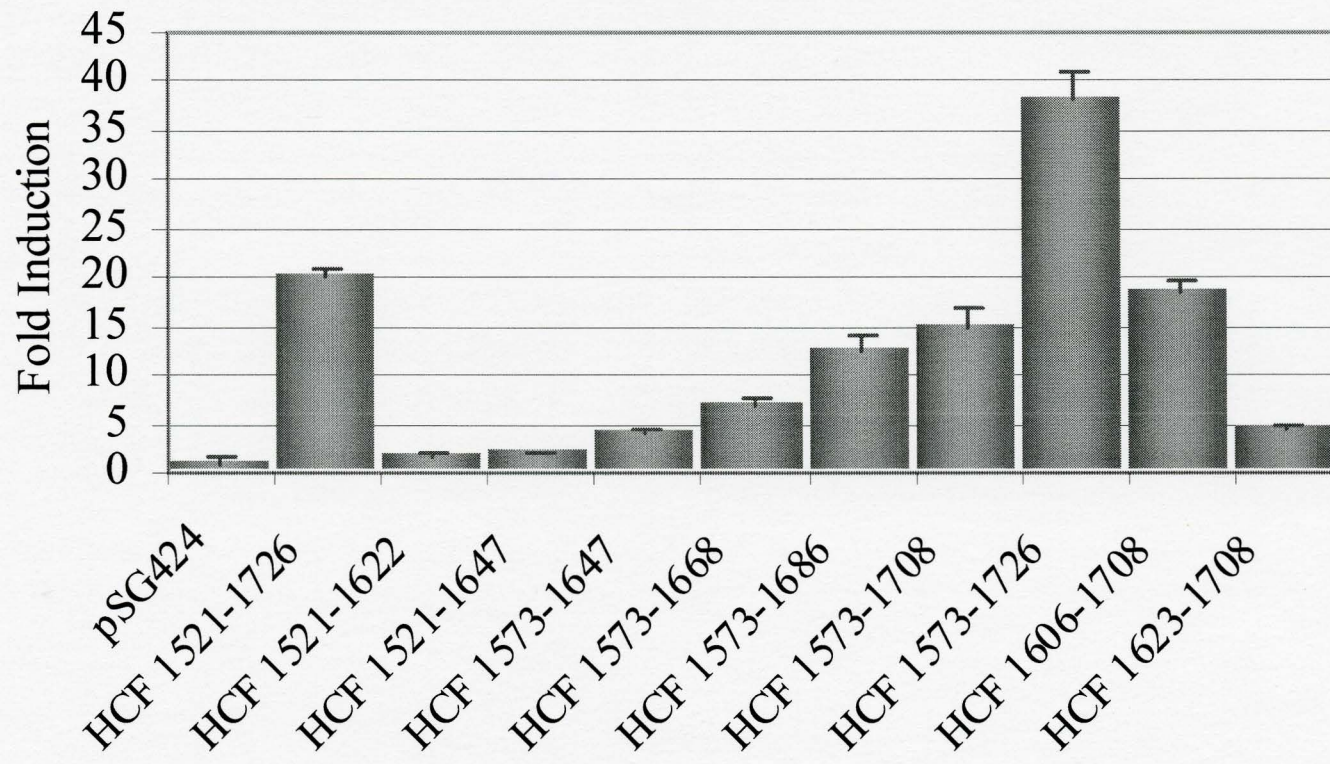
Figure 27. Transactivation Potential of HCF-1 Constructs in Mammalian COS-1 cells. The construct with the highest activation potential was HCF 1573-1726 showing approximately 37-fold induction of the luciferase reporter gene when compared to the control plasmid, pSG424, expressing GAL4DBD alone.
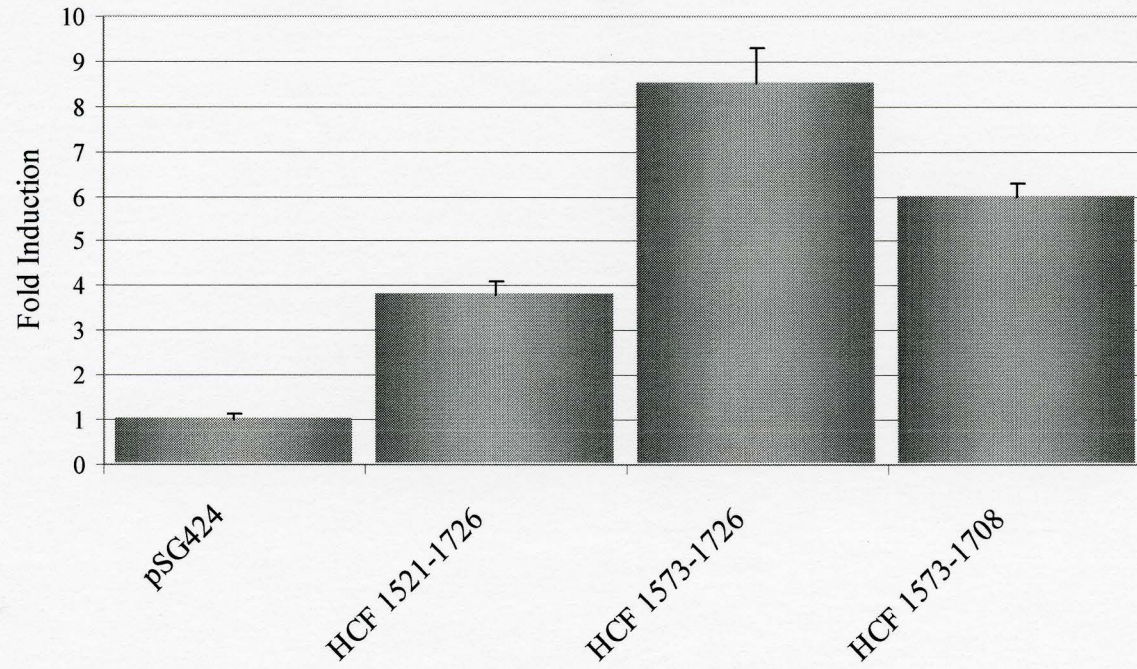
Figure 28. The HCF-1 Activation Domain HCF$_{1573-1726}$ is Active in BHK-21 Cells. Transfections in BHK-21 cells showed that HCF$_{1573-1726}$ contained the highest transactivation potential, similar to results in COS-1 cells.
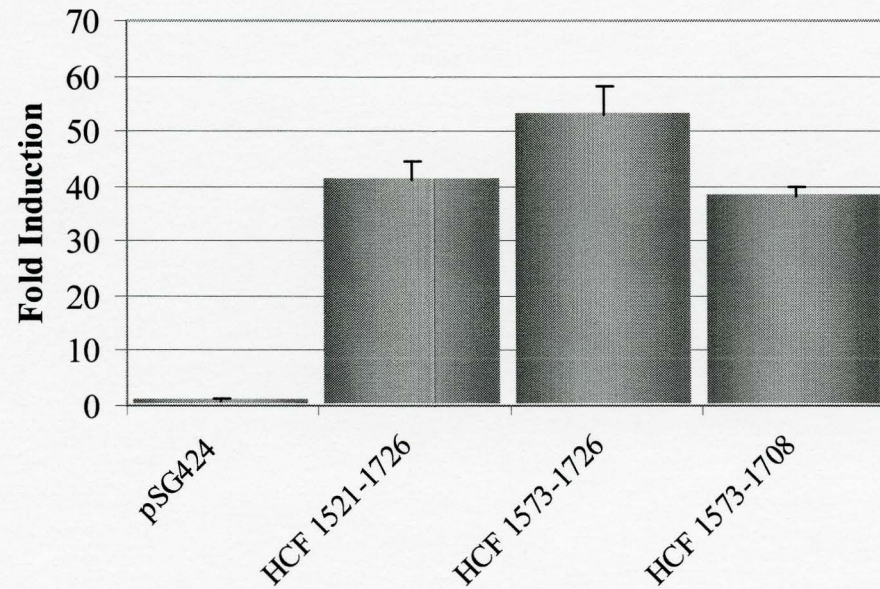
Figure 29. The HCF-1 Activation Domain HCF$_{1573-1726}$ is Active in HeLa Cells. These results also showed that HCF 1573-1726 possessed the highest transactivation potential suggesting that the activity is intrinsic to the region and not dependent on cell type. However, expression of HCF-1 led to higher levels of transcription initiation of reporter gene in comparison to COS-1 and BHK-21 cell lines.
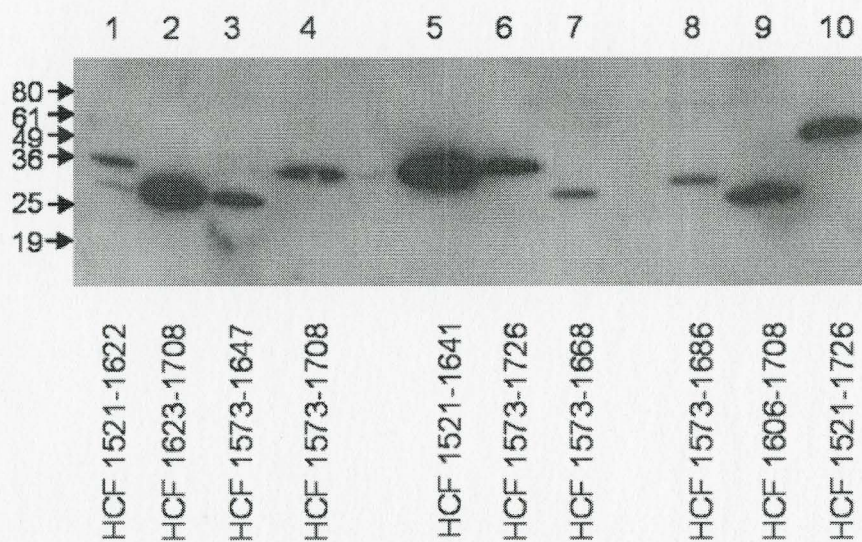
Figure 30. Western Blot Showing the Expression of HCF-1 Constructs in Cos-1 Cell Lysates. These data indicate that the various levels of activation potential are not caused by differences in protein expression. 25µg of protein was loaded per lane. The constructs were expressed at different levels but these differences did not correlate with their transactivation activity.

CHAPTER FOUR

DISCUSSION

4.1 Major Findings of the Studies

Recent evidence has shown that HCF-1 associates with numerous DNA-binding transcription factors and is capable of mediating their transactivation potential (Luciano and Wilson, 2000; Lu and Misra, 2000; Vogel and Kristie, 2000a; Lin *et al.*, 2002; David Piluso, personal communication).

Impairment of HCF-1 binding was shown to significantly reduce the transactivation potential of LZIP and Zhangfei, (Luciano and Wilson, 2000; Lu and Misra, 2000). Similarly, HCF-1 was demonstrated to increase the activation potential of the transcriptional coactivator PGC-1β, which also interacts with HCF-1$_{VID}$ using the same HBM as LZIP, Zhangfei and VP16 (Lin *et al.*, 2002). The ETS domain transcription factor GABP has been shown to bind at amino acids 813-847 of the HCF-1 basic domain (Vogel and Kristie, 2000a). Critical HCF-1-binding residues correlated directly to GABP activation, supporting the evidence from LZIP, Zhangfei and PGC-1β studies that suggest a co-activator role for HCF-1. However, in addition to the HCF-1-binding region, other regions of HCF-1 are likely required to enhance the activation potential of GABP. Co-transfection with increasing amounts of the fragment of HCF-1 required for GABP interaction (presumably acting as a competitive inhibitor for endogenous wild-type HCF-1) repressed the transactivation potential of GABP (Vogel and Kristie, 2000a).

Interestingly, although Miz-1 also interacts with the basic domain of HCF-1 (residues 750-902) using the region critical for its transactivation potential, HCF-1 acts as a co-repressor of Miz-1 transactivation (David Piluso, personal communication) and not a coactivator as previously seen.

HCF-1 is likely involved in promoting the cell cycle since an HCF-1 P134S (proline-to-serine change at residue 134) mutant in the tsBN67 cell line prevents cells from passing the G_o/G_1 boundary (Goto *et al.*, 1997). Therefore, it is very interesting that GABP, identified as a crucial transactivator in the life cycle of HSV (Jones and Tjian, 1985) can be rendered transcriptionally active with HCF-1 binding, while Miz-1, an activator of the cell cycle inhibitor $p15^{INK4b}$ is repressed. These results imply that HCF-1 can regulate pathways that both activate and inhibit the cell cycle.

Although HCF-1 is known to have a role in cell cycle control, the events underlying the G_o/G_1 cell cycle arrest observed are unknown. The detection of specific HCF-1 interacting proteins can provide insight into the signals that affect HCF-1 function and effector proteins downstream that HCF-1 regulates. The basic domain has been implicated in cell cycle control because it was required to rescue the cell proliferation defect in tsBN67 cells (Wilson *et al.*, 1997) and has therefore been the main focus of this research. The data in this thesis present the characterization of a novel human protein that interacts with the HCF-1 basic domain at residues 750-902. We have named this protein HCF-1 interacting zinc finger protein (HIZ). HIZ displays tissue-specific expression with high levels in the heart, skeletal muscle, liver and testis. Similar to other HCF-1 binding proteins: LZIP, GABP and Miz-1, we have found a functional autonomous activation domain that correlates to its HCF-1 interaction domain. This

activation domain is located at the carboxy-terminus. Co-transfections with large amounts of full-length HCF-1 as well as $HCF_{750-902}$ repressed HIZ activation potential, similarly to Miz-1.

Molecular dissections of coactivators have often shown that an activation domain is present and required for synergistic coactivation (Merika *et al.*, 1998; Luo *et al.*, 1998; Belaguli *et al.*, 2000). Thus, to investigate the newfound coactivating role of HCF-1, we constructed GAL4DBD fusions of HCF-1 regions and discovered the presence of an autonomous activation domain within the carboxy-terminal acidic domain.

4.2 HIZ Protein Structure

A partial HIZ clone was isolated from the two-hybrid screen. A BLAST search of this sequence against the human genome sequences showed that HIZ is encoded at 19q13.42, a region closely related to murine chromosome 17 (Stubbs *et al.*, 1996). The long arm of chromosome 19 has been identified as a region enriched in zinc finger genes (Lichter *et al.*, 1992). Our partial HIZ clone contains one cluster of 7 zinc fingers and by tracing the upstream sequence in the genomic sequence, we have identified an additional 9 zinc fingers including another cluster of 7 zinc fingers. Between the two clusters are 2 unlinked zinc fingers that are separated from the two clusters by regions enriched in proline and alanine. Interestingly, within one of the independent zinc fingers (ZF 9) is an LxxLL motif (see figure 11).

An LxxLL motif has been shown to be both necessary and sufficient for recognition and binding of certain coactivators to liganded nuclear hormone receptors (Heery *et al.*, 1997). For LZIP, its LxxLL motifs were not required for HCF-1 binding,

but were critical for transactivation (Luciano and Wilson, 2000). Therefore, although HIZ binds HCF-1 at a separate region from LZIP, they may synergistically recruit additional cofactors, to form a larger complex of transcription factors surrounding HCF-1. It is very interesting to note that in the chicken homolog, cKr2, there is also an LxxLL motif within the ninth ZF, similar to HIZ. In cKr2, this ZF is linked to ZF 8 and not independent as in HIZ, but this pair is separated from the two clusters of linked fingers. As well, studying the Miz-1 protein sequence reveals a similar position of an LxxLL motif within its third zinc finger. Zinc finger structures have been implicated in protein binding for several proteins (Mackay and Crossley, 1998). In the transcription factor Fog (friend of GATA), the ZF that serves as a protein interaction motif to interact with the N-terminal ZF of GATA-1 is separate from the array of sequential DNA-binding ZF structures (Mackay and Crossley, 1998). This arrangement is similar to the ZFs in HIZ. The ZF containing the LxxLL motif is separate from the two clusters of linked ZFs. Thus, future studies on whether the LxxLL motifs within the loop of zinc finger 9 bind nuclear hormone receptors may reveal a new mechanism of protein-binding which uses the structure of zinc fingers to expose the protein binding site. With respect to HIZ, this LxxLL motif may play a role in increased activation.

The primary sequence of HIZ suggests that it is modular with functionally distinct domains, like most transcription factors e.g. HCF-1, Oct-1 and VP16. In addition to the zinc finger DNA-binding domain in the amino-terminal, there are distinctive regions enriched in specific residues with unknown function. In particular, a glutamine rich domain is of interest because it contains a novel 11-residue motif: EVTTVQLQPAQ repeated identically three times in full and once with 9 out of the 11 residues (see figure

8).   Similarly, Schutz and Niessing noted a 24aa helix-turn-helix motif (TNVQLQXLPQPXXVTNIQLQAXEV) that was repeated seven times in the chicken homolog cKr2. The (I/V)QLQ motif also appears independent of the 24aa repeat in cKr2 an additional five times. Similarly, in the murine Stat2 (signal transducers and activators of transcription 2) carboxy-terminal, there exists a transactivation domain and adjacent to it (towards the amino-terminal) is also a novel glutamine-rich repeat: APQVLLEP APQVQLEP repeated six times (Park *et al.*, 1999). Comparisons between the human and murine Stat2 proteins indicated that the primary sequence is very well conserved with the exception of the repeated motifs that lie adjacent to their respective activation domains. However, the divergence in the motifs caused the two homologs to interact with a distinct set of proteins (Park *et al.*, 1999). These findings suggest that HIZ may contain a novel glutamine-rich protein-binding motif that may function to bind other transcription factors for regulating its activation domain.

The predicted HIZ protein sequence also contains many putative phosphorylation sites (see figure 8). These may be significant because the protein phosphatase PP1 has been shown to interact with HCF-1 and thus, phosphorylation may be a mechanism used in HCF-1 regulation of HIZ.

4.3 HIZ Contains a Transactivation Domain

Through the construction of various HIZ regions tethered to the GAL4DBD, a transactivation domain was mapped to the carboxy-terminal at residues 876-996. Deletions from either end of this region resulted in a significantly lower transactivation potential. Differences in transactivation potential of smaller fragments of the activation

domain in BHK-21 and HeLa cells suggest that the HIZ transactivation domain relies on several cofactors for its full activation potential, some that may be present at different concentrations in the two cell lines.

VP16 is one of the strongest transcription activators yet characterized. In comparison, the HIZ activation domain is approximately one-seventh of its transactivation potential demonstrating the strength of this novel transcription factor.

4.4 HCF-1 Can Interact with HIZ and Decrease its Transactivation Potential

HCF-1 was shown to interact with and regulate the activation potentials of the transcription factors LZIP, Zhangfei, PGC-1β, GABP and Miz-1 (Lu *et al.,* 1997, Lu and Misra, 2000, Lin *et al.,* 2002, Vogel and Kristie, 2000a). While acting as a co-activator for LZIP, Zhangfei, PGC-1β and GABP, HCF-1 has been demonstrated to repress the transactivation potential of the cell cycle inhibitor, Miz-1.

Using *in vitro* GST-binding assays, HIZ demonstrated specific binding of 7-14% to GST-HCF-1 within residues 876-974, 913-996 and 876-996. These results indicate that the critical residues in HIZ for interaction with HCF-1 should be within the overlapping region 913-974. However, the binding assay demonstrated insignificant levels of binding within this region. It is possible that additional residues on *either* side of residues 913-974 are important in HCF-1 binding or that expression of residues 913-974 lacked the necessary folding to bind HCF-1. Interestingly, in HeLa cells, although HIZ$_{876-974}$ and HIZ$_{913-996}$ are both capable of significant transactivation, HIZ$_{913-974}$ transactivates only minimally. Thus, it appears that the fragments capable of binding

HCF-1 correlate to those important in transactivation. This region could be a site responsible for regulation of the HIZ transactivation potential.

Co-transfection of full length HCF-1 and $HCF_{750-902}$ in HeLa cells has shown that HCF-1 can decrease HIZ transactivation potential to approximately a quarter of its maximum activation potential. It is possible that at endogenous levels of HCF-1 in HeLa cells, HCF-1 does not bind HIZ, but associates with other factors that have a higher binding affinity, possibly GABP and/or Miz-1, which bind within $HCF_{813-847}$ and $HCF_{750-902}$ respectively. Upon changes in the cell that either signal the release of HCF-1 from its partners or increase HCF-1 expression, HCF-1 can bind and repress HIZ transactivation potential.

The repression by HCF-1 was shown to be specific to HIZ because addition of the same amounts of HCF-1 did not repress the transcriptional activity of the VP16 activation domain. Addition of HCF-1 to HeLa cells actually resulted in an increase in the VP16 transactivation, which correlates to preliminary data from our lab that HCF-1 may function synergistically with VP16 (Pat Bilan, personal communication). These results also coincide with previous data that HCF-1 can increase the activation potential of some binding partners (Lin *et al.*, 2002).

As described below, we have isolated an activation domain within HCF-1 capable of activating approximately 40-fold in HeLa cells. It is possible that HCF-1 exerts a repressive effect on GAL4-HIZ mediated transcription by masking the HIZ activation domain and replacing it with its own activation domain. HCF-1 may also affect HIZ activation potential by recruiting the phosphatase PP1 for dephosphorylation of key regulatory residues. HCF-1 has already been shown to be able to bind an array of other

factors so it is likely that the HCF-1-HIZ interaction must be stabilized by other factors *in vivo* because the *in vitro* binding assays show that this interaction is weak on its own. As depicted in figure 31, HCF-1 may require the amino-terminal HCF$_{VIC}$ domain to bind PGC-1α or PGC-1β and form a stable complex with nuclear hormone receptors. Our series of GAL4-HIZ fusions demonstrated that only the carboxy-terminal residues, and not the LxxLL motif, was necessary for the large activation potential of HIZ.

4.5 HIZ Expression

HIZ has been shown in this study to be expressed in specific tissues. We saw strong expression in the testis, liver, pancreas, heart and skeletal muscle.

The high expression of HIZ in the testis suggests a possible role in spermatogenesis (Gromoll *et al.*, 1997; de Rooij, 1998).

The recent findings that members of the PGC-1 family contain the HBM and experimental evidence that PGC-1β and PGC-1α can interact with HCF-1 *in vivo* are very interesting because as shown in this study, HIZ expression is tissue specific and was found to be expressed at high levels in the liver, while HCF-1 has been shown previously to be expressed at very low levels (Wilson *et al.*, 1995). PGC-1α and PGC-1β are coactivators of numerous nuclear hormone receptors such as PPARγ, and increased expression of PGC-1α and β has been shown in the liver during fasting, suggesting a role for both in regulating gluconeogenesis, β-oxidation of fatty acids, and/or ketogenesis (Yoon *et al.*, 2001). At low levels of HCF-1, PGC-1α and PGC-1β, the high expression of HIZ can potentially be involved in negative regulation of PGC-1α and PGC-1β-linked pathways in the liver. Thus, HIZ may regulate the constitutive transcription of factors in

pathways that are repressed only upon fasting. During these conditions, the concentrations of HCF-1 are potentially increased thereby synergistically activating PCC-1β while repressing HIZ transcription.

4.6 A Possible Role for HIZ in Suppression of Cell Proliferation

Hybridization of HIZ to a matched tumour/normal tissue array showed that HIZ expression is decreased in lung and uterine tumour tissue in comparison to normal tissue from the same individual. Therefore, HIZ may have a role in inhibiting cell proliferation in lung and uterine tissues. Its absence may cause the lack of transcription of one or more cell cycle inhibitors. HCF-1 has been shown to promote the cell cycle and therefore, HCF-1 may repress HIZ transcription initiation *in vivo*. Critical residues for HCF-1 binding correlated with those required for maximum activation potential implying an important role for HCF-1 regulation of HIZ transactivation of its target gene(s).

4.7 Obtaining Full-Length HIZ

We had a lot of difficulty in obtaining the full-length HIZ clone. This may have been caused by a large degree of secondary structure in the cDNA due to the high GC content in the 5' region. Also, our northern expression analysis indicated that there were alternate splice variants in skeletal muscle, pancreas and liver. The presence of alternate splice variants is verified by EST no. BG290052, which represents a spliced mRNA. Therefore, our putative sequence may represent full length HIZ but there is a chance that this sequence may be spliced at several sites, one of which is at the 5' end. There did not appear to be alternative splicing in the testis seen in our northern blot analysis, but our

difficulties in amplifying HIZ from a testis library may still have been caused by splice variants that occur at a lower abundance undetected by the northern blot.

Similar to Zhangfei, which has a predicted 1676nt 3' untranslated region (UTR), our putative sequence of HIZ is 3168nt but its mRNA is approximately 4200nt as detected in our northern blot analysis. Therefore, HIZ may also contain a large 3' UTR of approximately 1000nt. Long 5' and 3' UTRs have been implicated in post-transcriptional regulation in a variety of proteins including the vascular endothelial growth factor (VEGF), human androgen receptor (hAR) and human estrogen receptor (hER) (Cohen *et al.*, 1996; Faber *et al.*, 1991, Keavenley *et al.*, 1993).

Obvious splice variants were found at 3500nt and 1400nt. Thus, alternatively our putative sequence could be a spliced message represented by the band at 3500 and the 1400nt message could indicate an even smaller sliced message, possibly matching to EST BG290052, which lacks regions found in ESTs AI912315 (contains our putative sequence) and BG290052.

We believe that our putative sequence for HIZ is correct because of the following:

a) It resembles the chicken zinc finger protein cKr2 in sequence, size and structure,

b) There are mouse and rat ESTs with significant sequence homology to HIZ,

c) The sequence flanking the ATG initiator codon of HIZ conforms to the Kozak consensus sequence for translation, and

d) Upstream from this start site are stop codons in all three frames.

4.8 HCF-1 Contains an Activation Domain

Through recent findings of novel interactions between HCF-1 and the transcription factors LZIP, Zhangfei, GABP and PGC-1β, an important role for HCF-1 in increasing the activation potential of cellular transcription has been shown. We have identified an autonomous activation domain within the acidic domain of HCF-1 (HCF$_{1573-1726}$), which may play a part in synergistically activating the collection of transcription factors that HCF-1 can interact with, including VP16.

PGC-1β was shown to be able to increase the transcriptional activity of GAL4DBD-HCF-1$_{N380}$ from minimal induction (approximately 5-fold) to approximately 140-fold induction of reporter gene upon cotransfection (Lin *et al.*, 2002). Whether this transactivation ability of the amino-terminal 380 residues of HCF-1 can be attributed to recruitment of the carboxy-terminal activation domain (through the self-association domains), or by another activation domain present in the HCF-1 amino-terminal is unknown. However, the recruitment of this activation domain to the amino-terminal could provide a reason for the unique self-association of the amino- and carboxy-terminals seen in HCF-1. Regardless, the isolation of one activation domain within HCF-1 demonstrates the capability of HCF-1 to function as a transcription activator in addition to its previously identified role simply as a coregulator of transactivation.

Whether HIZ interacts with other HCF-1 binding factors is a matter that will have to be determined in the future. However, the presence of LxxLL motifs in LZIP, Zhangfei, Miz-1 and HIZ demonstrates the potential for nuclear hormone receptor binding to form a complex with HCF-1. Mutation of the two LxxLL motifs in LZIP do no affect binding to HCF-1 but alter its activation potential, suggesting that nuclear

hormone receptors play a role in regulating the transactivation of LZIP. Similarly, HIZ may be a strong transcription factor regulated by HCF-1 in concert with other factors through its LxxLL motif.

As mentioned above, HCF-$1_{750\text{-}902}$ was sufficient for HIZ binding and for the repression effect it exerted on the GAL4-HIZ activation potential. However, HCF-1 requires regions in addition to the basic domain to act as a coactivator for GABP-mediated transcription (Vogel and Kristie, 2000a). Co-transfection of GABP with the region of HCF-1 responsible for HCF-1-GABP binding resulted in the repression of GABP-dependent transcription, likely because the HCF-1 fragment was acting as a competitive inhibitor for wild-type endogenous HCF-1 (Vogel and Kristie, 2000a). Interestingly, co-transfection with full-length HCF-1 resulted in higher levels of GAL4-GABP transcription (Vogel and Kristie, 2000a), supporting our evidence from HCF-1 cotransfections with GAL4-VP16 that HCF-1 can synergistically activate on its own, possibly through the activation domain we have identified.

4.9 Summary

In conclusion, the data presented in this thesis have addressed the objectives set and show the following:

a) HIZ is specifically expressed in abundance in the testis, liver, pancreas, heart and skeletal muscle and displays a higher expression in normal uterine and lung tissue in comparison to matching tumour tissue;

b) A putative sequence defines an extensive DNA-binding domain consisting of 16 zinc finger structures and regions enriched in proline/alanine, glycine and glutamine;

c) HIZ contains a very strong activation domain within residues 876-996;

d) Residues 750-902 within the HCF-1 basic domain bind specifically, but weakly to HIZ at residues 876-974, 876-996 and 913-996, regions that activate the strongest; and

e) Co-transfection with large levels of full length HCF-1 represses HIZ activation potential by approximately 75%.

Figure 31 is a schematic representation of the potential complex necessary for full HIZ transactivation *in vivo*. Residues at the both ends of the transactivation domain likely recruit other co-factors for transcription because deletions of either end showed different activation potential dependent on the cell line. Also, deletions of either end retained binding to HCF-1 but deletion of both abolished HCF-1 binding possibly because the HCF-1 interaction is stabilized by other factors that bind at either end.

We believe that HCF-1 does not act as a co-activator for HIZ as demonstrated for LZIP, Zhangfei, GABP and PGC-1β, but as a co-repressor, exerting its repression effect only at high levels in the cell, and/or through recruitment by other HIZ-binding factors as shown in figure 32.

In addition, we have identified an activation domain within the acidic region of HCF-1 that may be responsible for increasing the activation potential of proteins such as LZIP, Zhangfei, GABP and PGC-1β which require HCF-1 binding for full activation potential.

Thus, our results indicate that HCF-1 can play a role in both positive and negative regulation of transcription initiation. Contrary to previous published studies of HCF-1 binding proteins, we have isolated a novel HCF-1 interacting protein that is repressed by the presence of HCF-1. The function of this novel zinc finger is unknown but its expression pattern suggests a role in metabolism in the liver and/or regulation of a cell cycle inhibitor. Homologs of HIZ are found in chicken, mice and rats implying an important, evolutionarily conserved role for HIZ that must be determined in future studies.
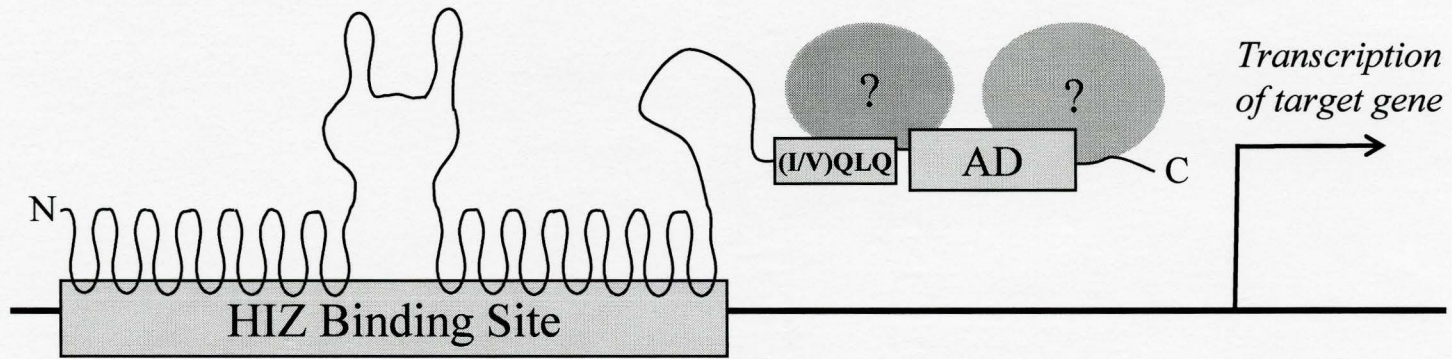
Figure 31. Lack of HCF-1 Allows HIZ to Activate its Target Gene. In the absence of HCF-1 interaction, HIZ likely recruits other factors for maximal transactivation because the HIZ AD showed varying transactivation potential in different cell lines, likely due to the presence or absence of co-factors. AD = transactivation domain
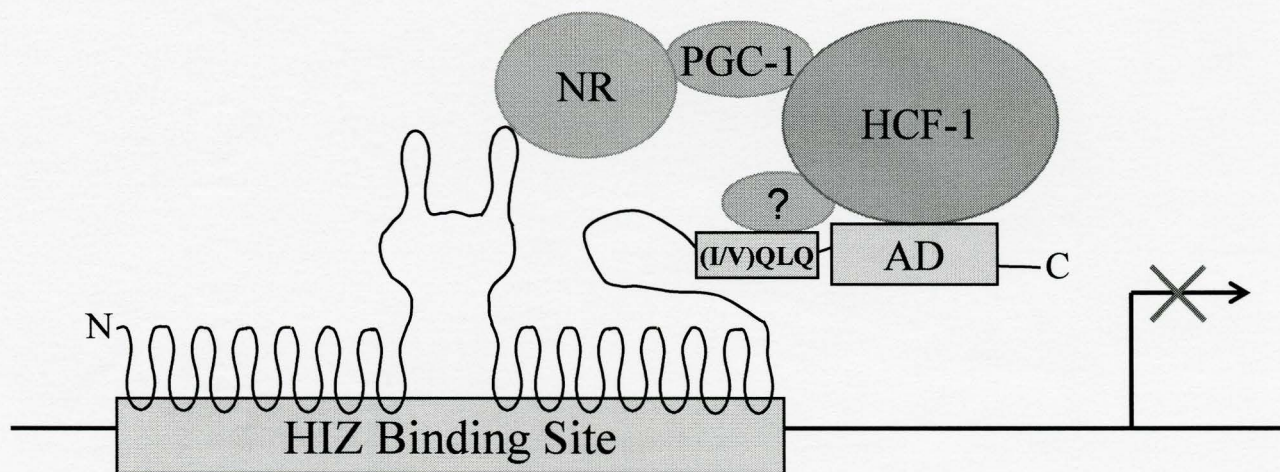
Figure 32. Potential Mechanism of HCF-1 Mediated Repression of HIZ Activation Potential. HCF-1 basic domain binding to the HIZ AD may be stabilized by other HIZ binding factors that bind through the novel glutamine motif adjacent to the AD and/or through interactions with NRs through the LxxLL motif in the unlinked zinc finger 9 of HIZ. NRs have been shown to bind PGC-1α and PGC-1β. PGC-1α and PGC-1β in turn, have been shown to bind HCF-1 at the HCF$_{KEL}$ amino-terminal domain. AD = activation domain; NR = nuclear hormone receptor

REFERENCES

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulous, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R. and J.C. Venter. 1991. Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science.* **252**:1651-1656.

Ajuh, P.M., Browne, G.J., Hawkes, N.A., Cohen, P.T.W., Roberts, S.G.E. and A.I. Lamond. 2000. Association of a protein phosphatase 1 activity with the human factor C1 (HCF) complex. *Nucleic Acids Res* **28**:678-686.

Aranda, A and A. Pascual. 2001. Nuclear hormone receptors and gene expression. *Physiol Rev* **81**:1269-1304.

Ausubel, F., Brent, R., Kingston, R.E., Moore, D.M., Seidman, J.G., Smith, J.A. and Struhl, K. 1997. Current Protocols in Molecular Biology. Toronto: J. Wiley & Sons Canada, Inc.

Beguli, N.S., Sepullveda, J.L., Nigam, V., Charron, F., Nemer, M. and R.J. Schwartz. 2000. Cardiac tissue enriched factors serum response factor and GATA-4 are mutual co-regulators. *Mol Cell Biol* **20**:7550-8.

Blom, N., Gammeltoft, S. and S. Brunak. 1999. Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol* **294**:1351-1362.

Bohm, S., Frishman, D. and H.W. Mewes. 1997. Variation of the C2H2 zinc finger motif in the yeast genome and classification of yeast zinc finger proteins. *Nucleic Acids Res* **25**:2464-2469.

Bork, P. and R.F. Doolittle. 1994. Drosophila kelch motif is derived from a common enzyme fold. *J Mol Biol* **236**:1277-1282.

Bowers, P.M., Schaugler, L.E., Klevit, R.E. 1999. A folding transition and novel zinc finger accessory domain in the transcription factor ADR1. *Nat Struct Biol* **6**:475-487.

Braunstein, M, Rose, A.B., Holmes, S.G., Allis, C.D., Broach, J.R. 1993. Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. *Genes Dev* **7**:592-604.

Brown, R.S., Sander, C., P. Argos. 1985. The primary structure of transcription factor TFIIIA has 12 consecutive repeats *FEBS Lett* **186**:271-274.

Burbelo, P.D., Gabriel, G.C., Kibbey, M.C., Yamada, Y, Kleinman, H.K., and B.S. Weeks. 1994. LZIP-1 and LZIP-2: Two Novel Members of the bZIP Family. *Gene* **139**:241-245.

Burge, C. and S. Karlin. 1997. Prediction of complete gene structures in human genomic DNA. *J Mol Biol* **268**:78-94.

Chen, J.L., Attardi, L.D., Verrijzer, C.P., Yokomori, K. and R. Tjian. 1994. Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators. *Cell* **79**:93-105.

Choo, Y. and A. Klug. 1993. A role in DNA binding for the linker sequences of the first three zinc fingers of TFIIIA. *Nucleic Acids Res* 21:3341-3346.

Clarke, N.D. and J.M. Berg. 1998. Zinc fingers in *Caenorhabditis elegans*: finding families and probing pathways. *Science* **282**:2018-2022.

Cleary, M.A., Stern, S., Tanaka, M. and W. Herr. 1993. Differential positive control by Oct-1 and Oct-2: Activation of a transcriptionally silent motif through Oct-1 and VP16 corecruitment. *Genes Dev* 7:72.

Cohen, T., Nahari, D., Cerem, L.W., Neufeld, G. and B.Z. Levi. 1996. Interleukin 6 induces the expression of vascular endothelial growth factor. *J Biol Chem* **271**:736-741.

Cousens, D.J., Greaves, R., Goding, C.R. and P. O'Hare. 1989. The C-terminal 79 amino acids of he herpes simplex virus regulatory protein, Vmw65, efficiently activate transcription in yeast and mammalian cells in chimeric DNA-binding proteins. *EMBO J* **8**:2337.

Crepin, M, Cukier-Kahn, R and F. Gros. 1975. Effect of a low-molecular-weight DNA binding protein, H1 factor, on the in vitro transcription of the lactose operon in Escherichia coli. *PNAS* **72**:333-337.

de Rooij, D.J. 1998. Stem cells in the testis. *Int J Exp Pathol* **79**:67-80.

Dynlacht, B.D., Hoey, T. and R. Tjian. 1991. Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* **66**:563-576.

Elrod-Erickson, M., Rould, M.A., Nekludova, L. and C.O. Pabo. 1996. Zif268 protein-DNA complex refined at 1.6 A: a model system for understanding zinc finger-DNA interactions. *Structure* **4**:1171-1180.

Faber, P.W., van Rooij, H.C., van der Korput, H.A., Baarends, W.M., Brinkmann, A.O., Grootegoed, J.A. and J. Trapman. 1991. Characterization of the human androgen receptor transcription unit. *J Biol Chem* **266**:10743-10749.

Fairall, L., Harrison, S.D., Travers, A.A. and D. Rhodes. 1992. Sequence-specific DNA binding by a two-zinc-finger peptide from the Drosophila melanogaster Tramtrack protein. *J Mol Biol* **225**:349-366.

Fairall, L., Schwabe, J.W.R., Chapman, L., Finch, J.T. and D. Rhodes. 1993. The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc finger/DNA recognition. *Nature* **366**:483-487.

Falcone, D. and D.W. Andrews. 1991. Both the 5' untranslated region and the sequences surrounding the start site contribute to efficient initiation of translation in vitro. *Mol Cell Biol* **11**:2656-64.

Flint, J. and T. Shenk. 1997. Viral transactivating proteins. *Annu. Rev. Genet.* **31**:177-212.

Foster, M.P., Wuttke, D.S., Radhakrishnan, I., Case, D.A., Gottesfeld, J.M. and P.E. Wright. 1997. Domain packing and dynamics in the DNA complex of the N-terminal zinc fingers of TFIIIA. *Nat Struct Biol* **4**:605-608.

Frattini, A., Faranda, S., Redolfi, E., Zucchi, I., Villa, A., Patrosso, M.C., Strina, D., Susani, L., and Vezzoni, P. 1994. Genomic Organization of the Human VP16 Accessory Protein, a Housekeeping Gene (HCFC1) Mapping to Xq28. *Genomics* **23**:30-35.

Freiman, R.N. and W. Herr. 1997. Viral mimicry: common mode of association with HCF by VP16 and the cellular protein LZIP. *Genes Dev* **6**:466-480.

Gerster, T., and R.G. Roeder. 1988. A herpesvirus trans-activating protein interacts with transcription factor OTF-1 and other cellular proteins. *Proc Natl Acad Sci* **85**:6347-6351.

Goto, H., Motomura, S., Wilson, A.D., Freiman, R.N., Nakabeppu, Y., Fukushima, K., Fujishima, M, Herr, W. and T. Nishimoto. 1997. A single-point mutation in HCF causes temperature-sensitive cell-cycle arrest and disrupts VP16 function. *Genes Dev.* **11**:726-37.

Greenblat, J. 1997. RNA polymerase II holoenzyme and transcriptional regulation. *Curr Opin Cell Biol* **9**:310-319.

Gromoll, J., Wessels, J., Rosiepen, G., Brinkworth, M.H. and G.F. Weinbauer. 1997. Expression of mitotic cyclin B1 is not confined to proliferating cells in the rat testis. *Biol Reprod* **57**:1312-1319.

Gunther, M., Laithier, M. and O. Brison. 2000. A set of proteins interacting with transcription factor Sp1 identified in a two-hybrid screening. *Mol Cell Biochem* **210**:131-1i42.

Hansen, S.K. and R. Tjian. 1995. TAFs and TFIIa mediate differential utilization of the tandem Adh promoters. *Cell* **82**:565-575.

Hebbes, T.R., Clayton, A.L. Thorne, A.W. and C. Crane-Robinson. 1994. A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J* 7:1395-1402.

Heery, D.M., Kalkhoven, E., Hoare, S. and M.G. Parker. 1997. A signature motif in transcriptional c-activators mediates binding to nuclear receptors. *Nature* **387**:733-736.

Herr, W. 1998. The Herpes Simplex Virus VP16-induced complex: mechanisms of combinatorial transcriptional regulation. *Cold Spring Harbor Symp Quant Biol* **63**:599-607.

Herr, W. and M.A. Cleary. 1995. Mechanisms for flexibility in DNA sequence recognition and VP16-induced complex formation by the Oct-1 POU domain. *Mol Cell Biol.* **15**:2090-100.

Hori, R. and M. Carey. 1994. The role of activators in assembly of RNA polymerase II transcription complexes. *Curr Opin Genet Dev* **4**:236-44.

Johnson, K.M., Mahajan, S.S. and A.C. Wilson. 1999. Herpes simplex virus transactivator VP16 discriminates between HCF-1 and a novel family member, HCF-2. *J Virol* **73**:3930-3940.

Jones, K.A. and R. Tjian. 1985. Sp1 Binds to Promoter Sequences and Activates Herpes Simplex Virus 'Immediate-Early' Gene Transcription *in vitro*. *Nature* **317**:179-182.

Kadosh, D. and K. Struhl. 1997. Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters. *Cell* **89**:365-371.

Katan, M.A., Haigh, C.P., Verrijzer, P.C., van der Vliet, P.C. and P. O'Hare. 1990. Characterization of a cellular factor which interacts functionally with Oct-1 in the assembly of a multicomponent transcription complex. *Nucleic Acids Res* **18**:6871-6880.

Keaveney, M., Parker, M.H. and F. Gannon. 1993. Identification of a functional role for the 3' region of the human oestrogen receptor gene. *J Mol Endocriniol* **10**:143-152.

Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**:283-292.

Kristie, T.M., LeBowitz, J.H. and P.A. Sharp. 1989. The octamer-binding proteins form multi-protein-DNA complexes with the HSV alpha TIF regulatory protein. *EMBO J* **8**:4229-4238.

Kristie, T.M., and P.A. Sharp. 1993. Purification of the Cellular C1 Factor Required For the Stable Recognition of the Oct-1 Homeodomain by the Herpes Simplex Virus Alpha-Trans-Induction Factor (VP16). *J Biol Chem* **268**:6525-6534.

Kristie, T.M., Pomerantz, J.L., Twomey, T.C., Parent, S.A. and Sharp, P.A. 1995. The cellular C1 factor of the herpes simplex virus enhancer complex is a family of polypeptides. *J Biol Chem* **270**:4387.

Kristie, T.M. 1997. The mouse homologue of the human transcription factor C1 (host Cell factor). *J Biol Chem* **272**:26749-26755.

Kristie, T.M., Vogel, J.L. and A.E. Sears. 1999. Nuclear localization of the C1 factor (host cell factor) in sensory neurons correlates with the reactivation of herpes simplex virus form latency. *Proc Natl Acad Sci* **96**:1229-1233.

LaBoissiere, S., Walker, S., and P. O'Hare. 1997. Concerted Activity of Host Cell Factor Subregions in Promoting Stable VP16 Complex Assemble and Preventing Interference by the Acidic Activation. *Mol Cell Biol* **17**:7108-7118.

La Boissiere, S., Hughes, T. and P. O'Hare. 1999. HCF-dependent nuclear import of VP16. *EMBO J* **18**:480-489.

Lai, J-S, Cleary, M.A. and W. Herr. 1992. A single amino acid exchange transfers VP16-induced positive control from the Oct-1 to the Oct-2 homeo domain. *Genes Dev* **6**:2058-65

Lai, J-S and W.Herr. 1997. Interdigitated residues within a small region of VP16 interact with Oct-1, HCF and DNA. *Mol Cell Biol* **17**:3937-3946.

LaMarco, K.L. and S.L. McKnight. 1989. Purification of a Set of Ceullular Polypeptides that Bind to the Purine-rich *cis*-regulatory Element of Herpes Simplex Virus Immediate Early Genes. *Genes Dev* **3**:1372-1383.

Lander, E.S., Linton, L.M., Birren, B. *et al.* 2001. Initial Sequencing and analysis of the human genome. *Nature* **409**:860-921.

Lee, M.S., Gippert, G.P., Soman, K.V., Case, D.A. and P.E. Wright. 1989. Three-dimensional solution structure of a single zinc finger DNA-binding domain. *Science* **245**:635-637.

Lee, S. and W. Herr. 2001. Stabilization but not the transcriptional activity of herpes simplex virus VP16-induced complexes is evolutionarily conserved among HCF family members. *J Virol* **75**:12402-11.

Lemon, B. and R. Tjian. 2000. Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev* **14**:2551-2569.

Lichter, P., Bray, P., Ried, T., Dawid, I.B. and D.C. Ward. 1992. Clustering of C2-H2 zinc finger motif sequences within telomeric and fragile site regions of human chromosomes. *Genomics* **13**:999-1007.

Lin, J., Puigserver, P., Donovan, J., Tarr, P. and B.M. Spiegelman. 2002. PGC-1β: A novel PGC-1 related transcription coactivator associated with host cell factor. *J Biol Chem.* **277**:1645-1648.

Liu, Y., Hengartner, M.O. and W. Herr. 1999. Selected Elements of Herpes Simplex Virus Accessory Factor HCF Are Highly Conserved in *Caenorhabditis elegans*. *Mol Cell Biol* **19**:909-915.

Lu, R., Yang, P., O'Hare, P. and V. Misra. 1997. Luman, a New Member of the CREB/ZTF Family, Binds to Herpes Simplex Virus Vp16-Associated Host Cellular Factor. *Mol Cell Biol* **17**:5117-5126.

Lu, R., Yang, P., Padmakumar, S. and V. Misra. 1998. The herpesvirus transactivator VP16 mimics a human basic domain leucine zipper protein, luman, in its interaction with HCF. *J Virol.* **72**:6291-7.

Lu, R and V, Misra. 2000. Zhangfei: a second cellular protein interacts with herpes simplex virus accessory factor HCF in a manner similar to Luman and VP16. *Nucleic Acids Res* **28**:2446-2454.

Luciano R.L. and A.C. Wilson. 2000. N-terminal transcriptional activation domain of LZIP comprises two LxxLL motifs and the host cell factor-1 binding motif. *Proc Natl Acad Sci* **97**:10757-62.

Luo, Y., Ge, H., Stevens, S., Xiao, H. and R.G. Roeder. 1998. Coactivation by OCA-B: definition of critical regions and synerism with general cofactors. *Mol Cell Biol.* **18**:3803-10.

MacKay, J.P. and M. Crossley. 1998. Zinc fingers are sticking together. *Trends Biochem Sci* **23**:1-4.

Merika, M., Williams, A.J., Chen, G., Collins, T. and D. Tahnos. 1998. Recruitment of CBP/p300 by the IFNβ enhanceosome is required for synergistic activation of transcription. *Mol Cell* **1**:277-287.

Michael, S.F., Kilfoil, V.J., Schmidt, M.H., Amann, B.T. and J.M. Berg. 1992. Metal binding and folding properties of a minimalist Cys$_2$His$_2$ zinc finger peptide. *Proc Natl Acad Sci* **89**:4796-4800.

Miller, J., McLachlan, A.D. and A. Klug. 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J* **4**:1609-14.

Nuchprayoon, I., Shang, J., Simkevich, C.P., Luo, M., Rosmarin, A.G. and A.D. Freidman. 1999. An Enhancer Located Between the Neutrophil Elastase and Proteinase 3 Promoters is Activated by Sp1 and and ETS Factor. *J Biol Chem* **274**:1085-1091.

Ogbourne, S. and T.M. Antalis. 1998. Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem J* **331**:1-14.

O'Hare, P. 1993. The Virion Transactivator of Herpes Simplex Virus. *Semin Virol* **4**:145-155.

Parraga, G., Horvath, S.J., Eisen, A., Taylor, W.E., Hood, L., Young, E.T., R.E. Klevit. 1988. Zinc-dependent structure of a single-finger domain of yeast ADR1. *Science* **241**:1489-1492.

Park, C., Lecomte, M.J. and C. Schindler. 1999. Murine Stat2 is uncharacteristically divergent. *Nucleic Acids Res* **27**:4191-4199.

Pavletich, N.P. and C.O.Pabo. 1991. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science* **252**:809-817.

Pugh, B.F. and R. Tjian. 1990. Mechanism of transcriptional activation by recruitment. *Nature* **386**:569-577.

Puigserver, P., Wu, Z., Park, C.W., Graves, R., Wright, M. and B.M. Spiegelman. 1998. A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell* **92**:829-839.

Roeder, R.G. and W.J. Rutter. 1969. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224**:234-237.

Rosmarin, A.G., Luo, M., Caprio, D.G., Shang, J. and C.P. Simkevich. 1998. Sp1 Cooperates With the ETS Transcription Factor, GABP, to Activate the CD18 (β2 leukocyte integrin) Promoter. *J Biol Chem* **273**:19097-13103.

Ryan, R.F. and M.K. Darby. 1998. The role of zinc finger linkers in p43 and TFIIIA binding to 5S rRNA and DNA. *Nucleic Acids Res* **26**:703-709.

Sadowski, I.., Ma, J, Triezenberg, S.J. and M. Ptashne. 1988. GAL4-VP16 is an unusually potent transcriptional activator. *Nature* **335**:563-564.

Sadowski, I. and Ptashne, M. 1989. A vector for expressing GAL4(1-147) fusions in mammalian cells. *Nucleic Acids Res* **17, 7539**

Schaeffer, L., Duclert, N., Huchet-Dymanus, M. and J.P. Changeux. Implication of a multisubunit Ets-related transcription factor in synaptic expression of the nicotinic acetylcholine receptor. *EMBO J* **17**:3078-3090.

Schutz, B and J. Niessing. 1994. Cloning and structure of a chicken zinc finger cDNA: restricted expression in developing neural crest cells. *Gene* **148**:227-36.

Schuh R, Aicher W, Gaul U, Cote S, Preiss A, Maier D, Seifert E, Nauber U, Schroder C and R. Kemler. 1986. A conserved family of nuclear proteins containing structural elements of the finger protein encoded by Kruppel, a Drosophila segmentation gene. *Cell* **47**:1025-1032.

Shastry, E.S. 1996. Transcription factor IIIA (TFIIIA) in the second decade. *J Cell Sci* **109**:535-539.

Staller, P., Peukert, K., Kiermaier, A., Seoane, J., Lukas, J., Karsunky, H., Moroy, T., Bartek, J., Massague, J, Hanel, F., Eilers, M. 2001. Repression of p15INK4b expression by Myc through association with Miz-1. *Nat Cell Biol* **3**:392-399.

Stern, S., Tanaka, M. and W. Herr. 1989. The Oct-1 homeodomain directs formation of a multiprotein-DNA complex with the HSV transactivator VP16. *Nature* **341**:624.

Stubbs, L., Carver, E.A., Shannon, M.E., Kim, J., Geisler, J., Generoso, E.E., Stanford, B.G., Dunn, W.C., Mohrenweiser, H., Zimmermann, W., Watt, S.M. and L.K. Ashworth. 1996. Detailed comparative map of human chromosome 19q and related regions of the mouse genome. *Genomics.* **35**:499-508.

Tanaka, M., Grossniklaus, U., Herr, W., and N. Hernandez. 1988. Activation of the U2 snRNA promoter by the octamer motif defines a new class of RNA polymerase II enhancer elements. *Genes Dev* **2**:1764-1778.

Tanaka, M. and W. Herr. 1990. Differential transcriptional activation by Oct-1 and Oct-2: Interdependent activation domains induce Oct-2 phosphorylation. *Cell* **60**:375.

Tanaka, M, Lai, J.S. and W. Herr. 1992. Promoter-selective activation domains in Oct-1 and Oct-2 direct differential activation of an snRNA and mRNA promoter. *Cell* **68**:755-767.

Tanese, N., Pugh, B.F. and R. Tjian. 1991. Coacivators for a proline-rich activator purified form the multisubunit human TFIID complex. *Genes Dev* **5**:2212-2224.

Takatsuji, H. 1998. Zinc-finger transcription factors in plants. *Cell Mol Life Sci* **54**:582-596.

Triezenberg, S.J., LaMarco, K.L. and S.L. McKnight. 1988. Evidence of DNA:protein interactions that mediate HSV-1 immediate early gene activation by VP16. *Genes Dev* **2**:730-742.

Vassias, I., Hazan, U., Michel, T., Sawa, C., Handa, H., Gouya, L., and F. Morinet. 1998. Regulation of Human B19 Parvovirus Promoter Expression by hGABP (E4TF1) Transcription Factor. *J Biol Chem* **273**:8287-8293.

Vega, R.B., Huss, J.M and D.P. Kelly. 2000. The coactivator PGC-1 cooperates with peroxisome proliferators-activated receptor alpha in transcriptional control of nuclear genes encoding mitochondrial fatty acid oxidation enzymes. *Mol Cell Biol.* **20**:1868-1876.

Vogel, J.L. and T.M. Kristie. 2000a. The novel coactivator C1 (HCF) coordinates multiprotein enhancer formation and mediates transcription activation by GABP. *EMBO J* **19**:683-690.

Vogel, J.L. and T.M. Kristie. 2000b. Autocatalytic proteolysis of the transcription factor-coactivator C1 (HCF): A potential role for proteolytic regulation of coactivator function. *PNAS* **97**:9425-9430.

Westuck, G.H. and J.P. Capone. 1993. An unusual cellular factor potentiates protein-DNA complex assembly between Oct-1 and Vmw65. *J Biol Chem* **268**:1272-1278.

Wilson, A.C., LaMarco, K., Peterson, M.G. and W. Herr. 1993. The VP16 Accessory Protein HCF is a Family of Polypeptides Processed from a Large Precursor Protein. *Cell* **74**:115-125.

Wilson, A.C., Parrish, J.E., Massa, H.F., Nelson, D.L., Trask, B.J. and W. Herr. 1995. The gene encoding the VP16-accessory protein HCF (HCFC1) resides in human Xq28 and is highly expressed in fetal tissues and the adult kidney. *Genomics* **25**:462-268.

Wilson, A.C., Freiman, R.N., Goto, H., Nishimoto, T., and W. Herr. 1997. VP16 targets an amino-terminal domain of HCF involved in cell-cycle progression. *Mol Cell Biol* **17**:6139-6146.

Wolfe, S.A., Nekludova, L., Pabo, C.O. 2000. DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struc.* **29**:183-212.

Wolffe, A. and D. Pruss. 1996. Targeting chromatin disruption: transcription regulators that acetylate histones. *Cell* **84**:817-819.

Wu, Z., Puigserver, P., Andersson, U., Zhang, C., Adelmant, G., Mootha, V., Troy, A., Cinti, S., Lowell, B., Scarpulla, R.C. and B.M. Spiegelman, 1999. Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell* **98**:115-124.

Wysocka, J., Reilly, P.T. and W. Herr. 2001. Loss of HCF-1-Chromatin Association Precedes Temperature-Induced Growth Arrest of tsBN67 Cells. *Mol Cell Biol* **21**:3820-3829.

Yoon, J.C., Puigserver, P., Chen, G., Donovan, J., Wu, Z., Rhee, J., Adelmant, G., Stafford, J., Kahn, C.R., Granner, D.K., Newgard, C.B. and B.M. Spiegelman. 2001. *Nature* **413**:131-138.

Xue, F., and L. Cooley. 1993. Kelch encodes a component of intercellular bridges in Drosophila egg chambers. *Cell* **72**:681-693.

Zhang, Y., LeRoy, G., Seelig, H.P., Lane, W.S. and D. Reinberg. 1998. The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodelling activities. *Cell* **95**:279-289.