2009-06-16

Biophysical Analysis of the AP1-DNA Interaction

Kenneth Ladd Seldeen
University of Miami, editor@scientificameriken.com

UNIVERSITY OF MIAMI

BIOPHYSICAL ANALYSIS OF THE AP1-DNA INTERACTION

By

Kenneth L. Seldeen

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

June 2009

UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

BIOPHYSICAL ANALYSIS OF THE AP1-DNA INTERACTION

Kenneth L. Seldeen

Approved:

Thomas K. Harris, Ph.D. Associate Professor of Biochemistry & Molecular Biology	Terri A. Scandura, Ph.D. Dean of the Graduate School
Amjad Farooq, Ph.D., D.I.C. Assistant Professor of Biochemistry & Molecular Biology	Arun Malhotra, Ph.D. Associate Professor of Biochemistry & Molecular Biology
Zafar Nawaz, Ph.D. Associate Professor of Biochemistry & Molecular Biology	Fenfei Leng, Ph.D. Associate Professor of Biochemistry Florida International University

SELDEEN, KENNETH L. (Ph.D., Biochemistry and Molecular Biology)
Biophysical Analysis of the AP1-DNA Interaction (June 2009)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Assistant Professor Amjad Farooq.
No. of pages in text. (139)

Jun and Fos are components of the AP1 family of transcription factors that bind to the promoters of a diverse multitude of genes involved in critical cellular responses such as cell growth and proliferation, cell cycle regulation, embryonic development and cancer. The specific protein-DNA interactions are driven by the binding of basic zipper (bZIP) domains of Jun and Fos to TPA response element (TRE) and cAMP response element (CRE) within the promoters of target genes.

Here, using a diverse array of biophysical techniques, including in particular isothermal titration calorimetry in conjunction with molecular modeling and semi-empirical analysis, I characterize AP1-DNA interactions in thermodynamic and structural terms. My data show that the binding of bZIP domains of Jun-Fos heterodimer to TRE and CRE are under enthalpic control accompanied by entropic penalty at physiological temperatures. This is in agreement with the notion that protein-DNA interactions are largely driven by electrostatic interactions and intermolecular hydrogen bonding. A larger than expected heat capacity change suggests that the basic regions within the bZIP domains are unstructured in the absence of DNA and interact in a coupled folding and binding manner. Further analysis demonstrates that Jun-Fos heterodimer can tolerate single nucleotide variants of the TRE consensus sequence and binds in the biologically relevant micromolar to submicromolar range. Of particular interest is the observation that

the Jun-Fos heterodimer binds to specific variants in a preferred orientation. 3D atomic models reveal that such preference in orientation results from asymmetric binding and may in part be attributable to chemically distinct but structurally equivalent residues within the basic regions of Jun and Fos. I further demonstrate that binding of the biologically relevant Jun-Jun homodimer to TRE and CRE occurs with favorable enthalpic contributions accompanied by entropic penalty at physiological temperatures in a manner akin to the binding of Jun-Fos heterodimer. However, anomalously large negative heat capacity changes provoke a model whereby Jun loads onto DNA as unfolded monomers coupled with subsequent folding and homodimerization upon association. The data also reveal that the heterodimerization of leucine zippers is modulated by the basic regions and these regions may undergo at least partial folding upon heterodimerization. Large negative heat capacity changes accompanying the heterodimerization of leucine zippers are consistent with the view that leucine zippers do not retain α-helical conformation in isolation and the formation of the native coiled coil α-helical dimer is attained through a coupled folding-dimerization mechanism.

Taken together, this dissertation marks the first comprehensive thermodynamic analysis of an otherwise well-studied and vitally important transcription factor. My studies shed new light on the forces driving the AP1-DNA interaction in thermodynamic and structural terms. The implications of these novel findings on the development of novel therapies for the treatment of disease with greater efficacy coupled with low toxicity cannot be overemphasized.

DEDICATION

To my wife,
Claudia

To my good friend,
Dr. Martin Gish

ACKNOWLEDGEMENTS

The road to my Ph.D. was, without question, unlike that of any other. Certainly challenge was to be expected. I was a foreigner to the culture of science as much as I was a foreigner to the city of Miami. My background was computers, years since I was last in a lab – a marine biology lab at that. Yet, I desired to enter the realm of science, to inject my personality, background and drive to discover something and improve life for others. The path was not easy and simply would not be possible if not for the amazing people I have met on this journey. For that reason, I would like take a moment to acknowledge those that have helped me towards the completion of this thesis.

First and foremost, I would like to express my appreciation to Amjad Farooq for his extraordinary mentorship. Amjad's vision to provide an intellectually stimulating and highly friendly scientific environment was a key factor in the pursuit of my research goals in a timely manner, not to mention the constant availability of granola bars to meet my daily energy demands. In particular, access to blistering technology and a suite of state-of-the-art equipment undoubtedly paved the way for my rapid progress to completion of my doctoral studies. Additionally, Amjad's enthusiasm, energy and his passion for science coupled with stimulating discussions, virtually on a daily basis, were instrumental toward my development into an independent scientist. I would also like to add that it has been a privilege to work with two wonderful colleagues in Caleb and Brian in the Farooq Laboratory. It was fun working together, we could overcome any problem the lab faced and the good times and good humor, I will always remember.

I also take the opportunity to express my deep thanks to my thesis committee members, Thomas Harris, Arun Malhotra and Zafar Nawaz. I am very grateful for their

input and critical evaluation of my work over the past several years. I also thank Dr. Fenfei for agreeing to serve as External Examiner on my thesis committee and for critically evaluating my work.

I would also like to thank my parents, Richard and Audra, for their loving support throughout my time here in Miami. I cannot express in its entirety the gratefulness within for everything they have done and just how much it means to me to see that Christmas stocking arrive through the mail. I would also like to thank my siblings, Dana, Bill and Doug, for keeping in contact and for their love and support. I would especially like to thank my new parents, Claudio and Delia, for being like a second family to me and helping me throughout the years. I would also like to thank all the friends I have made, at Miami and in California, who have always been supportive during these years.

Reserved in a special place in my heart, I would like to say “Thank you” to my wife, Claudia. We began our relationship as lab mates and made great scientific progress from there. She has been so supportive throughout my Ph.D. and has brightened so many of my days. I am truly fortunate that I was able to share every part of this wonderful experience with a person I love and adore. The completion of this thesis marks a new phase in my life and I look forward to new adventures with Claudia on my side.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 Chapter 1: Introduction

1.1 Jun and Fos modulate gene expression in a multitude of pathways and play a significant role in cancer

Eukaryotic cells have evolved many strategies to transduce extracellular signals through the plasma membrane to the nucleus. Some strategies may be as simple as the diffusion of chemicals, such as steroid hormones, through the membrane to nuclear receptors, while others may involve complex signaling cascades [1, 2]. Signals pass from the plasma membrane through the cytoplasm, via protein-protein interactions, ultimately leading to the activation of transcription factors. Through critical protein-DNA interactions, transcription factors such as Jun and Fos, known collectively as Activator Protein 1 (AP1), can modulate gene expression. As an example, by binding to the promoters of genes like metallothionein IIa, collagenase, interleukin2 and cyclin D1 among many others (Figure 1-1), AP1 can orchestrate a multitude of cellular processes including cell growth, proliferation, cell cycle regulation and embryonic development,

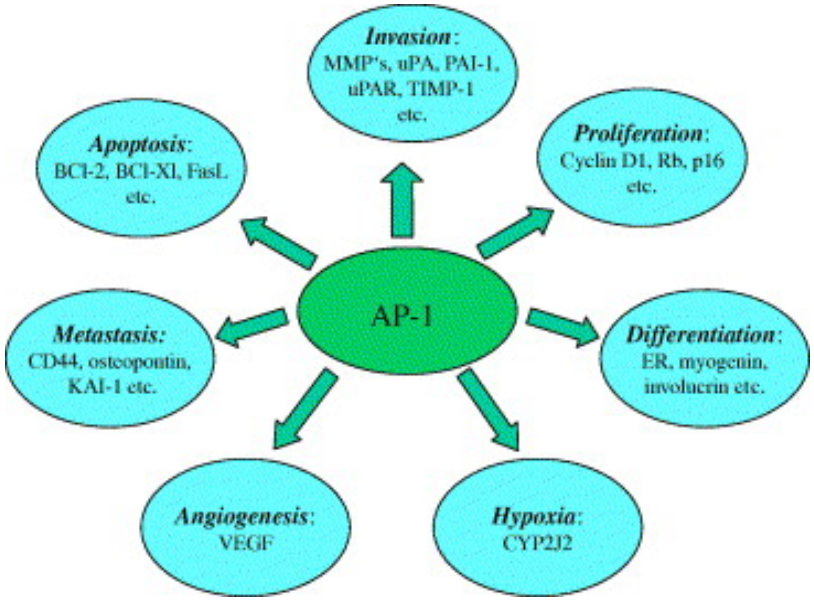

Figure 1-1. Pathways involving genes regulated by AP1 [3].

among others [4, 5]. Given its vital role in normal cellular processes it is imperative that cellular signaling be tightly regulated as failure to do so leads to pathogenesis and cancer [6]. As is the case, Jun and Fos play a significant role in cancer and was initially discovered by its elevated expression in response to the carcinogen 12-0-tetradecanolyphorbol-13-acetate (TPA) [4]. It has since been linked to breast, cervical, ovarian and prostate cancer among others [1, 7-10].

In particular, the role of AP1 in breast cancer has been extensively studied. The analysis of gene expression in tissues from breast cancer patients has indicated elevated AP1 expression [11-15]. Over expression of AP1 in MCF-7 breast cancer cell lines was seen to mimic the effect of estrogen activation in regards to the expression of estrogen induced target genes [16]. It was also shown that over expression results in increased motility and invasion of MCF-7 cells through induction of genes required for metastasis, including a known metastatic oncogene SPARC [17]. This metastatic potential was further demonstrated *in vivo* as AP1 caused weakly metastatic MCF-7 cells to metastasize to other tissues following tail injections in nude mice [7]. Complimentary to these over expression studies, a dominant negative AP1 was shown to have a significant role in preventing MCF-7 cell proliferation through the down regulation of Cyclin D1 and E2F transcription factors [8].

AP1 may also play a role in estrogen receptor positive (ER+) breast cancer resistance to tamoxifen, a common breast cancer therapeutic. Tamoxifen, a selective estrogen receptor modulator (SERM), binds to the ligand binding domain of estrogen receptor α (ERα) creating conformational changes that increase the propensity for transcriptional repressors to bind to ERα [18]. It has been documented that there is an

elevated expression profile of AP1 in ER+ tamoxifen resistant cells and further shown that ERα can interact with Jun [19, 20]. Taken together, these studies highlight the significant relationship between AP1 and cancer and mark AP1 as an excellent target for further study and ultimately drug therapy.

1.2 Regulation of transcriptional activity of Jun and Fos

Jun and Fos are considered immediate early genes, defined as being connected directly to biochemical signaling pathways and requiring no new translation products for their induction [21]. The role of immediate early genes, like Jun and Fos, is vastly important to cellular function and it is critical that they be tightly regulated – although it is of worthy note that the pathways that regulate these two proteins are markedly different. Jun is constitutively expressed at basal levels until further induced in response to various stimuli [22]. Upon activation, Jun positively regulates itself when it binds AP-1 sites within its own promoter. Jun can further maximize its own induction by dimerizing with another transcription factor, ATF-2 [22, 23]. Fos differs from Jun as it experiences stronger temporal regulation as dependent on mRNA expression and stability [21]. Transcriptional regulation of Fos is known to be under the control of transcription factors Elk-1 and CREB/ATF [24]. However, binding sites for Serum Response Element (SRE) and STAT transcription factors are also found within its promoter region [21].

Once expressed, activation of Jun and Fos occurs through a multitude of receptors, cytokines and activators including: G-Protein Coupled Receptors, TGFB, Interlukin 1 receptors, TNFA, growth factor receptors (VEGF, EGFR, ect.), serum, UV radiation, TPA, as well as oncoproteins such as v-Src and Ha-Ras [23, 25]. As one example for Jun and Fos activation, Figure 1-2 describes their up-regulation by the

growth factor pathway via the MAP kinase cascade [26, 27]. This activation leads to many of the post translational modifications that modulate Jun-Fos activity (Figure 1-3). Activation of the regulatory domains of Jun normally proceed through phosphoylation of residues 63 and 73 by the MAP kinase, Jun N-Terminal Kinase (JNK) [28, 29]. It has also been demonstrated that further activation occurs when residues 91 and 93 are phosphorylated by JNK, given conditions of prolonged stress or pro-inflammatory signals[30]. It has also been shown that phosphorylation of residue 95 may be required

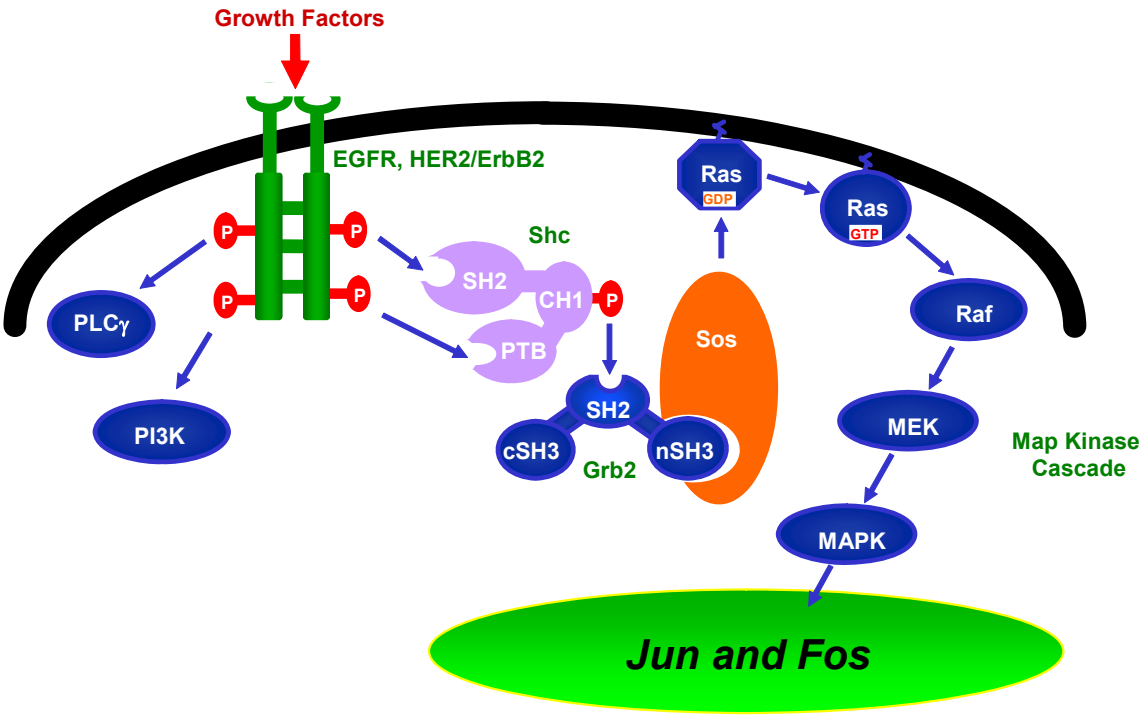

Figure 1-2. Activation of Jun and Fos through the growth factor receptor pathway. Growth factor binding to receptor tyrosine kinases induces dimerization and autophosphorylation of tyrosine residues on the cytoplasmic tails. The adaptor protein Shc can then recognize the phosphorylated tyrosines by either an SH2 domain or PTB domain. Once localized to the membrane the CH1 domain of Shc can be phosphorylated by kinases such as Src. The phosphorylated CH1 domain can then recruit another adaptor protein, Grb2, to the membrane via the SH2 domain of Grb2. Grb2 recruits Sos guanine nucleotide exchange factor, via interaction of its N-terminal SH3 domain with proline rich domains in Sos. At the membrane, Sos facilitates the GDP to GTP transfer in Ras, a small GTPase, leading to its activation and ability to activate the MAP kinase pathway. Ras induces conformational change in Raf, MAP kinase-kinase-kinase, leading to an activated form capable of phosphorylating the MAP kinase-kinase MEK. Once activated MEK can dually phosphorylate the TXY motif in the activation loops of MAP kinases, such as JNK, ERK or P38, leading to their activation. Activated MAP kinases will translocate into the nucleus where they can phosphorylate and activate transcription factors, such as Jun and Fos, enabling them to recruit transcriptional machinery to the sites of gene promoters [26, 27].

for the phosphorylation of 91 and 93 [31]. Activation of the regulatory domains of Fos necessitates phosphorylation of residue 232 by MAP kinases Extracellular Signal-Regulated Kinase (ERK) or p38 [32, 33]. Additionally, Fos activation may be modulated by phosphorylation of residues 325 and 374 by p38 [33].

It has been observed that other post translational modification events can attenuate Jun-Fos activity, including phosphorylation. Phosphorylation of Jun residues 231 and 243, by Casein Kinase II, was shown to abrogate DNA binding and it was also shown that their dephosphorylation is required for activity [34, 35]. Ubiquitinylation plays an unique role in Fos regulation as it leads to nuclear export and degradation, yet it was shown that ERK phosphorylation of residue 32 can inhibit this modification [35, 36]. Residues 229 and 257 in Jun as well as 265 in Fos are also subject to sumoylation resulting in a decrease in activity. Sumoylation was shown to be reversible and in Fos was further shown to be prevented by phosphorylation of residue 232 by ERK [37].

In conjunction with regulation of gene expression and activation, Jun and Fos can be regulated by the transcriptional complexes they form. Jun and Fos and 51 other human proteins belong to a group of transcription factors that each contain the so called basic zipper (bZIP) domain. bZIP containing proteins are further divided into families of which Jun and Fos are representative members of families containing 3 and 4 proteins, respectively [23, 38]. A comprehensive study that had analyzed the binding propensity of each bZIP containing protein to one another demonstrated Jun can interact with 13 of the 53 while Fos can interact with 9 [38]. Given the uniqueness of each member with, regards to differences in activation domains and DNA binding preferences, these findings

suggest an exceptional complexity to the Jun-Fos story. Added to this is the fact that Jun and Fos have a well documented promiscuity to interact with a multitude of non-bZIP containing transcription factors. Extensively reviewed by Chinenov & Kerrpola in 2001 [39], it has been demonstrated that Jun and Fos interact with 34 other transcription factors, including: ERα, SP1, NF-κB, Gata-2, SMAD and nFAT. The significance of these interactions are exemplified by the afore mentioned interaction of Jun with ERα and its potential role in conferring resistance to tamoxifen [19, 20]. Taken together, the regulatory mechanisms in place to modulate Jun-Fos interaction confer tight control yet give these proteins the propensity to control a wide array of cellular processes. Aberrant activity of Jun and Fos would come with great detriment and thus justifies the need for further study of these proteins.

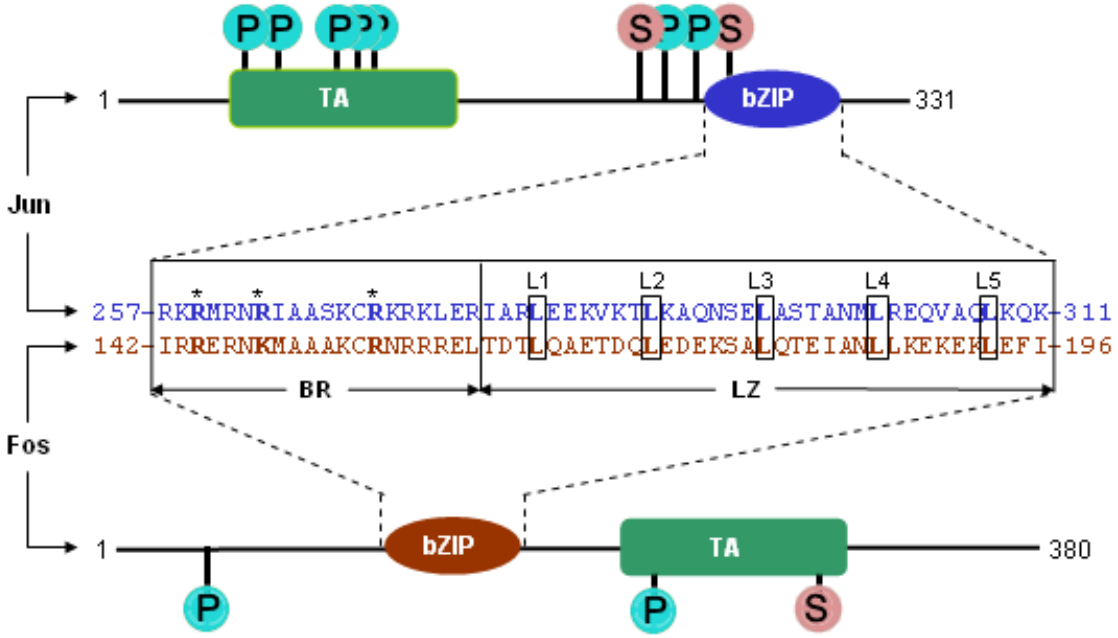

Figure 1-3. Domain organization of Jun and Fos. Jun and Fos are modular proteins that contain a transactivation domain (TA) used for the recruitment of transcriptional machinery and a basic zipper (bZIP) domain used for dimerization and DNA binding. The TA domain in Jun is N-terminal to the bZIP domain and comprised primarily of proline, glycine and acidic residues. In contrast, the TA domain of Fos is C-terminal to the bZIP domain and is primarily acidic rich [40]. Phosphorylation events are indicated by a blue P while sumoylation events are indicated by a red S.

1.3 Domain organization of Jun and Fos transcription factors

Jun and Fos are modular proteins with each containing a transactivation (TA) domain and a basic zipper (bZIP) domain (Figure 1-3). The bZIP domain, roughly 55 residues in length, is characterized by two adjacent subdomains termed the basic region (BR) at the N-terminus followed by the leucine zipper (LZ) at the C-terminus. The LZ subdomain contains a signature leucine at every seventh position within the five successive heptads of amino acid residues. The LZ adapts a continuous α-helical conformation and can induce either heterodimerization in the case of Jun and Fos, or homodimerization in the case of Jun, by virtue of their ability to wrap around each other in a coiled coil dimer [41, 42]. Such intermolecular arrangement brings the BR subdomains into close proximity, thereby enabling them to insert into the major grooves of DNA in a manner akin to a pair of forceps [43]. The 7 residues within each heptad are given the nomenclature ABCDEFG where leucines usually fill the D position, hydrophobic residues fill the A position and charged residues typically fill G and E positions [44]. This positioning allows A and D residues to form a hydrophobic face that stabilizes two zippers (Figure 1-4) through hydrophobic and Van der Waals interactions [45]. Additionally, electrostatic interactions between G and E residues not only add stability but more importantly confer specificity between bZIP containing proteins [46]. This observation is well exemplified by electrostatic repulsion between glutamates in the G and E positions of Fos, destabilizing the Fos-Fos homodimer, yet allowing for strong interaction with particular lysines in Jun in forming the Jun-Fos heterodimer [47].

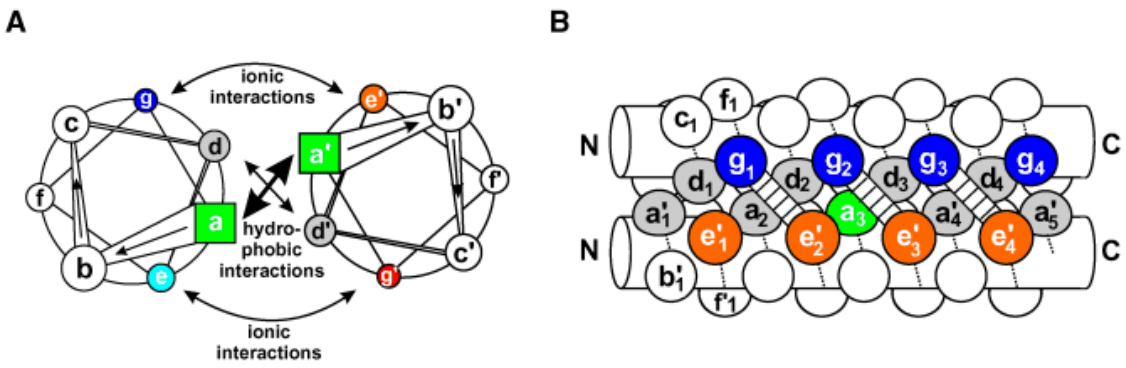

Figure 1-4. Schematic representation of a coiled-coil. This representation depicts the relationship of the residues within the heptad repeat in the context of a dimer. The helical wheel diagram in (A) looks down the axis of the α-helices from N-terminus to C-terminus, while (B) provides a side view. Hydrophilic interactions between G and E residues are indicated in (B) [44].

1.4 Jun and Fos specifically recognize TRE and CRE consensus sequences

When binding DNA, Jun and Fos specifically interact with the TPA response element (TRE) and the cAMP Response Element (CRE). The pseudo-palindromic TRE consensus sequence, TGACTCA, was initially identified and characterized in the human metallothionein IIA gene [4, 48]. The binding dissociation constant (K_d) has been determined for the binding of Jun and Fos to TRE through use of the qualitative techniques of mobility shift and pull down assays and have given dissociation constants of around 50nM [49, 50].

Current knowledge on the flexibility of Jun and Fos to bind variants of the TRE consensus sequence is largely based upon a published study that, using mobility shift assays, indicated nearly all bases are critical, with several noted exceptions [51]. The stringency of the consensus sequence was further demonstrated by the use of the yeast Jun-Fos homologue GCN4 that showed that symmetrical mutations (equivalent mutations placed on both sides of the center nucleotide) of the TRE site completely abrogate binding [52]. However, some degeneracy has been observed and many examples involve

variations in the consensus sequence when Jun and Fos interact with another protein (i.e. Jun-Fos and NFAT1 complexes bind a non-consensus TGAAACA) [39]. Interestingly, one study demonstrated that Jun-Fos bound strongest to a non-consensus TGACTAA site in a promoter that contained 4 other canonical TRE sites, suggesting the strength of interaction may also be context dependent [53].

Additionally, Jun and Fos bind to the fully palindromic CRE site, TGACGTCA, first identified in the somatostatin gene [54]. Although initially characterized as a consensus sequence for other transcription factors, evidence for Jun and Fos binding to CRE was first demonstrated by mobility shift assays that showed substantial binding for CRE yet significantly stronger binding for TRE [55]. These data were refuted by a second study, using a DNA affinity precipitation assay, that demonstrated only a slight difference between TRE and CRE binding [56]. Taken together, these studies shed light on the ability of AP1 to interact with variants of its known consensus sequences. However, despite these sites being characterized nearly 20 years ago, quantitative techniques, such as isothermal titration calorimetry (ITC), have never been used to fully characterize AP1-DNA interaction. Data from such analysis is vital to our understanding of how cis-acting elements with the DNA modulate these interaction with Jun and Fos and how this affects the range of genes and pathways under their control.

1.5 3D structures of the various AP1 complexes provide critical information about bZIP-DNA interaction

The tertiary structure of Jun-Fos heterodimer in complex with the TRE consensus sequence was published in 1995 [43]. Figure 1-5a presents this structure and shows both subunits to form continuous α-helices that position the basic regions to make specific hydrogen bonds and Van der Waals contacts with the DNA (Fig 1-5b). Interestingly, the

X-ray data suggest that Jun-Fos binds to the asymmetrical TRE sequence with no preferred orientation. These data are rationalized by the only asymmetrical contacts between Jun-Fos and the DNA bases occur between two conserved arginine residues, residue 255 in Jun and residue 155 in Fos, and the guanine nucleotide in the central base pair. Such energetically equivalent contacts allow the protein to freely exchange and thus bind in a non-oriented manner. This notion is further supported by findings from studies involving chimeric Jun-Fos proteins [57] and basic region mutations that can confer particular orientations to AP1-DNA binding [58].

Data from the Jun-Fos-TRE complexed structure also support the notion that Jun-Fos heterodimer is more stable than either Jun-Jun or Fos-Fos homodimers due to unfavorable charge repulsion between side chain residues in the G and E positions within the homodimers [59]. This notion of homodimer destabilization by G/E residues was

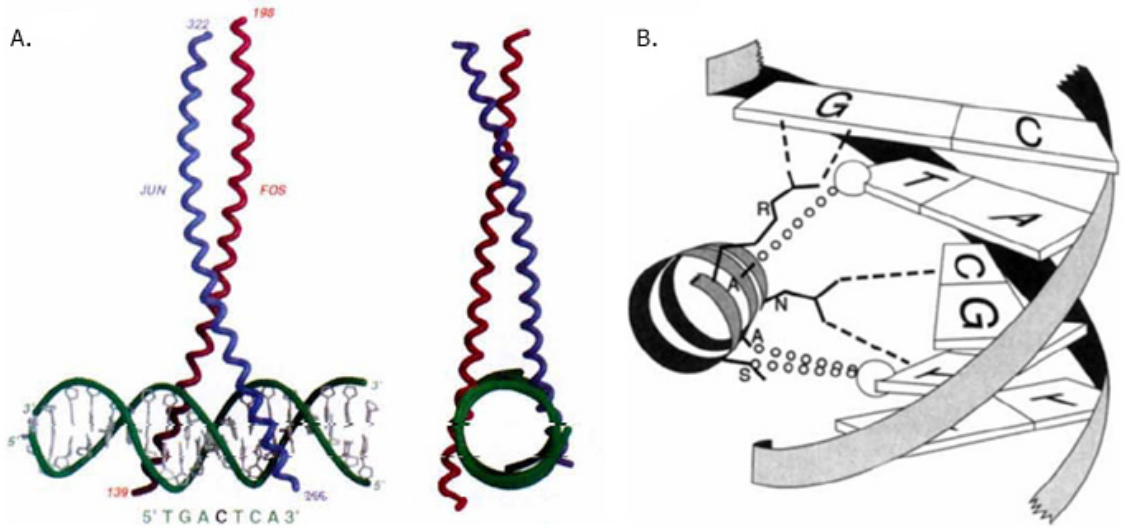

Figure 1-5. X-ray crystal structure of Jun and Fos in complex with the TRE consensus sequence. (A) Crystal structure showing the heterodimeric nature of Jun and Fos and their interaction with the TGACTCA binding site and the same complex turned 90º. (B) Diagram illustrating the conserved residues within the basic region of Jun, Fos and yeast homologue GCN4 interacting with the TGAC half site. Hydrogen bond interactions are denoted with dashed lines while Van der Waals interactions are indicated by dotted lines [43].

further shown in the 3D NMR structure of the Jun-Jun homodimer published in 1996 [60]. Here data show the ε–amino moiety of K283 in one monomer and K288 in the other are positioned 6.7Å apart, in the context of the homodimer, creating unfavorable charge repulsion and that these residues would otherwise make favorable contacts with glutamate residues in the context of the Jun-Fos heterodimer [43].

Three other 3D structures have been solved including the Jun homodimer in complex with the TRE and CRE sites (unpublished but deposited in the PDB database under the PDB ID 2H7H & 1JNM, respectively) and Jun, Fos and nFAT in complex with DNA. The latter complex further suggests that Jun-Fos interaction with other proteins may allow it to tolerate variation in the consensus sequence yet confer a preferred orientation [61]. Taken together these structural data provide essential insight into the various AP1 interactions on a molecular level providing a context for the further understanding of mechanistic data.

1.6 The bZIP domains of Jun and Fos are unstructured as monomers

Several lines of evidence support the notion that Jun and Fos are unstructured as monomers. One study utilized circular dichroism (CD) to demonstrate that the α-helical content of the bZIP domain increases from 70% in the absence of DNA to 95% with the addition of DNA containing a consensus binding site – suggesting basic regions are unstructured and fold to bind DNA [62]. Moreover, on the basis of CD and differential scanning calorimetry (DSC) performed on GCN4, it has been demonstrated that the entire bZIP domain exists as either a folded dimer or an unfolded monomer [63]. This salient observation is further supported by kinetic data that suggested bZIP domains preferentially bind as monomers to DNA and couple dimerization with DNA binding [64,

65]. Taken together, these studies add to our understanding of the underlying mechanisms that guide bZIP interaction further than what can be simply observed from 3D structural data. It is also of worthy note that few studies have examined, mechanistically, the underlying thermodynamics of AP1 interaction using techniques such as ITC.

1.7 Significance of these studies

Signaling is critical to cell survival and must be tightly regulated to avoid aberrant activity that can disrupt normal cellular processes. Stimuli, such as cytokines, hormones and growth factors, can activate signaling cascades ultimately leading to the recruitment of transcription factors to the promoters of target genes. Transcription factors, such as Jun and Fos, modulate important pathways and may be responsible as the cause of various cancers [6, 8]. Jun and Fos are modular proteins that contain two functional domains. The first, a transactivation domain, assists in recruitment of the RNA polymerase complex [23]. The second, the bZIP domain enables heterodimerization, in the case of Jun-Fos heterodimer, or homodimerization, in the case of Jun-Jun homodimer, via the formation of leucine zippers. Additionally, the bZIP domain contains a region of basic residues that enables interaction with DNA [42]. Jun and Fos recognize the TRE and CRE consensus sites, found in the promoters of many genes including metallothionein IIa, collagenase, interleukin 2 and cyclin D1 [5]. Given its significant and varied roles in cellular pathways and importance to cancer, Jun and Fos merit further study. In particular, our understanding of AP1-DNA interaction in biophysical terms is severely lacking. Although crystal structures of AP1-DNA complexes are available, the knowledge of underlying “invisible” thermodynamic forces that drive the binding of AP1 to DNA remains elusive. In an effort to elucidate the role of such invisible forces driving the

formation of AP1-DNA complexes, this dissertation will cover an extensive thermodynamic and structural analysis of AP1-DNA interaction.

2 Chapter 2: Materials and Methods

2.1 Molecular cloning

bZIP domain of human Jun (residues 251-331), bZIP domain of human Fos (residues 136-216), LZ domain of human Jun (residues 277-331) and LZ domain of human Fos (residues 162-216) were cloned into the pET102 bacterial expression vector – using Invitrogen TOPO technology following manufacturer's protocol. The pET102 vector features an N-terminal thioredoxin (Trx)-tag and a C-terminal polyhistidine (His)-tag. The Trx-tag markedly improves solubility of fusion proteins. The His-tag was used to aid in protein purification using Ni-NTA affinity chromatography. Thrombin cleavage sites (LVPRGS) were introduced at both the N- and C-termini of the proteins to allow the removal of Trx and His-tags after protein purification, should this become necessary.

2.2 Protein expression and purification

Escherichia coli Rosetta2(DE3) or BL21*(DE3) strains (Novagen) were transformed with pET102 vectors containing various constructs of Jun and Fos. Cells were cultured in LB media and grown at 20ºC to an optical density of 0.5 at 600nm prior to induction with 0.5 mM isopropyl ß-D-1-thiogalactopyranoside (IPTG). Cells were allowed to express overnight at 20ºC and were subsequently harvested and resuspended in Lysis Buffer (50 mM Tris, 500mM NaCl, 2M Urea, 2mM ß-mercaptoethanol (ß-ME), 10% Triton X-100 at pH 8.0). Cells were then disrupted using a Biospec Bead-Beater® and subjected to high speed centrifugation to remove cell debris. Cell lysate was then applied to a Ni-NTA affinity chromatography column. The column was washed with Wash Buffer (50 mM Tris, 500 mM NaCl, 2M Urea, 20 mM Imidizole, 2mM ß-ME at pH 8.0) to remove non-specific binding of bacterial proteins. Protein was then eluted

using Elution Buffer (50 mM Tris, 500 mM NaCl, 2M Urea, 500 mM Imidizole, 2mM ß-ME at pH 8.0) and dialyzed in SEC Buffer (50 mM Tris, 200 mM NaCl, 1 mM EDTA, 5 mM ß-ME at pH 8.0) overnight. Proteins were further purified on a HiLoad 26/60 Superdex 200 Prep Grade size exclusion chromatography (SEC) column coupled to a GE Akta FPLC system. Protein was purified to apparent homogeneity as judged by SDS-PAGE analysis (Fig 2-1A). The treatment of all proteins with thrombin protease significantly destabilized them and all proteins appear to be proteolytically unstable. Protein concentrations for all proteins were determined by the fluorescence-based Quant-It assay (Invitrogen) and spectrophotometrically using extinction co-efficients calculated from the online software ProtParam at ExPasy Server [66].

2.3 SDS-PAGE analysis

SDS-PAGE is a widely used technique to separate proteins according to size [67]. SDS is an anionic detergent that denatures proteins and coats them uniformly with a negative charge proportional to the mass of the protein. During PAGE, the uniform negative charge causes proteins to be pulled with the same force towards the gel apparatus cathode. Separation according to their size occurs as larger proteins have greater difficulty traveling through the gel matrix. The size of proteins can then be estimated using a standardized set of protein markers of known molecular weight.

For the SDS-PAGE analysis of recombinant proteins (Figure 2-1), gels were run between 120-150 V for up to 1 hour using a VWR AccuPower power supply and a Bio-Rad Protean Cell. Gels were visualized by staining with coomassie-blue stain and destaining with a 10% acetic acid, 10% methanol solution. Images of the gels were captured using a UVP MultiDoc-It Gel Imaging System.

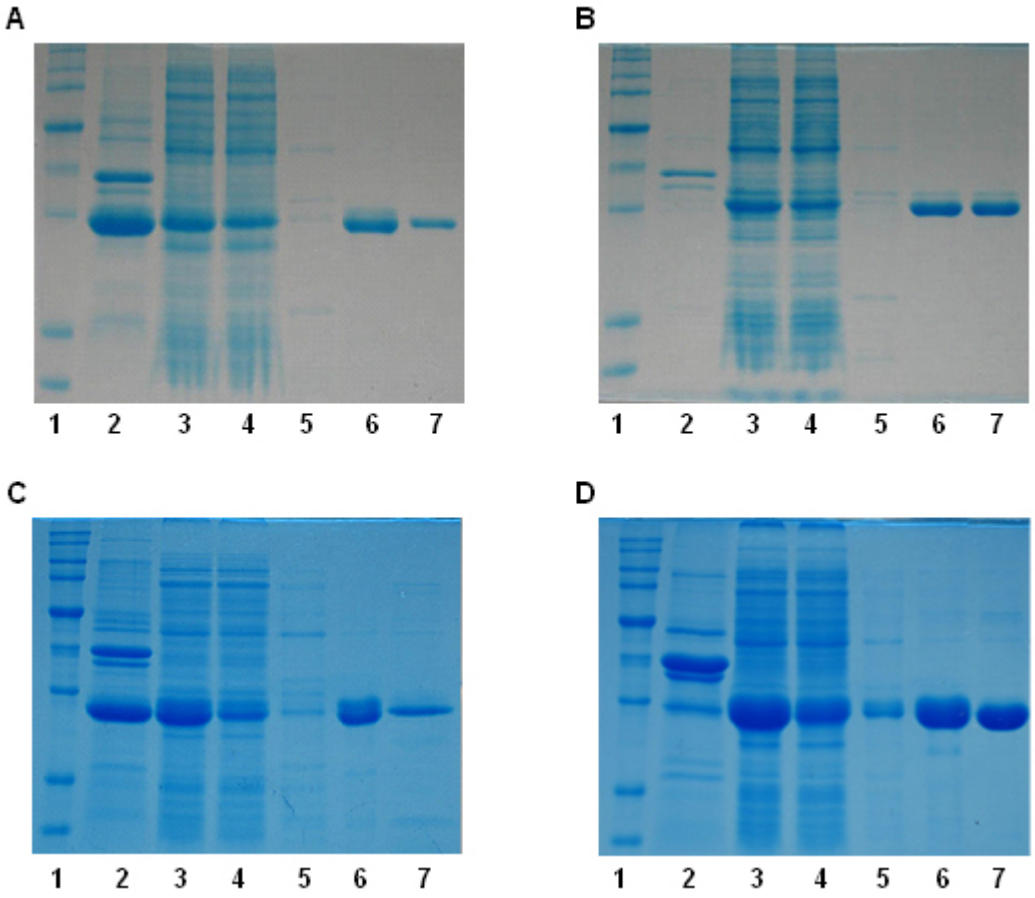

Figure 2-1. SDS-PAGE analysis of the Ni-NTA purification of recombinant proteins. bZIP domain of human Jun (A), bZIP domain of human Fos (B), LZ domain of human Jun (C), LZ domain of human Fos (D) were expressed in E. Coli Rosetta2 or BL21* strains at 20ºC to an optical density of 0.5 at 600nm and induced with 0.5 mM IPTG overnight. Cells were then harvested and disrupted using a Biospec Bead-Beater. After disruption cell debris was separated from the soluble lysate fraction using high speed centrifugation. A portion of the cell pellet was solubilized in 10% SDS to determine whether protein went to an inclusion body (lane 2). The lysate (lane 3) was subjected to a Ni-NTA chromatography column. Proteins that flowed through and were not retained by the Ni column were sampled (lane 4). Column was then washed with 20 mM imidazole to remove non-specific binding of bacterial proteins (lane 5). Next, the recombinant proteins were eluted with 500mM imidazole (lane 6) and dialyzed into appropriate physiological buffers (lane 7). Protein markers (Promega) were added for reference (lane 1).

2.4 Oligo annealing

DNA oligos for the TRE, CRE and single nucleotide variants of TRE were commercially obtained from Sigma Genosys. The complete nucleotide sequences of these oligos are presented in corresponding chapters. Oligo concentrations were determined spectrophotometrically on the basis of the extinction co-efficients derived from the nucleotide sequences using the online software OligoAnalyzer 3.0 (Integrated DNA

Technologies) based on the nearest-neighbor model [68]. To obtain double-stranded DNA (dsDNA) annealed oligos, equimolar amounts of sense and antisense oligos were mixed together and heated at 95^{o}C for 10 min and then slowly cooled to room temperature, using a Techne TC-312 thermocycler. The efficiency of oligo annealing to generate dsDNA was close to 100% as judged by Native-PAGE and SEC.

2.5 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a powerful technique capable of direct measurement of the heat associated with a given process, which, under isobaric conditions, is equal to the enthalpy change in the process [68]. Through the titration of a protein with a ligand, an ITC experiment can directly measure enthalpy (ΔH). Further analysis allows calculation of affinity (K_d), Gibbs free energy (ΔG), entropy (ΔS), and stoichiometry (n). ITC can be further utilized to unlock key mechanistic data through determination of heat capacity change (ΔC_P), change in solvent accessible surface area (ΔSASA) [69-73], counter-ion uptake [74] and protonation/deprotonation events [75]. Given its range of capabilities the ITC undoubtedly proves its value for biophysical applications. Specific details regarding ITC experiments contained within this dissertation can be found in the experimental procedures section of corresponding chapters.

2.6 SASA calculations

The magnitude of changes in polar and apolar solvent-accessible surface area (SASA) in the bZIP domains of Jun-Fos heterodimer upon binding to dsDNA oligos containing the TRE and CRE consensus sites were calculated from thermodynamic data obtained using ITC and compared with those obtained from structural data based on the

3D structural models. The details of this procedure can be found in the experimental procedures section of corresponding chapters.

2.7 Structural modeling

3D structures of the various AP1-DNA complexes were modeled using the MODELLER software based on homology modeling [76]. MODELLER employs molecular dynamics and simulated annealing protocols to optimize the modeled structure through satisfaction of spatial restraints derived from amino acid sequence alignment with a corresponding template in Cartesian space. Modeled structures should be expected to adopt 3D folds similar to the template structure except for sidechain conformations of specific amino acids due to the introduction of specific hydrogen bonding, the rearrangement of domains and DNA spatially to one-another or the modeling of loops not rendered in template structures. Such hydrogen bonding restraints being introduced herein were necessary to bring the sidechain atoms of respective residues within optimal hydrogen bonding distance in agreement with our thermodynamic data reported. Atomic distances set for hydrogen bonding restraints between specific pairs of oxygen and nitrogen atoms were 2.8±0.5Å. Thus, MODELLER will force the sidechain oxygen and nitrogen atoms of specific hydrogen bonding partners to lie within approximately 2.8Å of each other through the rotation of backbone N-Cα and Cα-C' bonds with little effect on the overall global fold. Additionally, unfolded regions were allowed to adopt an open compact conformation and allowed to reach the energy minima without any restraints. In each case, a total of 100 structural models were calculated and the structure with the lowest energy, as judged by the MODELLER Objective Function, was selected for further energy minimization in MODELLER prior to analysis. The structures were

rendered using RIBBONS [77]. Specific modifications made within each model are detailed in the experimental procedures section of corresponding chapters.

3 Chapter 3: Coupling of Folding and DNA-Binding in the bZIP Domains of Jun-Fos Heterodimeric Transcription Factor

3.1 Summary

In response to mitogenic stimuli, the heterodimeric transcription factor Jun-Fos binds to the promoters of a diverse array of genes involved in critical cellular responses such as cell growth and proliferation, cell cycle regulation, embryogenic development and cancer. In so doing, Jun-Fos heterodimer regulates gene expression central to physiology and pathology of the cell in a specific and timely manner. Here, using the technique of isothermal titration calorimetry (ITC), we report detailed thermodynamics of the bZIP domains of Jun-Fos heterodimer to synthetic dsDNA oligos containing the TRE and CRE consensus promoter elements. Our data suggest that binding of the bZIP domains to both TRE and CRE is under enthalpic control and accompanied by entropic penalty at physiological temperatures. Although the bZIP domains bind to both TRE and CRE with very similar affinities, the enthalpic contributions to the free energy of binding to CRE are more favorable than TRE, while the entropic penalty to the free energy of binding to TRE is smaller than CRE. Despite such differences in their thermodynamic signatures, enthalpy and entropy of binding of the bZIP domains to both TRE and CRE are highly temperature-dependent and largely compensate each other resulting in negligible effect of temperature on the free energy of binding. From the plot of enthalpy change versus temperature, the magnitude of heat capacity change determined is much larger than that expected from the direct association of bZIP domains with DNA. This observation is interpreted to suggest that the basic regions in the bZIP domains are largely unstructured in the absence of DNA and only become structured upon interaction with DNA in a coupled folding and binding manner. Our new findings are rationalized in

the context of 3D structural models of bZIP domains of Jun-Fos heterodimer in complex with dsDNA oligos containing the TRE and CRE consensus sequences. Taken together, our study demonstrates that enthalpy is the major driving force for a key protein-DNA interaction pertinent to cellular signaling and that protein-DNA interactions with similar binding affinities may be accompanied by differential thermodynamic signatures. Our data corroborate the notion that the DNA-induced protein structural changes are a general feature of the bZIP family of transcription factors.

3.2 Background

Protein-DNA interactions play a critical role in coupling extracellular information in the form of growth factors, cytokines, hormones and stress to DNA transcription and, in so doing, regulate a diverse array of cellular processes such as cell growth and proliferation, cell cycle regulation, embryogenic development and cancer. Discovered as components of the transcription factor AP1, Jun and Fos recognize — as Jun-Jun homodimer or Jun-Fos heterodimer — the pseudo-palindromic TGACTCA and palindromic TGACGTCA consensus sequences found in the promoters of a multitude of genes such as metallothionein IIa, collagenase, interleukin 2 and cyclin D1 [4, 41, 48, 56, 78-81]. The consensus sequences TGACTCA and TGACGTCA, respectively referred to as the TPA (12-O-tetradecanoylphorbol-13-acetate) response element (TRE) and the cAMP response element (CRE), occur with a high frequency in the human genome [39, 82]. Jun and Fos are expressed in a wide variety of tissues and are subject to activation by a diverse array of mitogenic inputs, including up-regulation by MAP kinases [26, 27]. Upon activation, Jun and Fos can switch on gene transcription via their direct involvement and through their co-operation with other transcription factors in the

recruitment of the transcriptional machinery to the site of DNA [3, 5, 39, 83, 84]. Jun and Fos are potent activators of mitogenic transcription and, as such, their hyperactivity is positively correlated with oncogenic transformations of cells [3, 5, 84]. To combat such undesirable effects, the activity of Jun and Fos is tightly regulated at various levels, including gene expression, post-translational phosphorylation, and interaction with other cellular proteins [5].

Jun and Fos are modular proteins and contain regions with conserved leucine residues at every seventh position that enable them to form coiled coils termed leucine zippers [41, 42]. Located N-terminal to leucine zippers in both proteins are clusters of basic residues that together with leucine zippers constitute what have come to be known as the basic zipper (bZIP) domains of Jun and Fos (Figure 3-1a). The bZIP domains enable the recruitment of Jun and Fos to the site of transcription by virtue of their ability to recognize the TRE and CRE consensus sequences in DNA at the promoters of the target genes [5, 41]. Once recruited to the site of transcription, Jun and Fos unleash their full transactivation potential and participate in the transcriptional machinery through regions that lie outside the bZIP domains [5, 85]. While Jun can homodimerize with itself or heterodimerize with Fos via the formation of leucine zippers to form transcriptionally-active species, the existence of Fos as a homodimer has never been observed under physiological conditions and Fos alone does not possess any transcriptional activity [42].

The availability of 3D structure of the bZIP domains of Jun-Fos heterodimer in complex with dsDNA oligo containing the TRE site has significantly contributed to our understanding of the molecular mechanism of action of this transcription factor at

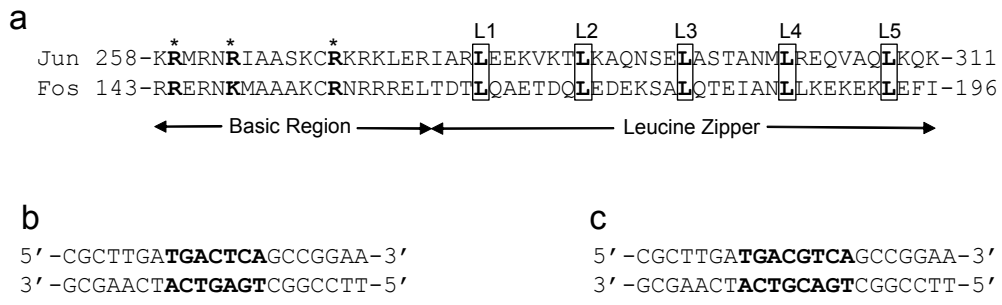

Figure 3-1. Protein and DNA sequences. (a) Amino acid sequence alignment of the leucine zipper and basic region in human Jun and Fos that dimerize to form the bZIP domains. The five signature leucines (L1-L5) characteristic of leucine zipper, spaced exactly six residues apart, are boxed and bold faced. The basic residues in the basic region that contact the DNA bases and the backbone phosphates are marked by asterisks and bold faced. (b) Nucleotide sequence of 21-mer dsDNA oligo containing the TRE site (bold faced). (c) Nucleotide sequence of 22-mer dsDNA oligo containing the CRE site (bold faced).

a structural level [43]. In this structure, the heterodimeric bZIP domains adopt continuous α-helical conformations of about 15 turns and wrap around each other like a pair of forceps that inserts into the major grooves of DNA via the N-terminal basic regions. While the α-helices are held together by numerous inter-helical hydrophobic contacts and salt bridges, hydrogen bonding between the sidechains of basic residues in the bZIP domains and the sidechains of nucleotides accounts for high affinity binding of bZIP domains to DNA. Further knowledge of how bZIP domains interact with their cognate sequences in DNA comes from kinetic studies that suggest that although bZIP domains of Jun and Fos can heterodimerize rapidly in the absence of DNA, the rate of heterodimerization is significantly enhanced in the presence of DNA and that the pathway in which the bZIP monomers associate with DNA prior to heterodimerization appears to be highly favored on kinetic grounds [64].

Despite such wealth of structural and kinetic data, little is known about the thermodynamic mechanism of the binding of bZIP domains of Fos and Jun to DNA. Here, using the technique of isothermal titration calorimetry (ITC), we report detailed

thermodynamics of the bZIP domains of Jun-Fos heterodimer to synthetic dsDNA oligos containing the TRE and CRE consensus promoter elements. Our data suggest that binding of the bZIP domains to both TRE and CRE is under enthalpic control and accompanied by entropic penalty at physiological temperatures. Although the bZIP domains bind to both TRE and CRE with very similar affinities, the enthalpic contributions to the free energy of binding to CRE are more favorable than TRE, while the entropic penalty to the free energy of binding to TRE is smaller than CRE. Despite such differences in their thermodynamic signatures, enthalpy and entropy of binding of the bZIP domains to both TRE and CRE are highly temperature-dependent and largely compensate each other resulting in negligible effect of temperature on the free energy of binding. From the plot of enthalpy change versus temperature, the magnitude of heat capacity change determined is much larger than that expected from the direct association of bZIP domains with DNA. This observation is interpreted to suggest that the basic regions in the bZIP domains are largely unstructured in the absence of DNA and only become structured upon interaction with DNA in a coupled folding and binding manner. Our new findings are rationalized in the context of 3D structural models of bZIP domains of Jun-Fos heterodimer in complex with dsDNA oligos containing the TRE and CRE consensus sequences. Taken together, our study demonstrates that enthalpy is the major driving force for a key protein-DNA interaction pertinent to cellular signaling and that protein-DNA interactions with similar binding affinities may be accompanied by differential thermodynamic signatures. Our data corroborate the notion that the DNA-induced protein structural changes are a general feature of the bZIP family of transcription factors.

3.3 Experimental procedures

3.3.1 Protein preparation

bZIP domains of human Jun (residues 251-331) and human Fos (residues 136-216) were cloned into pET102 bacterial expression vector — with an N-terminal thioredoxin (Trx)-tag and a C-terminal polyhistidine (His)-tag — using Invitrogen TOPO technology. Trx-tag was included to maximize protein expression in soluble fraction, while the His-tag was added to aid in protein purification through Ni-NTA affinity chromatography. Additionally, thrombin protease sites were introduced at both the N- and C-termini of the proteins to aid in the removal of tags after purification. Proteins were subsequently expressed in Escherichia coli Rosetta2(DE3) bacterial strain (Novagen) cultured in LB media and purified on Ni-NTA affinity column using standard procedures. Briefly, bacterial cells were grown at 20°C to an optical density of 0.5 at 600nm prior to induction with 0.5mM isopropyl β-D-1-thiogalactopyranoside (IPTG). After further overnight growth at 15°C, the cells were harvested and disrupted using a beadbeater. After separation of cell debris at high speed centrifugation, the cell lysate was subjected to Ni-NTA column and washed extensively with low concentrations of imidazole to remove non-specific binding of bacterial proteins to the column. The recombinant bZIP domains of Jun and Fos were subsequently eluted with 500mM imidazole and dialyzed against an appropriate buffer to remove imidazole. Further treatment of bZIP domains of Jun and Fos on MonoQ ion-exchange column coupled to GE Akta FPLC system led to purification of recombinant domains to apparent homogeneity as judged by SDS-PAGE analysis. The identity of recombinant proteins was confirmed by MALDI-TOF mass spectrometry analysis. Final yields were typically

between 10-20mg protein of apparent homogeneity per liter of bacterial culture. The treatment of recombinant proteins with thrombin protease significantly destabilized the bZIP domains of both Jun and Fos and both domains appeared to be proteolytically unstable. For this reason, all experiments reported herein were carried out on recombinant fusion bZIP domains of Jun and Fos containing a Trx-tag at the N-terminus and a His-tag at the C-terminus. The tags were found to have no effect on the binding of these domains to DNA under all conditions used here. Protein concentrations were determined by the fluorescence-based Quant-It assay (Invitrogen) and spectrophotometrically using extinction co-efficients of 14,230$M^{-1}cm^{-1}$ and 14,230$M^{-1}cm^{-1}$ at 280nm for the bZIP recombinant domains of Jun and Fos, respectively. The extinction co-efficients were calculated using the online software ProtParam at ExPasy Server [86]. Results from both methods were in an excellent agreement. Jun-Fos bZIP heterodimers were generated by mixing equimolar amounts of the purified bZIP domains of Jun and Fos. The efficiency of bZIP heterodimerization was close to 100% as judged by Native-PAGE and size exclusion chromatography (SEC) analysis using a Hiload Superdex 200 column.

3.3.2 DNA synthesis

HPLC-grade DNA oligos containing the consensus TRE and CRE sites were commercially obtained from Sigma Genosys. The complete nucleotide sequences of these oligos are presented in Figures 3-1b and 3-1c. Oligo concentrations were determined spectrophotometrically on the basis of their extinction co-efficients derived from their nucleotide sequences using the online software OligoAnalyzer 3.0 (Integrated DNA Technologies) based on the nearest-neighbor model [68]. To obtain double-stranded DNA (dsDNA) annealed oligos, equimolar amounts of sense and anti-sense oligos were

mixed together and heated at 95°C for 10min and then allowed to cool to room temperature. The efficiency of oligo annealing to generate dsDNA was close to 100% as judged by Native-PAGE and size exclusion chromatography (SEC) analysis using a Hiload Superdex 200 column.

3.3.3 ITC measurements

Isothermal titration calorimetry (ITC) experiments were performed on Microcal VP-ITC instrument and data were acquired and processed using fully automized features in Microcal ORIGIN software. All measurements were repeated 2-3 times. Briefly, the protein and DNA samples were prepared in 50mM Tris, 200mM NaCl, 5mM EDTA and 5mM β-mercaptoethanol at pH 8.0 and de-gassed using the ThermoVac accessory for 10min. The experiments were initiated by injecting 25 x 10μl injections of 100-200μM of dsDNA oligo from the syringe into the calorimetric cell containing 1.8ml of 5-10μM of Jun-Fos heterodimer at a fixed temperature in the narrow range 15-35°C. The bZIP domains of Jun and Fos form heterodimers with an affinity of less than 0.1μM [64]. Thus, under the ITC conditions, the bZIP domains of Jun and Fos would be expected to predominantly exist as heterodimers with negligible amounts of monomers. The change in thermal power as a function of each injection was automatically recorded using Microcal ORIGIN software and the raw data were further processed to yield binding isotherms of heat release per injection as a function of DNA to protein molar ratio. The heats of mixing and dilution were subtracted from the heat of binding per injection by carrying out a control experiment in which the same buffer in the calorimetric cell was titrated against the dsDNA oligos in an identical manner. Control experiments with scrambled dsDNA oligos generated similar thermal power to that obtained for the buffer

alone — as did the titration of dsDNA oligos containing TRE and CRE sites against a protein construct containing thioredoxin with a C-terminal His-tag (Trx-His). Titration of concentrated Trx-His protein construct into the calorimetric cell containing the bZIP domains of Jun-Fos heterodimer produced no observable signal, implying that neither Trx-tag nor His-tag interact with the bZIP domains of Jun and Fos. To extract various thermodynamic parameters, the binding isotherms were iteratively fit to the following in-built function by non-linear least squares regression analysis using the integrated Microcal ORIGIN software:

$$q(i) = (n\Delta HVP/2)\ \{[1+(L/nP)+(K_d/nP)] - [[1+(L/nP)+(K_d/nP)]^2 - (4L/nP)]^{1/2}\} \qquad [1]$$

where q(i) is the heat release (kcal/mol) for the ith injection, n is the binding stoichiometry, ΔH is the binding enthalpy (kcal/mol), V is the effective volume of protein solution in the calorimetric cell (1.46ml), P is the total protein concentration in the calorimetric cell (μM), L is the total concentration of DNA added (μM) and K_d is the apparent equilibrium binding constant (μM). The above equation is derived from the binding of a ligand to a macromolecule using the law of mass action assuming a 1:1 binding stoichiometry [74]. The iterative fit of binding isotherms to the above in-built function thus directly generated values for K_d and ΔH. The free energy change (ΔG) upon ligand binding was calculated from the relationship:

$$\Delta G = RT\ln K_d \qquad [2]$$

where R is the universal molar gas constant (1.99 cal/K/mol), T is the absolute temperature in Kelvins and K_d is in the units of mol/L. The entropic contribution ($T\Delta S$) to the free energy of binding was calculated from the relationship:

$$T\Delta S = \Delta H - \Delta G \qquad [3]$$

where ΔH and ΔG are as defined above. Heat capacity change (ΔC_p) in the protein upon DNA binding was calculated by measuring ΔH as a function of temperature (T) in the narrow range 15-35°C — with the slope of the ΔH versus T plot yielding the value of ΔC_p. To improve the accuracy of ITC measurements, the c value was controlled in the approximate range 10-200. The c-value is a dimensionless parameter and defined by the ratio of total protein concentration in the calorimetric cell divided by K_d. Measurements with the c-value outside the range 5-500 are subject to error. Despite the stability conferred upon the bZIP domains of Jun and Fos by the presence of Trx- and His-tags, only the recombinant Jun-Fos heterodimer was stable at protein concentrations above 5μM needed for ITC experiments, while the recombinant Jun-Jun homodimer lacked the stability for any reliable ITC measurements. For this reason, only ITC measurements on the binding of Jun-Fos heterodimer to DNA are reported in this study.

3.3.4 SASA calculations

The magnitude of changes in polar and apolar solvent-accessible surface area (SASA) in the bZIP domains of Jun-Fos heterodimer upon binding to dsDNA oligos containing the TRE and CRE consensus sites were calculated from thermodynamic data obtained using ITC and compared with those obtained from structural data based on the 3D structural models (see below).

For calculation of changes in polar SASA ($\Delta SASA_{polar}$) and apolar SASA ($\Delta SASA_{apolar}$) upon the binding of dsDNA oligos containing the TRE and CRE consensus sites to bZIP domains of Jun-Fos heterodimer from thermodynamic data, it was assumed that ΔC_p and ΔH at 60°C (ΔH_{60}) are additive and linearly depend on the change in

$\Delta SASA_{polar}$ and $\Delta SASA_{apolar}$ as embodied in the following empirically-derived expressions [69-73]:

$$\Delta C_p = a[\Delta SASA_{polar}] + b[\Delta SASA_{apolar}] \quad [4]$$

$$\Delta H_{60} = c[\Delta SASA_{polar}] + d[\Delta SASA_{apolar}] \quad [5]$$

where a, b, c and d are empirically-determined co-efficients with values of –0.26 cal/mol/K/Å^2, +0.45 cal/mol/K/Å^2, +31.34 cal/mol/Å^2 and –8.44 cal/mol/Å^2, respectively. The co-efficients a and b are independent of temperature, while c and d are referenced against a temperature of 60°C, which equates to the median melting temperature of the proteins from which these constants are derived [69-71]. ΔC_p was calculated from the slope of a plot of ΔH versus T in the narrow temperature range 15-35°C for the binding of TRE (-0.87 kcal/mol/K) and CRE (-0.81 kcal/mol/K) dsDNA oligos to the bZIP domains of Jun-Fos heterodimer using the ITC instrument. ΔH_{60} was calculated by the extrapolation of a plot of ΔH versus T to 60°C for the binding of TRE (-60.56 kcal/mol) and CRE (-64.95 kcal/mol) dsDNA oligos to the bZIP domains of Jun-Fos heterodimer using the ITC instrument. With ΔC_p and ΔH_{60} experimentally determined using ITC and the knowledge of co-efficients a-d from empirical models [69-73], equations [4] and [5] were simultaneously solved to obtain the magnitude of changes in $\Delta SASA_{polar}$ and $\Delta SASA_{apolar}$ independent of structural information upon the binding of dsDNA oligos to the bZIP domains of Jun-Fos heterodimer.

To determine changes in $\Delta SASA_{polar}$ and $\Delta SASA_{apolar}$ upon the binding of dsDNA oligos to bZIP domains of Jun-Fos heterodimer from structural data, two models of binding were assumed — the Rigid Body model and the Induced Fit model. In the Rigid Body model, it was assumed that the bZIP domains of Jun-Fos heterodimer undergo no

conformational change upon binding to dsDNA oligos containing TRE and CRE sites. In the Induced Fit model, it was assumed that the basic regions in the bZIP domains of Jun-Fos heterodimer are fully unstructured and only become structured upon binding to dsDNA oligos containing TRE and CRE sites. Changes in $\Delta SASA_{polar}$ and $\Delta SASA_{apolar}$ upon the binding of dsDNA oligos to bZIP domains of Jun-Fos heterodimer from structural data were calculated using the following relationships:

$$\Delta SASA_{polar} = SASA_{bp} - (SASA_{fp} + SASA_{dp}) \quad [6]$$

$$\Delta SASA_{apolar} = SASA_{ba} - (SASA_{fa} + SASA_{da}) \quad [7]$$

where $SASA_{bp}$ and $SASA_{ba}$ are the polar and apolar SASA of bZIP domains of Jun-Fos heterodimer bound to DNA, $SASA_{fp}$ and $SASA_{fa}$ are the polar and apolar SASA of free bZIP domains of Jun-Fos heterodimer, and $SASA_{dp}$ and $SASA_{da}$ are the polar and apolar SASA of free dsDNA oligos. $SASA_{bp}$ and $SASA_{ba}$ were calculated from structural models of bZIP domains of Jun-Fos heterodimer in complex with dsDNA oligos (containing atomic coordinates of both the bZIP domains and the corresponding sense and antisense dsDNA oligos), $SASA_{fp}$ and $SASA_{fa}$ were calculated from structural models of bZIP domains of Jun-Fos heterodimer in complex with dsDNA oligos (but containing atomic coordinates of the bZIP domains only) for the Rigid Body model, while $SASA_{fp}$ and $SASA_{fa}$ were calculated from structural models of bZIP domains of Jun-Fos heterodimer determined in the absence of DNA with basic regions allowed to adopt extended conformations (see below) in the Induced Fit model, and $SASA_{dp}$ and $SASA_{da}$ were calculated from structural models of bZIP domains of Jun-Fos heterodimer in complex with dsDNA oligos (but containing atomic coordinates of only the

corresponding sense and antisense dsDNA oligos). All SASA calculations were performed using the online software GETAREA with a probe radius of 1.4Å [87].

3.3.5 Structural modeling

3D structures of bZIP domains of Jun-Fos heterodimer alone and in complex with dsDNA oligos containing TRE and CRE sites were modeled using the MODELLER software based on homology modeling [76]. The model of bZIP domains of Jun-Fos heterodimer in the absence of DNA was obtained using the crystal structure of bZIP domains of Jun-Fos heterodimer in complex with a dsDNA oligo containing the TRE consensus sequence as a template [PDB code: 1FOS]. However, only the structural coordinates of atoms within leucine zipper regions were used to model the free structure of the bZIP domains of Jun-Fos heterodimer, while the residues in the N-terminal basic regions were allowed to adopt an open extended conformation and allowed to reach the energy minima without any restraints supplied by corresponding residues in the template. The model of bZIP domains of Jun-Fos heterodimer in complex with 21-mer dsDNA oligo containing the TRE site was obtained using the crystal structure of bZIP domains of Jun-Fos heterodimer in complex with a dsDNA oligo containing the TRE consensus sequence TGACTCA but differing in flanking sequences as a template [PDB code: 1FOS]. The model of bZIP domains of Jun-Fos heterodimer in complex with 22-mer dsDNA oligo containing the CRE site was obtained using the crystal structure of bZIP domains of Jun-Jun homodimer in complex with a dsDNA oligo containing the CRE consensus sequence TGACGTCA but differing in flanking sequences as a template [PDB code: 1JNM]. In each case, a total of 100 structural models were calculated and the structure with the lowest energy, as judged by the MODELLER Objective Function, was

selected for further energy minimization in MODELLER prior to analysis. The structures were rendered using RIBBONS [77] and superimposed in MOLMOL [88]. All other calculations were performed on the lowest energy-minimized structural model.

3.4 Results and discussion

3.4.1 Enthalpy drives the protein-DNA interaction

In an attempt to unravel the thermodynamic mechanism of the binding of bZIP domains of Jun-Fos heterodimer to dsDNA oligos containing the TRE and CRE sites, we employed the powerful technique of ITC (Figure 3-2). Comparison of the various thermodynamic parameters is presented in Table 3-1. Our data suggest that the bZIP domains of Jun-Fos heterodimer bind to TRE and CRE sites with very similar affinities of 0.15μM and 0.21μM, respectively (Table 3-1). These observations are in an excellent agreement with previous studies based on semi-quantitative analysis [42, 49, 56, 57, 89]. What sets our study apart from anything previously reported is the striking observation that binding of the bZIP domains of Jun-Fos heterodimer to TRE and CRE sites in DNA is under strong enthalpic control accompanied by entropic penalty at physiological temperatures. This observation is consistent with enthalpic-driven nature of the interaction of the bZIP domains of the yeast transcription factor GCN4 to TRE and CRE sites in DNA [90, 91].

The favorable enthalpic change is most likely due to the formation of hydrogen bonding, hydrophobic contacts and electrostatic interactions between the bZIP domains structure [43]. Interestingly, despite the slightly weaker interaction of the bZIP domains of Jun-Fos heterodimer to CRE site relative to TRE site, the binding to CRE site appears

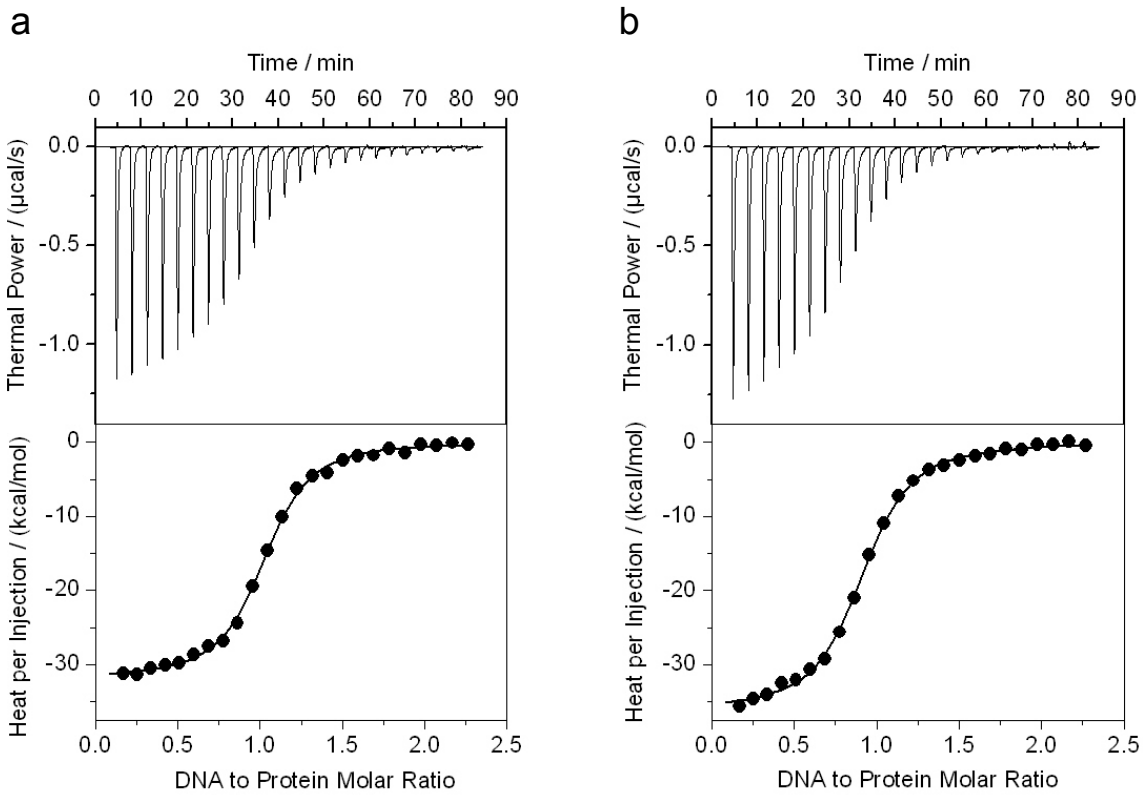

Figure 3-2. ITC analysis of the binding of the bZIP domains of Jun-Fos heterodimer to dsDNA oligos containing TRE (a) and CRE (b) sites. bZIP domains of Jun-Fos heterodimer in the calorimetric cell were titrated with 25 x 10µl injections of dsDNA oligo from the injection syringe at 25°C. The first injection and the corresponding heat release are not shown due to systematic uncertainties in the measurement. The proteins and the DNA were both in a final buffer of 50mM Tris, 200mM NaCl, 5mM EDTA and 5mM β-mercaptoethanol at pH 8.0. The solid lines represent the fit of the data to the function based on the binding of a ligand to a macromolecule using the Microcal ORIGIN software [92].

to be enthalpically more favorable by about 5 kcal/mol (Table 3-1), implying that the presence of an extra base pair between the TGA and TCA half-sites in CRE accounts for favorable electrostatic and hydrophobic interactions with the protein. However, the enthalpic advantage of CRE over TRE is offset by an equally greater entropic penalty, resulting in the overall weaker binding of CRE relative to TRE. This salient observation is in stark contrast to the finding that ΔH is more negative for the binding of bZIP domains of GCN4 to TRE versus CRE [90, 91]. However, in these previous studies [90, 91], the more favorable ΔH for the binding of bZIP domains of GCN4 to TRE is offset by

Table 3-1. Experimentally determined thermodynamic parameters for the binding of bZIP domains of Jun-Fos heterodimer to dsDNA oligos containing TRE and CRE consensus sequences using ITC at 25°C and pH 8.0.

	K_d/μM	ΔH/kcal mol^{-1}	$T\Delta S$/kcal mol^{-1}	ΔG/kcal mol^{-1}
TRE	0.15 ± 0.01	-31.62 ± 0.09	-22.29 ± 0.08	-9.33 ± 0.01
CRE	0.21 ± 0.02	-36.23 ± 0.06	-27.10 ± 0.11	-9.12 ± 0.04

The values for the various parameters shown were obtained from the fit of a function — based on the binding of a ligand to a macromolecule using the law of mass action assuming a 1:1 binding stoichiometry [23] — to the ITC isotherms shown in Figure 3-2. The binding stoichiometries to the fits agreed to within ±10%. Errors were calculated from 2-3 independent measurements. All errors are given to one standard deviation.

a much greater entropic penalty relative to CRE and the net result is that, unlike the scenario reported here, it is CRE site that overall appears to bind to GCN4 with a slightly higher affinity relative to TRE site. It is worthy of note that enthalpy-driven nature of protein-DNA interactions observed here is neither a rule nor an exception to the rule, as numerous examples of protein-DNA interactions under enthalpic as well as entropic control have been reported previously [90, 93-100].

Although protein-DNA interactions can be driven by either enthalpic or entropic or a combination of both factors [90, 94-97, 101-103], the large unfavorable entropic change — -75 cal/mol/K for the binding of TRE and -91 cal/mol/K for the binding of CRE — observed here is of significant interest. What might be the molecular basis of such an unfavorable entropy change observed here? Net entropic changes (ΔS) upon protein-DNA interactions are widely considered to result from an interplay between three major entropic forces symbolized as ΔS_{solv}, ΔS_{conf} and ΔS_{mix}. ΔS_{solv} is the favorable entropy change due to enhancement in the degrees of freedom of solvent molecules as a result of their restructuring and displacement, particularly around apolar groups, upon molecular associations. In the case of protein-DNA association, the release of counterions from DNA upon binding to protein is also likely to contribute favorably to ΔS_{solv}. ΔS_{conf}

is the unfavorable entropic change that arises from the restriction of conformational degrees of freedom of the backbone and sidechain atoms upon molecular associations. It has been suggested that the basic regions in the bZIP domains of Jun-Fos heterodimer are unstructured in the absence of DNA and undergo folding only upon DNA binding [104].

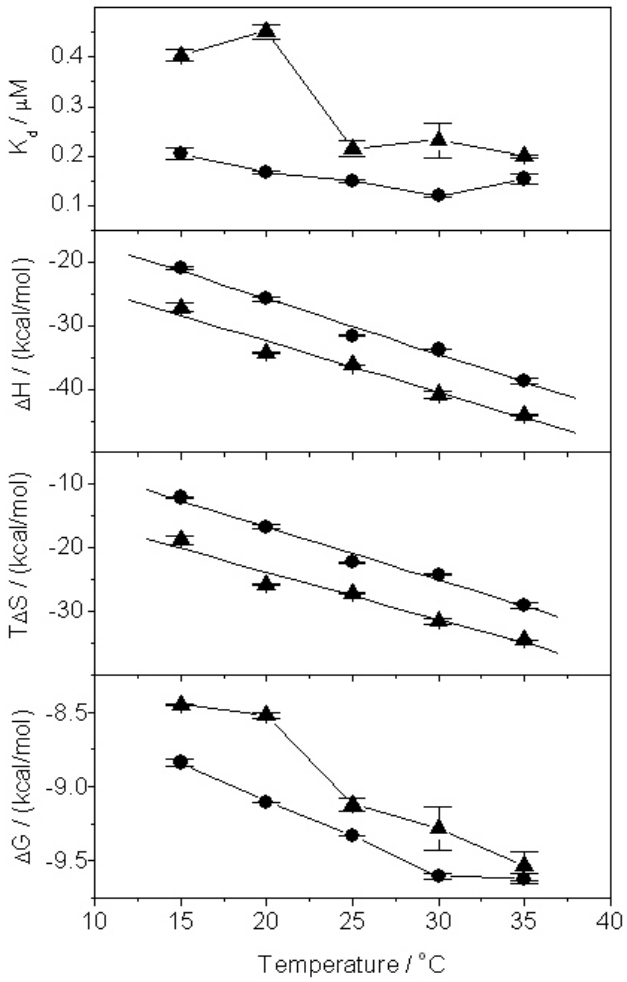

Figure 3-3. Dependence of thermodynamic parameters K_d, ΔH, TΔS and ΔG on temperature for the binding of bZIP domains of Jun-Fos heterodimer to dsDNA oligos containing TRE (●) and CRE (▲) sites. bZIP domains of Jun-Fos heterodimer in the calorimetric cell were titrated with 25 x 10μl injections of dsDNA oligo from the injection syringe at various temperatures in the range 15-35°C. The proteins and the DNA were both in a final buffer of 50mM Tris, 200mM NaCl, 5mM EDTA and 5mM β-mercaptoethanol at pH 8.0. To determine the various thermodynamic parameters, the ITC isotherms were fit to the function based on the binding of a ligand to a macromolecule using the Microcal ORIGIN software [23]. Each data point is the arithmetic mean of 2-3 experiments. All error bars are given to one standard deviation. The solid lines for the ΔH versus temperature and TΔS versus temperature plots show linear fits to the data, while the solid lines for the K_d versus temperature and ΔG versus temperature plots show straight lines merely connecting the data points for clarity.

Thus, such restructuring of protein upon DNA binding could further negatively contribute to the ΔS_{conf}. Finally, ΔS_{mix} is the unfavorable entropic change due to the restriction in the translational, rotational and vibrational degrees of freedom of molecules upon binding. Several lines of evidence suggest that ΔS_{mix} typically contributes no more than about –10 cal/mol/K of entropy penalty to the overall entropic change upon binding [72, 105-107]. On the basis of these arguments, we attribute the unfavorable entropic change incurred upon the binding of bZIP domains of Jun-Fos heterodimer to DNA largely to the loss of conformational degrees of freedom of backbone and sidechain atoms in both the protein and DNA as embodied in the term ΔS_{conf}.

3.4.2 Enthalpy and entropy compensate the effect of temperature on binding

Thermodynamic behavior of intermolecular associations can be highly dependent on the ambient temperature and knowledge of how thermodynamics vary as a function of temperature can provide invaluable insights into the mechanism of molecular recognition. In an effort to determine the effect of temperature on the various thermodynamic parameters, we analyzed the binding of the bZIP domains of Jun-Fos heterodimer to dsDNA oligos containing the TRE and CRE consensus sites in the temperature range 15-35°C (Figure 3-3). Our data indicate that both the enthalpic (ΔH) and entropic ($T\Delta S$) contributions to the overall free energy of binding (ΔG) show strong temperature-dependence and that both ΔH and $T\Delta S$ largely compensate for each other to generate ΔG that is virtually independent of temperature — while ΔH and $T\Delta S$ experience more than 20 kcal/mol change in their contributions to binding in going from 15°C to 35°C, ΔG gains no more than 1 kcal/mol over the same temperature range. Consistent with this

observation is the relatively constant nature of the binding affinity (0.1-0.4μM) over the same temperature range for the interaction of both the TRE and CRE dsDNA oligos with the bZIP domains of Jun-Fos heterodimer. It is of worthy note that the binding affinity for the interaction of bZIP domains with CRE appears to experience a sudden increase from a value of about 0.4μM at a temperature of 20°C and lower to a value of about 0.2μM at a temperature of 25°C and higher. This unexpected break cannot be accounted for and we do not believe that it is of any significance. We limited our investigations of the effect of temperature on the thermodynamics of protein-DNA interaction within this narrow temperature range due to a number of technical hurdles — below a temperature of 15°C, the observable ITC signal in the form of thermal power becomes substantially attenuated such that no reliable measurements can be made, while above a temperature of 35°C, bZIP domains have the tendency to undergo melting and thus measurements at higher temperatures may lead to inconsistencies [90, 91]. Regardless of these limitations, our data suggest that within this temperature range, the thermodynamics of binding of the bZIP domains of Jun-Fos heterodimer to TRE and CRE sites are overwhelmingly driven by favorable enthalpic changes with unfavorable entropic contributions to the overall free energy of binding.

The linear and opposing dependence of ΔH and $T\Delta S$ as a function of temperature, while maintaining a more or less constant ΔG, is a common feature observed in protein folding and binding reactions [70, 108, 109]. This phenomenon implies that there exists a temperature T_H where ΔH contribution to the free energy of binding changes sign. In the case of the binding of bZIP domains of Jun-Fos heterodimer to their DNA duplexes, ΔH will become endothermic and hence thermodynamically unfavorable below this T_H. Our

analysis of the dependence of ΔH versus temperature suggests that T_H for the binding of bZIP domains of Jun-Fos heterodimer to TRE and CRE sites is -10°C and -20°C, respectively — the binding of Jun-Fos heterodimer to TRE and CRE sites will be expected to be endothermic below temperatures of about -10°C and -20°C, respectively, and thus under these conditions, entropy not enthalpy is likely to drive the binding process. For further curiosity, we also determined the temperatures at which TΔS and ΔG will change sign for the binding of bZIP domains of Jun-Fos heterodimer to TRE and CRE sites in DNA. Extrapolation of the plots of TΔS versus temperature suggests that TΔS will become positive and hence thermodynamically favorable below temperatures of about 0°C and -10°C for binding to TRE and CRE, respectively. Similarly, extrapolation of the plots of ΔG versus temperature suggests that ΔG will reach zero at approximate temperatures of -210°C and –125°C for binding to TRE and CRE, respectively — the temperatures at which enthalpic and entropic contributions to the free energy of binding will have exactly equal and opposing components.

It is thus striking that despite an entropy penalty encountered at physiological temperatures, the entropy change will be favorable below a temperature of about -10°C for the binding of Jun-Fos heterodimer to both the TRE and CRE sites in DNA and, below a temperature of about -20°C, the binding will be predominantly driven by entropic factors with unfavorable contributions from enthalpic forces. We believe that the driving force for binding to switch from being under enthalpic control at physiological temperatures to being under entropic control at lower temperatures is largely due to the change in the ΔS_{conf} component of the overall entropy change of the system. Unlike at physiological temperatures, the change in ΔS_{conf} is unlikely to be the major entropic

penalty at lower temperatures due to the loss of reduction in the degrees of freedom of backbone and sidechain atoms available to molecules prior to binding as a result of loss of kinetic energy. Thus, molecular associations may not suffer from the loss of ΔS_{conf} as much as those encountered at higher physiological temperatures. With the ΔS_{conf} penalty being significantly reduced at lower temperatures, the ΔS_{solv} gain as a result of dehydration of water molecules surrounding molecular surfaces that become occluded from the solvent upon binding is likely to generate favorable entropic changes needed to drive the binding to completion against the backdrop of unfavorable enthalpic changes — which may also be accounted for by the loss of kinetic energy available for molecular collisions at lower temperatures.

3.4.3 Heat capacity change results from both folding and binding

The temperature-dependence of ΔH is related to heat capacity of binding (ΔC_p) by Kirchhoff's relationship $\Delta C_p = d(\Delta H)/dT$ — the slope of a plot of ΔH versus temperature equates to ΔC_p. Heat capacity is an important thermodynamic parameter in that it is related to the extent of the burial, occlusion and dehydration of molecular surfaces from surrounding solvent molecules upon intermolecular association — technically referred to as the change in solvent-accessible surface area (ΔSASA) [70, 94, 110, 111]. As such, this information is critical to understanding the mechanism of molecular recognition and, in the context of protein-DNA interactions, such information can further help us understand the role of thermodynamics pertinent to the regulation of transcriptional machinery.

In an attempt to understand how the binding of Jun-Fos heterodimer to DNA affects SASA, we calculated ΔC_p of –0.87 kcal/mol/K and –0.81 kcal/mol/K from the

slopes of ΔH versus temperature plots obtained for the binding of bZIP domains of Jun-Fos heterodimer to TRE and CRE sites, respectively (Figure 3-3). What might be the significance of the negative values of ΔC_p observed here? A positive value of ΔC_p implies that the occlusion of polar surfaces dominates the intermolecular association over apolar surfaces [70, 112, 113]. The fact that the values of ΔC_p are negative suggests that the occlusion of apolar surfaces dominates the occlusion of polar surfaces for the binding of the bZIP domains of Jun-Fos heterodimer to both the TRE and CRE sites. The slightly more negative ΔC_p for the binding of the bZIP domains of Jun-Fos heterodimer to TRE likely reflects a greater burial of apolar surfaces over polar surfaces for the binding of CRE. Experimental determination of values of ΔC_p combined with ΔH_{60} (enthalpy change at 60°C) have been widely used to quantitatively calculate changes in polar SASA ($\Delta SASA_{polar}$), apolar SASA ($\Delta SASA_{apolar}$) and total SASA ($\Delta SASA_{total}$) upon intermolecular association [69-73]. Such changes in SASA upon the binding of bZIP domains of Jun-Fos heterodimer to DNA from our thermodynamic measurements are reported in Table 3-2.

It has been suggested that the basic regions in bZIP domains are largely unstructured in the absence of DNA and undergo folding only upon binding to DNA [62, 63, 114-116]. Although the solution or crystal structure of the bZIP domains of Jun-Fos heterodimer in the absence of DNA has never been determined, circular dichroism studies provide stern evidence that the basic regions in the bZIP domains of Jun-Fos heterodimer are also unstructured in the absence of DNA and undergo folding only upon DNA binding [104]. In light of this foregoing argument, it is likely that change in SASA determined from our thermodynamic measurements above represents the folding of the

Table 3-2. Changes in polar SASA ($\Delta SASA_{polar}$), apolar SASA ($\Delta SASA_{apolar}$) and total SASA ($\Delta SASA_{total}$) upon the binding of bZIP domains of Jun-Fos heterodimer to dsDNA oligos containing TRE and CRE sites.

DNA site →	TRE			CRE		
Method →	Thermodynamic	Structural		Thermodynamic	Structural	
Model →	None	Rigid Body	Induced Fit	None	Rigid Body	Induced Fit
$\Delta SASA_{polar}/Å^2$	−2905	−1371	−2602	−3028	−1430	−2507
$\Delta SASA_{apolar}/Å^2$	−3612	−1401	−3489	−3550	−1484	−2499
$\Delta SASA_{total}/Å^2$	−6517	−2772	−6091	−6578	−2914	−5006

The changes in SASA are calculated and compared from both the thermodynamic and structural data. SASA values based on thermodynamic data were obtained from the measurement of ΔC_p for the binding of the bZIP domains of Jun-Fos heterodimer to dsDNA oligos containing TRE and CRE sites (Figure 3-3) using expressions [4] and [5], while SASA values based on structural data were derived from 3D structural models of the bZIP domains of Jun-Fos heterodimer alone and in complex with dsDNA oligos containing TRE and CRE sites (Figure 3-4) using expressions [6] and [7]. For SASA values calculated from structural data, two models were assumed — the Rigid Body model and the Induced Fit model. In the Rigid Body model, it was assumed that the bZIP domains of Jun-Fos heterodimer undergo no conformational change upon binding to DNA. In the Induced Fit model, it was assumed that the basic regions in the bZIP domains of Jun-Fos heterodimer are fully unstructured and only become structured upon binding to DNA. ΔSASA values calculated from thermodynamic data make no assumptions and are thus model-independent.

basic regions in the bZIP domains of Jun-Fos heterodimer in concert with binding to DNA. To test that this is so, we also determined changes in SASA upon the binding of the bZIP domains of Jun-Fos heterodimer to DNA from structural data independent of our thermodynamic measurements. To calculate changes in SASA upon the binding of the bZIP domains of Jun-Fos heterodimer to DNA from structural data, we assumed two models of binding — the Rigid Body model and the Induced Fit model. In the Rigid Body model, it was assumed that the bZIP domains of Jun-Fos heterodimer undergo no conformational change upon binding to dsDNA oligos containing TRE and CRE sites. In this model, the 3D structure of the bZIP domains of Jun-Fos heterodimer and the dsDNA oligos containing TRE and CRE sites were identical in both the free and the bound states — the Rigid Body model of binding assumes that no folding occurs in the protein or the DNA upon association and that the protein-DNA interaction is solely driven by the binding process alone. In the Induced Fit model, it was assumed that the basic regions in the bZIP domains of Jun-Fos heterodimer are fully unstructured and only become structured upon binding to dsDNA oligos containing TRE and CRE sites — the Induced

Fit model of binding assumes that partial folding of the bZIP domains in the basic regions occurs concomitantly upon association and that the protein-DNA interaction is accompanied by both the folding and the binding process.

Table 3-2 summarizes and compares values for $\Delta SASA_{polar}$, $\Delta SASA_{apolar}$ and $\Delta SASA_{total}$ upon the interaction of the bZIP domains of Jun-Fos heterodimer to TRE and CRE sites, as calculated from our thermodynamic and structural data. Our analysis shows that there are significant deviations between the values calculated from thermodynamic data and the Rigid Body model of binding. In contrast, the values determined from thermodynamic data agree par excellence with those calculated from the Induced Fit model. While the changes in SASA determined from thermodynamic data are off by more than 50% relative to those determined from structural data assuming the Rigid Body model for both TRE and CRE binding, these values agree within about 5% and 20% to those calculated from structural data assuming the Induced Fit model for TRE and CRE binding, respectively. The small anomalies in the values for changes in SASA between those obtained from thermodynamic data versus those calculated from structural data assuming the Induced Fit model are likely due to errors in the atomic coordinates of the structural models. These anomalies particularly become more pronounced in the case of the binding of the bZIP domains of Jun-Fos heterodimer to CRE site due to the poor quality of the 3D structural model of bZIP domains of Jun-Fos heterodimer in complex with CRE versus TRE site. While the structural model of the bZIP domains of Jun-Fos heterodimer in complex with dsDNA oligo containing TRE site was determined from reliable atomic coordinates of the crystal structure of this complex [PDB code: 1FOS], experimental structure of the bZIP domains of Jun-Fos heterodimer in complex with CRE

site is not available and thus had to be modeled on the basis of sequence homology with the bZIP domains of Jun-Jun homodimer in complex with CRE site [PDB code: 1JNM]. An alternative explanation for the anomalies observed in the values for changes in SASA between those obtained from thermodynamic data versus those calculated from structural data assuming the Induced Fit model may be due to the assumption that DNA experiences no conformational change upon interaction with the protein in spite of the evidence that it undergoes bending upon binding [50, 58, 117, 118]. Nonetheless, this latter assumption is an excellent approximation in our a priori calculations of changes in SASA from structural data due to negligible occlusion of molecular surface in DNA upon bending compared to rather large surface area buried upon protein-DNA contacts coupled with protein folding. It is thus not surprising that, despite small anomalies, our values for changes in SASA upon protein-DNA interaction calculated from thermodynamic dataversus those calculated from structural data assuming the Induced Fit model show remarkable consistency. In short, we believe that the significant underestimation of changes in SASA calculated from structural data assuming the Rigid Body model are due to the unrealistic assumption that neither protein nor DNA underwent any conformational change upon the formation of protein-DNA complex. However, in the Induced Fit model of binding, this unrealistic assumption is eliminated. The fact that the changes in SASAobtained from the structural data assuming the Induced Fit model and thermodynamic studies agree to within experimental error unequivocally demonstrates that the bZIP domains of Jun-Fos heterodimer undergo coupled folding and binding to DNA.

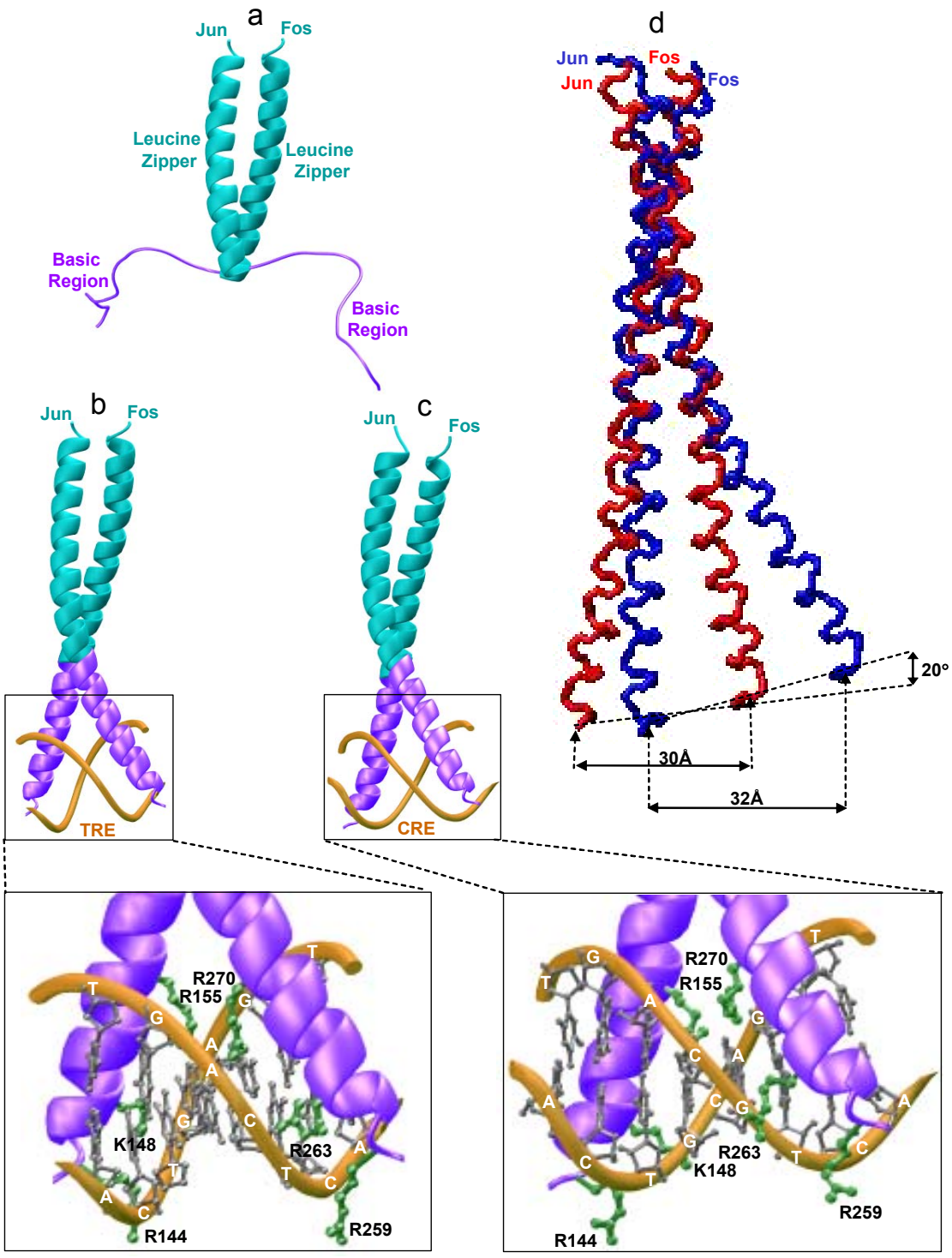

Figure 3-4. Modeled structures of bZIP domains of Jun-Fos heterodimer alone and in complex with dsDNA oligos containing TRE and CRE sites. For clarity, only the core regions of bZIP domains of Jun (residues 258-311) and Fos (residues 143-196) are shown. For the same reason, nucleotides flanking the TRE (TGACTCA) and CRE (TGACGTCA) sites are also omitted. (a) bZIP domains of Jun-Fos heterodimer in the absence of DNA. Leucine zippers are shown in cyan and the basic regions in purple. While the leucine zippers adapt α-helical conformations, the basic regions that contact the DNA are largely

unstructured and most likely adapt extended conformations. (b) bZIP domains of Jun-Fos heterodimer in complex with TRE site (yellow). Leucine zippers in the bZIP domains are shown in cyan and the basic regions in purple. The inset shows intermolecular interactions between critical basic residues (green) in bZIP domains and the DNA bases (gray) plus the backbone phosphates (yellow). The structural model was generated using the crystal structure of bZIP domains of Jun-Fos heterodimer in complex with dsDNA oligo containing the TRE site [PDB code: 1FOS] as a template in MODELLER. (c) bZIP domains of Jun-Fos heterodimer in complex with CRE site (yellow). Leucine zippers in the bZIP domains are shown in cyan and the basic regions in purple. The inset shows the intermolecular interactions between critical basic residues (green) in bZIP domains and the DNA bases (gray) plus the backbone phosphates (yellow). The structural model was generated using the crystal structure of bZIP domains of Jun-Jun homodimer in complex with dsDNA oligo containing the CRE site [PDB code: 1JNM] as a template in MODELLER. (d) Superimposition of the orientation of the bZIP domains of Jun-Fos heterodimer in complex with TRE site (red) versus its orientation in complex with CRE site (blue). Superimposition was performed by the alignment of the backbone N, CA and C atoms of the five signature leucines L1, L2, L3, L4 and L5 in each bZIP pair of Jun-Fos heterodimer. The double-headed horizontal arrows compare the distance between the two ends of the bZIP forceps in each of the two complexes in contact with TRE and CRE sites. The double-headed vertical arrow compares the rotation of the bZIP forceps in contact with TRE site relative to CRE site.

3.4.4 Structural modeling allows rationalization of thermodynamic data

In an attempt to rationalize our thermodynamic data in structural terms, we modeled 3D structures of the bZIP domains of Jun-Fos heterodimer alone and in complex with dsDNA oligos containing TRE and CRE sites on the basis of homology modeling (Figure 3-4). Given that the insertion of an extra base pair at the center of the CRE site would increase the distance between the TGA and TCA half-sites, our structural models would also test the extent to which the bZIP domains of Jun-Fos heterodimer may have to undergo conformational change to accommodate the wider spacing between these half-sites in the major grooves of DNA.

In agreement with our thermodynamic data, our 3D structural model of the bZIP domains of Jun-Fos heterodimer shows the basic regions unstructured and more or less fully extended in the absence of DNA (Figure 3-4a). It is believed that such unfolding occurs due to unfavorable interactions between closely spaced basic residues. However, the basic regions become structured upon the introduction of dsDNA oligos containing TRE and CRE sites due to neutralization of basic charge with negative DNA backbone

phosphates (Figures 3-4b and 3-4c). Our model predicts that the key residues in the basic regions that contact the DNA bases and backbone phosphates would have to undergo significant conformational changes from being fully exposed to solvent to becoming dehydrated upon contact with DNA — a scenario that fully accounts for the magnitudes of ΔC_p observed in our thermodynamic measurements for the binding of bZIP domains of Jun-Fos heterodimer to TRE and CRE sites (Table 3-2). The conformational changes occurring in the protein upon DNA binding can be best visualized in the form of a movie accessible on the WWW @ http://labs.med.miami.edu/farooq/movies/AP1bZIP.

Consistent with the crystal structure of the protein-DNA complex [43], the bZIP domains of Jun-Fos heterodimer bind to half-sites within the major grooves formed by the TRE site related by a 180° rotation about the dyad axis of the complex (Figure 3-4b). One half-site of TRE packs against the basic region of Jun bZIP while the other against the basic region of Fos bZIP. Since the TRE site is pseudo-palindromic, it is believed that the basic regions within Jun and Fos may assume no preferred orientation and that these could easily switch between the two half-sites [43, 57]. However, the bZIP domains of Jun-Fos assume asymmetry in that the helical axis of Fos is somewhat straighter than that of Jun and the latter appears to wrap around the straighter axis of Fos. Although the central axis of the coiled-coil bZIP domains of Jun-Fos is almost perpendicular to the double-helical axis of DNA, it is believed that this may not always be the case and that the central axis of the bZIP domains could undergo bending to some degree and that such a flexibility may represent a general feature of the interaction of bZIP domains with DNA [43]. A number of residues in the basic regions of Jun and Fos engage in hydrogen bonding, hydrophobic contacts and electrostatic interactions with the bases and backbone

phosphates in the TRE site and are believed to be essential for high affinity binding of Jun-Fos heterodimer to DNA [57, 119]. In particular, these include the basic residues R259, R263 and R270 in the basic region of Jun, and R144, K148 and R155 in the basic region of Fos (inset to Figure 3-4b).

How does an increase in the distance between the TGA and TCA half-sites in the consensus sequence TGACTCA, as a result of insertion of an extra base pair to generate the CRE site, affect the binding of Jun-Fos heterodimer to DNA? As for TRE site, the bZIP domains of Jun-Fos heterodimer bind to CRE site in a very similar manner — the more curved Jun wraps around a straighter Fos and the central axis of the coiled coil Jun-Fos heterodimer is almost perpendicular to the double-helical axis of DNA (Figure 3-4c). However, our structural model of the bZIP domains of Jun-Fos in complex with CRE site suggests that the increase in the distance between the TGA and TCA half-sites would be compensated by a slight re-orientation of the basic regions in the bZIP domains of Jun-Fos heterodimer so as to splay apart the bZIP forceps and re-align the key residues in the basic regions that contact the DNA bases and backbone phosphates (inset to Figure 3-4c). Additionally, the extra G insert at the center of the CRE site would contribute to strong favorable hydrogen bonding and hydrophobic interactions with the bZIP domains of Jun-Fos heterodimer. These additional protein-DNA interactions may also account for the more favorable enthalpic change observed upon the binding of bZIP domains of Jun-Fos heterodimer to CRE site relative to TRE site (Table 3-1).

Superimposition of bZIP domains of Jun-Fos heterodimer in complex with TRE site versus CRE site reveals that the backbone atoms of basic regions in Jun and Fos open up by about 2Å at their N-termini in contact with major grooves of DNA (Figure 3-4d).

This increase in the distance is consistent with the knowledge that the addition of an extra base pair to DNA adds about 3Å to its double-helical length. Furthermore, there is also a rotation of about 20° of the bZIP domains of Jun-Fos heterodimer in complex with CRE site relative to TRE site due to the insertion of an extra base pair. Binding of bZIP domains of Jun-Fos heterodimer to CRE site is further aided by the slight bending of DNA relative to that in TRE site. Our model of bZIP domains of Jun-Fos heterodimer in complex with CRE site also predicts that the residues in the basic regions of both Jun and Fos that are critical for interaction with TRE site are also likely to play a key role for interaction with CRE site and that the insertion of an extra base pair does not alter the pattern of protein-DNA interactions. These observations are in an excellent agreement with the crystal structure of the bZIP domains of GCN4 in complex with CRE site [120, 121].

Apart from the subtle differences in the protein and DNA conformations mentioned above, both the TRE and CRE sites appear to occlude similar surface areas at the protein-DNA interface upon binding (Table 3-2). Although the larger CRE site would be expected to engage in more extensive contacts with DNA and hence lead to greater occlusion of surface area upon binding, the conformational change in the protein necessary to accommodate the larger CRE site appears to counterbalance this due to slight re-arrangement of residues at the protein-DNA interface. Nonetheless, the additional contacts made by the introduction of an extra base pair between the TGA and TCA half-sites manifests itself in the form of additional release of heat of about 5 kcal/mol relative to TRE site upon binding to the bZIP domains of Jun-Fos heterodimer. This favorable gain in enthalpy is, however, offset by an equal but opposing amount of

unfavorable entropic cost for the binding of CRE site relative to TRE site, making the latter site equally favorable, if not more favorable, for binding the Jun-Fos heterodimer.

3.5 Concluding remarks

Although critical role of the Jun-Fos heterodimeric transcription factor in cellular signaling was reported over two decades ago [4, 41, 48, 56, 78-81], thermodynamics of this key protein-DNA interaction have hitherto not been reported. Knowledge of thermodynamics is central to understanding the intrinsic forces that determine the structure and stability of protein-DNA interactions. In an attempt to elucidate the role of various thermodynamic forces at play in the binding of the Jun-Fos heterodimer to DNA, we have reported herein a detailed ITC analysis of this key protein-DNA interaction pertinent to cellular signaling and cancer. The lack of availability of such thermodynamic information over the past two decades underlies the difficulties of biophysical analysis of free bZIP domains and this has been particularly problematic in the case of the bZIP domains of Jun-Fos heterodimer. However, the knowledge that tagging recombinant proteins with thioredoxin may enhance their stability in solution led to our success with being able to work with recombinant thioredoxin-tagged bZIP domains of Jun-Fos heterodimer under conditions necessary for their biophysical analysis and thus gaining invaluable insights into the thermodynamics of this key protein-DNA interaction.

One might be tempted to question the validity of our thermodynamic data due to the enhanced stability afforded by the presence of thioredoxin tag at the N-termini of bZIP domains of Jun and Fos. However, we have shown herein that the thioredoxin tag does not physically interact with either bZIP domain and that its role here is simply to enhance the bZIP stability through its ability to offset the balance of destabilizing

interactions of bZIP domains with water molecules. In this regard, the thioredoxin tag may act as a blessing-in-disguise in that it may impart upon bZIP domain a conformation more reminiscent to that found in the context of full-length Jun or Fos protein. On the contrary, thermodynamic studies on untagged isolated bZIP domains could also raise the question of their validity as these domains may behave in a slightly differential manner when taken out of the context of a full-length protein. Ideally, one would like to carry out studies of this nature on bZIP domains in the context of their full-length proteins but given that their expression and purification could pose new challenges, one will always be reduced to work with some sort of assumptions.

Given the reasonable assumption that the thioredoxin tag is unlikely to alter the thermodynamic properties of bZIP domains in any significant way, our study shows for the first time that the binding of Jun-Fos heterodimer to DNA is under enthalpic control and that this process is accompanied by an unfavorable loss of entropy at physiological temperatures. We attribute this entropic penalty to the loss of conformational degrees of freedom of backbone and sidechain atoms in both the protein and DNA upon binding. The net entropic change (ΔS) can be decomposed into the following major entropic contributions:

$$\Delta S = \Delta S_{solv} + \Delta S_{conf} + \Delta S_{mix}$$

where ΔS_{solv}, ΔS_{conf} and ΔS_{mix} are the entropic contributions due to restructuring of solvent, changes in conformational degrees of freedom of backbone and sidechain atoms, and changes in translational, rotational and vibrational degrees of freedom upon binding, respectively. While ΔS_{solv} has been shown to be equal to $\Delta C_p \ln[T/385]$ (where T is the absolute temperature), ΔS_{mix} is essentially the cratic entropy for a bimolecular reaction

and equates to –8 cal/mol/K [72, 106]. From the experimentally determined values of ΔC_p here (Figure 3-3), ΔS_{solv} for the binding of bZIP domains of Jun-Fos heterodimer to TRE and CRE can be shown to be equal to +226 cal/mol/K and +210 cal/mol/K, respectively, at 25°C. Thus, given that ΔS for the binding of bZIP domains of Jun-Fos heterodimer to TRE and CRE is respectively –75 cal/mol/K and –91 cal/mol/K at 25°C (Table 3-1), ΔS_{conf} for the binding of bZIP domains of Jun-Fos heterodimer to both TRE and CRE turns out to be –293 cal/mol/K at 25°C. In other words, the entropic penalty due to the loss of conformational degrees of freedom of backbone and sidechain atoms in both the protein and DNA upon binding is –293 cal/mol/K at 25°C. However, this penalty is largely compensated by the favorable entropic gain of +226 cal/mol/K and +210 cal/mol/K at 25°C due to the restructuring and displacement of water molecules upon the binding of bZIP domain of Jun-Fos heterodimer to TRE and CRE sites, respectively.

Although it has been known that the bZIP domains of Jun-Fos heterodimer could recognize both the TRE and CRE sites [56, 57, 89], our study here shows that the binding of Jun-Fos heterodimer to both sites occurs with almost indistinguishable affinities and that differences in the thermodynamic parameters ΔH and $T\Delta S$ largely compensate for each other without any significant impact on the overall ΔG for binding. This is remarkably striking given that the insertion of an extra base pair between the TGA and TCA half-sites is expected to increase the distance between them by about 3Å. Our structural models of bZIP domains of Jun-Fos heterodimer in complex with TRE and CRE sites suggest that this increase in distance between the TGA and TCA half-sites is counteracted by an equally accommodating conformational change in the protein that

allows the basic regions in the bZIP domains to open up by about 2Å at the protein-DNA interface and further undergo a rotation of about 20° relative to their positions in contact with TRE site in DNA. The magnitude of heat capacity changes and associated changes in SASA upon protein-DNA interaction determined from ITC measurements suggest strongly that the basic regions in the bZIP domains of Jun-Fos heterodimer are partially unstructured and become structured only upon interaction with DNA in a coupled folding and binding manner. This finding corroborates the notion that the coupled folding and DNA-binding is a general feature of the bZIP family of transcription factors [62, 63, 114-116].

Finally, current strategies for the design of drugs that can inhibit the oncogenic action of Jun-Fos heterodimer on cellular machinery are based on molecules that either interfere with the heterodimerization, or alternatively, compete with TRE and CRE sites for binding to Jun-Fos heterodimer. Our demonstration that the Jun-Fos heterodimer undergoes conformational change upon DNA binding may offer novel opportunities for the design of drugs that lock the Jun-Fos heterodimer in its partially unstructured state so as to completely abrogate its DNA binding potential.

4 Chapter 4: Thermodynamic Analysis of the Heterodimerization of Leucine Zippers of Jun and Fos Transcription Factors

4.1 Summary

Jun and Fos are components of the AP1 family of transcription factors and bind to the promoters of a diverse multitude of genes involved in critical cellular responses such as cell growth and proliferation, cell cycle regulation, embryonic development and cancer. Here, using the powerful technique of isothermal titration calorimetry, we characterize the thermodynamics of heterodimerization of leucine zippers of Jun and Fos. Our data suggest that the heterodimerization of leucine zippers is driven by enthalpic forces with unfavorable entropy change at physiological temperatures. Furthermore, the basic regions appear to modulate the heterodimerization of leucine zippers and may undergo at least partial folding upon heterodimerization. Large negative heat capacity changes accompanying the heterodimerization of leucine zippers are consistent with the view that leucine zippers do not retain α-helical conformations in isolation and that the formation of the native coiled coil α-helical dimer is attained through a coupled folding-dimerization mechanism.

4.2 Background

Jun and Fos are components of the AP1 transcription factor and are expressed in a wide variety of tissues [5, 41]. Upon activation by a diverse multitude of mitogenic signals, including up-regulation by MAP kinases, Jun and Fos heterodimerize and bind to specific DNA sequences found in the promoters of genes such as metallothionein IIa, collagenase, interleukin 2 and cyclin D1. In this manner, the Jun-Fos heterodimeric transcription factor plays a central role in coupling mitogenic stimuli to DNA

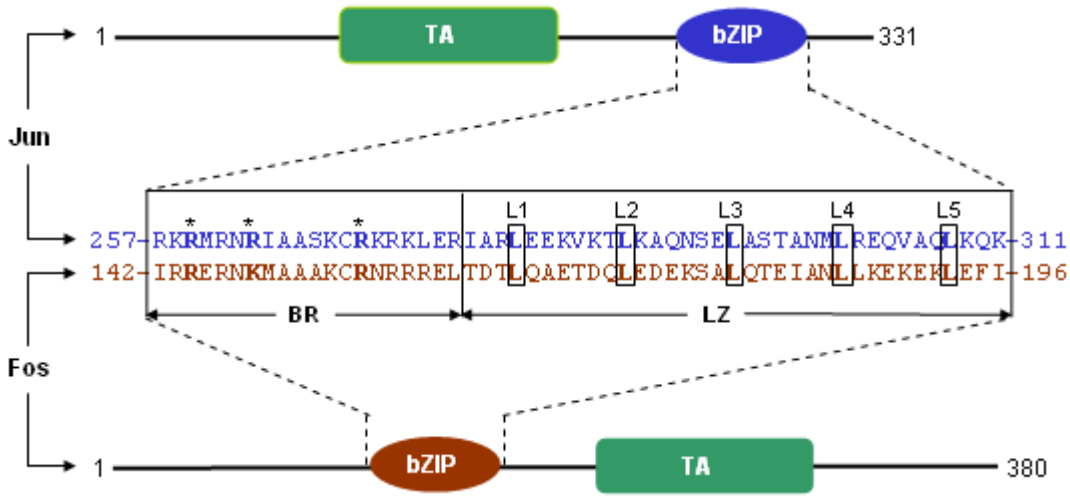

Figure 4-1. A schematic showing domain organization of Jun and Fos transcription factors containing the basic zipper (bZIP) domain and the transactivation (TA) domain. The positions of the N-terminal basic region (BR) and the C-terminal leucine zipper (LZ) subdomains relative to each other are indicated. The five signature leucines (L1-L5) characteristic of LZ subdomains are boxed and bold faced.

transcription and, in so doing, regulates a wide array of cellular processes such as cell growth and proliferation, cell cycle regulation and embryonic development [5].

The ability of Jun and Fos to recognize specific DNA sequences at the promoter regions resides in a region that has come to be known as the basic zipper (bZIP) domain (Figure 4-1). The bZIP domain can be further dissected into two well-defined functional subdomains termed the basic region (BR) at the N-terminus followed by the leucine zipper (LZ) at the C-terminus. The LZ subdomain is a highly conserved protein module found in a wide variety of transcription factors and structural proteins and contains a signature leucine at every seventh position within the five successive heptads of amino acid residues. The LZ subdomains adapt continuous α-helical conformations and induce heterodimerization of Jun and Fos by virtue of their ability to wrap around each other in a coiled coil dimer [41, 42]. Such intermolecular arrangement brings the BR subdomains at the N-termini of bZIP domains into close proximity and thereby enables them to insert

into the major grooves of DNA at the promoter regions in an optimal fashion in a manner akin to a pair of forceps [43].

Although the mechanism of the binding of bZIP domains to DNA is well-established, little is known about how protein-protein interactions dictate the heterodimerization of leucine zippers. To address this important issue, we have employed here the powerful technique of isothermal titration calorimetry (ITC) to measure and characterize thermodynamics of heterodimerization of leucine zippers of Jun and Fos from their respective homodimers in the context of the LZ subdomains and bZIP domains. Our data provide novel insights into the thermodynamics of a key protein-protein interaction pertinent to cellular transcriptional machinery.

4.3 Experimental procedures

4.3.1 Protein preparation

LZ subdomain (residues 277-331) and bZIP domain (residues 251-331) of human Jun as well as LZ subdomain (residues 162-216) and bZIP domain (residues 136-216) of human Fos were cloned into pET102 bacterial expression vector, with an N-terminal thioredoxin (Trx)-tag and a C-terminal polyhistidine (His)-tag, using Invitrogen TOPO technology. Recombinant proteins were expressed, purified and characterized as described previously [122].

4.3.2 ITC measurements

Isothermal titration calorimetry (ITC) measurements were performed on Microcal VP-ITC instrument and data were acquired and processed using fully automized features in Microcal ORIGIN software. All measurements were repeated 3-4 times. Briefly, the protein samples were prepared in 50mM Tris, 200mM NaCl, 1mM EDTA and 5mM β-

mercaptoethanol at pH 8.0 and de-gassed using the ThermoVac accessory for 10min. The experiments were initiated by injecting 25 x 10μl injections of 50-100μM of LZ subdomain or bZIP domain of Fos from the syringe into the calorimetric cell containing 1.8ml of 5-10μM of LZ subdomain or bZIP domain of Jun at a fixed temperature in the narrow range 20-30°C. The change in thermal power as a function of each injection was automatically recorded using Microcal ORIGIN software and the raw data were further processed to yield ITC isotherms of heat release per injection as a function of Fos to Jun molar ratio. The negligible heats of mixing and unfolding were subtracted from the heat of binding per injection by carrying out a control experiment in which the same buffer in the calorimetric cell was titrated against the LZ subdomain or bZIP domain of Fos in an identical manner. All other control experiments were performed as necessary. To extract the apparent equilibrium constant (K_d) and the enthalpy change (ΔH) associated with heterodimerization, the ITC isotherms were iteratively fit to the following built-in function by non-linear least squares regression analysis using the integrated Microcal ORIGIN software:

$$q(i) = (n\Delta HVP/2)\ \{[1+(L/nP)+(K_d/nP)] - [[1+(L/nP)+(K_d/nP)]^2 - (4L/nP)]^{1/2}\} \qquad [1]$$

where q(i) is the heat release (kcal/mol) for the ith injection, n is the stoichiometry of heterodimerization, V is the effective volume of Jun in the calorimetric cell (1.46ml), P is the total Jun concentration in the calorimetric cell (μM) and L is the total Fos concentration in the calorimetric cell at the end of each injection (μM). The above equation is derived from the binding of two molecules using the law of mass action assuming a 1:1 stoichiometry [74].

4.4 Results and discussion

4.4.1 Heterodimerization of leucine zippers is under enthalpic control

In an attempt to elucidate the thermodynamic forces governing the heterodimerization of leucine zippers of Jun and Fos, we employed the powerful technique of ITC. To shed light on the role of basic regions in the heterodimerization of leucine zippers, we performed ITC analysis on both the LZ subdomains (containing leucine zippers only) and bZIP domains (containing basic regions located N-terminal to leucine zippers). The basic regions and leucine zippers are also alternatively referred to as BR subdomains and LZ subdomains, respectively, throughout this study. Our ITC data indicate that the heterodimerization of leucine zippers in the context of both the LZ subdomains and bZIP domains is under strong enthalpic control and accompanied by an unfavorable loss of entropy at physiological temperatures (Figure 4-2 and Table 4-1). However, the heterodimerization of LZ subdomains proceeds with an affinity (0.06μM) that is over 2-fold greater than that observed for the heterodimerization of bZIP domains (0.13μM) — implying that the BR subdomains play an inhibitory role in the heterodimerization of leucine zippers. These results are in good agreement with previous studies based on non- calorimetric methods [123, 124]. Our study, however, provides complete thermodynamic signatures of the heterodimerization of the leucine zippers of Jun and Fos that hitherto have not been reported. What is the molecular basis of a largely enthalpy-driven nature of the heterodimerization of leucine zippers with unfavorable entropic contributions observed here? The leucine zippers of Jun and Fos adopt a coiled coil conformation comprised of parallel α-helices wound around each other like a pair of

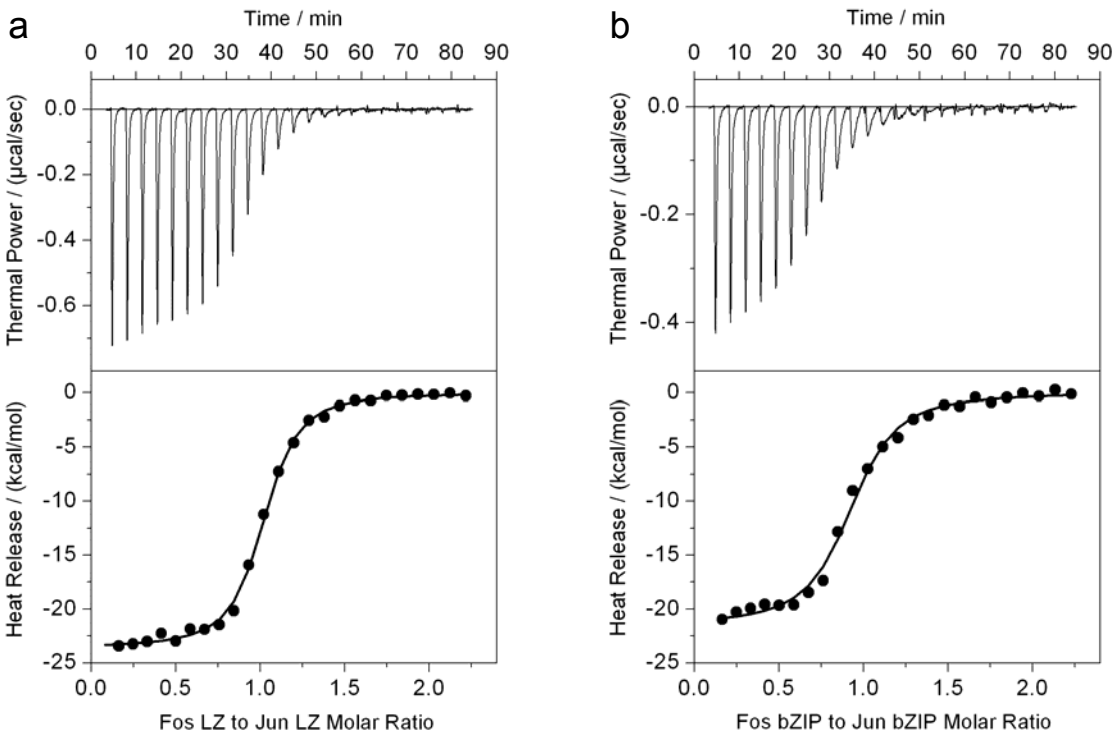

Figure 4-2. ITC analysis of the heterodimerization of LZ subdomains (a) and bZIP domains (b) of Jun and Fos. LZ subdomains or bZIP domains of Jun in the calorimetric cell were titrated with 25 x 10µl injections of corresponding LZ subdomains or bZIP domains of Fos from the injection syringe at 25°C. The solid lines to the data in the lower panel represent the fit to expression [1].

forceps. The α-helices are held together by molecular glue comprised of largely hydrophobic contacts with minor but important contributions from electrostatic interactions due to the salt bridging of oppositely charged residues [43]. In light of such structural observations, it is thus highly likely that the establishment of an extensive network of hydrophobic contacts and electrostatic interactions between a pair of α-helices could account for the favorable enthalpic contributions to the free energy of heterodimerization. The molecular basis of heat release upon the heterodimerization of leucine zippers may seem intuitive but that of entropic penalty begs a little more thought. The entropy change observed here is the net change resulting from a combination of favorable and unfavorable entropic forces upon molecular associations. The major

Table 4-1. Experimentally determined thermodynamic parameters for the heterodimerization of LZ subdomains and bZIP domains of Jun and Fos using ITC at 25°C and pH 8.0

	$K_d/\mu M$	$\Delta H/kcalmol^{-1}$	$T\Delta S/kcalmol^{-1}$	$\Delta G/kcalmol^{-1}$
LZ	0.06 ± 0.01	−23.21 ± 0.23	−13.52 ± 0.42	−9.69 ± 0.26
bZIP	0.13 ± 0.03	−21.75 ± 0.28	−12.35 ± 0.42	−9.40 ± 0.14

The values for the apparent equilibrium constant (K_d) and the enthalpy change (ΔH) associated with heterodimerization were obtained from the fit of expression [1] to the ITC isotherms shown in Figure 4-2. Free energy of heterodimerization (ΔG) was calculated from the relationship $\Delta G=RT\ln K_d$, where R is the universal molar gas constant (1.99 cal/mol/K) and T is the absolute temperature (K). Entropic contribution ($T\Delta S$) to heterodimerization was calculated from the relationship $T\Delta S=\Delta H-\Delta G$. The stoichiometries to the fits agreed to within ±10%. Errors were calculated from 2-3 independent measurements. All errors are given to one standard deviation.

favorable entropic force upon molecular associations is the enhancement in the degrees of freedom of water molecules as a result of their restructuring and displacement from molecular surfaces coming into contact with each other, particularly from around the apolar groups — this favorable contribution is usually denoted ΔS_{solv} or the change in solvent entropy. However, the favorable ΔS_{solv} is largely offset by the loss of conformational degrees of freedom of the backbone and sidechain atoms upon molecular associations — this unfavorable contribution is usually denoted ΔS_{conf} or the change in conformational entropy. There is also a slight negative contribution to the overall entropic change due to the restriction of translation, rotational and vibrational degrees of freedom of molecules upon association. In light of the foregoing argument, we attribute the large unfavorable entropy change observed here upon the heterodimerization of LZ subdomains and bZIP domains to the loss of conformational degrees of freedom of the backbone and sidechain atoms as embodied in the term ΔS_{conf}.

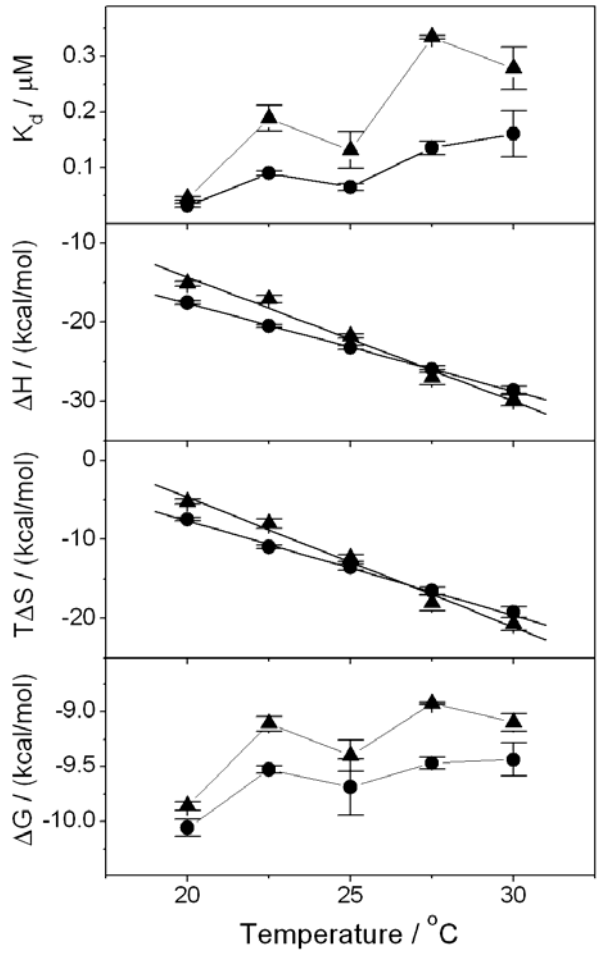

Figure 4-3. Dependence of thermodynamic parameters K_d, ΔH, TΔS and ΔG on temperature for the heterodimerization of LZ subdomains (●) and bZIP domains (▲) of Jun and Fos. Each data point is the arithmetic mean of 3-4 experiments. All error bars are given to one standard deviation. The solid lines for the ΔH and TΔS plots show linear fits to the data, while the solid lines for the K_d and ΔG plots show straight lines connecting the data points for clarity.

4.4.2 Enthalpic and entropic factors compensate the effect of temperature on the heterodimerization of leucine zippers

Thermodynamics of protein-protein interactions can be highly dependent on the ambient temperature and knowledge of how thermodynamics vary as a function of temperature can provide invaluable insights into the mechanism of protein oligomerization. In an effort to determine the effect of temperature on the various thermodynamic parameters, we analyzed heterodimerization of LZ subdomains and bZIP

domains of Jun and Fos in the narrow temperature range 20-30°C using ITC (Figure 4-3). Our data indicate that both the enthalpic (ΔH) and entropic (TΔS) contributions to the overall free energy of heterodimerization (ΔG) show strong temperature-dependence and that both ΔH and TΔS largely compensate each other to generate ΔG that is virtually independent of temperature— while ΔH and TΔS experience nearly 20 kcal/mol change in their contributions to heterodimer formation in going from 20°C to 30°C, ΔG gains no more than about 1 kcal/mol over the same temperature range. Consistent with this observation is the relatively constant nature of the apparent equilibrium constant of heterodimerization (0.05-0.3μM) over the same temperature range.

The linear and opposing dependence of ΔH and TΔS as a function of temperature, while maintaining a more or less constant ΔG, is a common feature observed in protein folding and binding reactions. This phenomenon gives rise to two key temperature points T_H and T_S — the temperatures where enthalpic (ΔH) and entropic (TΔS) contributions to the free energy change sign, respectively. In the case of the heterodimerization of leucine zippers of Jun and Fos, ΔH will become negative and hence thermodynamically favorable above T_H, while TΔS will become negative and hence thermodynamically unfavorable above T_S. Table 4-2 provides the values for T_H and T_S accompanying the heterodimerization of leucine zippers of Jun and Fos in the context of LZ subdomains and bZIP domains. As evidenced in Table 4-2, both T_H and T_S fall well below the physiological temperature of 37°C, implying that the heterodimerization of leucine zippers of Jun and Fos will be largely under enthalpic control accompanied by entropic penalty at physiological temperatures.

Table 4-2. Experimentally determined thermodynamic parameters for the heterodimerization of LZ subdomains and bZIP domains of Jun and Fos obtained from ITC measurements at various temperatures in the narrow range 20-30°C and pH 8.0

	T_H/°C	T_S/°C	ΔC_p/kcalmol^{-1} K^{-1}
LZ	+4.15 ± 0.42	+13.36 ± 0.39	−1.11 ± 0.03
bZIP	+10.78 ± 0.88	+17.10 ± 0.46	−1.57 ± 0.10

The values for the various parameters shown were obtained as follows. The values for T_H, the temperature at which ΔH is zero, were obtained from the extrapolation of linear fits to the ΔH versus temperature plots (Figure 4-3). The values for T_S, the temperature at which TΔS is zero, were obtained from the extrapolation of linear fits to the TΔS versus temperature plots (Figure 4-3). The values for ΔH_{60}, the enthalpy at 60°C, were obtained from the extrapolation of linear fits to the ΔH versus temperature plots (Figure 4-3). The values for ΔC_p, the heat capacity change, were obtained from the slopes of linear fits to the ΔH versus temperature plots (Figure 4-3). Errors were calculated from 3-4 independent measurements. All errors are given to one standard deviation.

The temperature-dependence of ΔH is related to heat capacity of binding (ΔC_p) by Kirchhoff's relationship $\Delta C_p = d(\Delta H)/dT$ — the slope of a plot of ΔH versus temperature equates to ΔC_p. Heat capacity is an important thermodynamic parameter in that it is related to the extent of the burial and dehydration of molecular surfaces from surrounding solvent molecules upon intermolecular association [70, 94, 110, 111]. As indicated in Table 4-2, the association of leucine zippers of Jun and Fos into a heterodimer results in large negative changes in heat capacity. What might be the significance of such large negative values of ΔC_p observed here? A positive value of ΔC_p implies that the residues being occluded from the solvent and hence residing at the interface of two molecular surfaces coming together are largely of polar nature with little or negligible contributions from apolar groups. The fact that the heat capacity changes are largely negative suggests strongly that the heterodimerization of LZ subdomains and bZIP domains involves substantial burial of hydrophobic residues with little contributions from polar residues. It should be noted here that protein-ligand interactions typically result in the magnitude of ΔC_p of less than –1000 cal/mol/K, while values of ΔC_p in the range –1000 to -2000

cal/mol/K are characteristic of proteins undergoing folding due to burial of a large number of apolar groups as a result of hydrophobic effect. Several lines of evidence suggest that the LZ subdomains are unstable as α-helices when in isolation and only fold into α-helices in the context of a coiled coil dimer [63, 114]. This is believed to be due to the fact that the dimer interface of a coiled coil is comprised of hydrophobic residues, created largely by the interdigitation of signature leucines from each α-helix, and thus each LZ α-helix is thermodynamically unstable in isolation. In light of these arguments, we believe that the large negative changes in heat capacity observed here most likely arise due to the association of unfolded leucine zippers of Jun and Fos into α-helical heterodimers. In other words, heterodimerization of Jun and Fos appears to be coupled to folding of α-helices of their respective leucine zippers.

4.4.3 Basic regions modulate the heterodimerization of leucine zippers

Although it is widely believed that electrostatic repulsions between basic residues in the BR subdomains prevent them from becoming structured in the absence of DNA [62, 104], our demonstration that the affinity of heterodimerization of leucine zippers is 2-fold greater in the context of LZ subdomains relative to bZIP domains suggests that the basic regions inhibit the heterodimerization of leucine zippers (Table 4-1). In an attempt to understand the thermodynamic basis of such differences, we generated differential thermodynamic signatures for the heterodimerization of LZ subdomains relative to bZIP domains (Figure 4-4a). In this analysis, a negative value of ΔΔH implies that the enthalpy change is more favorable for the heterodimerization of LZ subdomains relative to bZIP domains, a negative value of TΔΔS implies that the entropy change is less favorable for the heterodimerization of LZ subdomains relative to bZIP domains, and a negative value

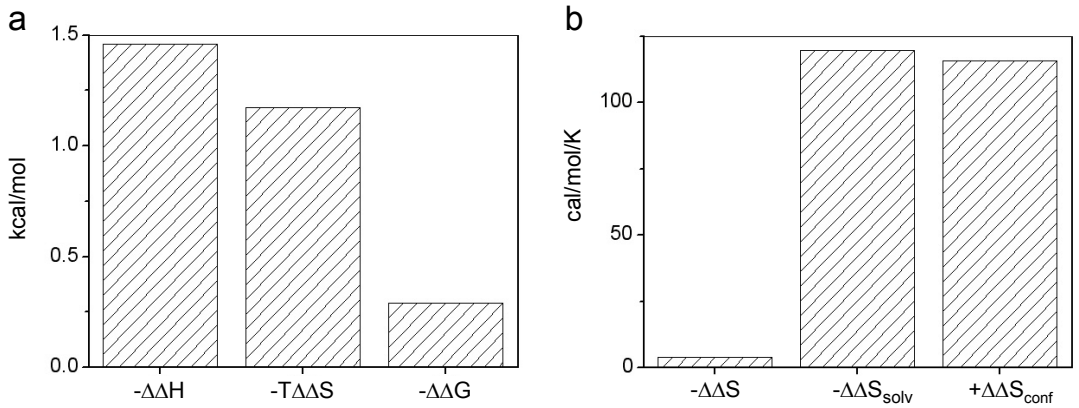

Figure 4-4. Differential energetics for the heterodimerization of LZ subdomains versus bZIP domains. (a) Differential thermodynamic signature for the heterodimerization of LZ subdomains relative to bZIP domains. $\Delta\Delta H$, $T\Delta\Delta S$ and $\Delta\Delta G$ were calculated from the relationships $\Delta\Delta H=\Delta H_{LZ}-\Delta H_{bZIP}$, $T\Delta\Delta S=T\Delta S_{LZ}-T\Delta S_{bZIP}$ and $\Delta\Delta G=\Delta G_{LZ}-\Delta G_{bZIP}$, where the subscripts LZ and bZIP denote the corresponding thermodynamic parameters for the heterodimerization of LZ subdomains and bZIP domains, respectively (Table 4-1). (b) Differential entropic signature for the heterodimerization of LZ subdomains relative to bZIP domains. $\Delta\Delta S$, $\Delta\Delta S_{solv}$ and $\Delta\Delta S_{conf}$ were calculated from the relationships $\Delta\Delta S=\Delta S_{LZ}-\Delta S_{bZIP}$, $\Delta\Delta S_{solv}=\Delta S_{solv(LZ)}-\Delta S_{solv(bZIP)}$ and $\Delta\Delta S_{conf}=\Delta S_{conf(LZ)}-\Delta S_{conf(bZIP)}$, where the subscripts LZ and bZIP denote the corresponding thermodynamic parameters for the heterodimerization of LZ subdomains and bZIP domains, respectively. ΔS_{solv} was calculated from the relationship $\Delta S_{solv}=\Delta C_p \ln[298/385]$ and ΔS_{conf} from the relationship $\Delta S_{conf}=\Delta S-\Delta S_{solv}$ for the heterodimerization of LZ subdomains or bZIP domains with the ΔS and ΔC_p being the corresponding thermodynamic parameters (Tables 4-1 and 4-2).

of $\Delta\Delta G$ implies that the free energy change is more favorable for the heterodimerization of LZ subdomains relative to bZIP domains. As evidenced in Figure 4-4a, the enthalpy of heterodimerization is more favorable by about –1.46 kcal/mol for the LZ subdomains relative to the bZIP domains. This small but clearly more favorable release of heat upon the heterodimerization of LZ subdomains could contribute to the magnitude of affinity of heterodimerization by as much as 10-fold instead of 2-fold relative to bZIP domains observed here. That this is not the case is best understood in terms of the more unfavorable entropic contribution of about -1.17 kcal/mol for the heterodimerization of LZ subdomains relative to the bZIP domains. In other words, the favorable gain of enthalpy for the heterodimerization of LZ subdomains relative to bZIP domains is largely offset by nearly an equal but opposing loss of entropy. In contrast, the unfavorable loss of

enthalpy for the heterodimerization of bZIP domains relative to LZ subdomains is largely counterbalanced by nearly an equal but opposing gain of entropy. While the slightly less favorable enthalpy change observed for the heterodimerization of bZIP domains relative to LZ subdomains could be rationalized in terms of the unfavorable electrostatic repulsions between the basic residues in the BR subdomains, the rationale for the slightly less unfavorable loss of entropy observed for the heterodimerization of bZIP domains relative to LZ subdomains begs further tuition.

In Figure 4-4b, we decompose the additional loss of entropy ($\Delta\Delta S$) upon the heterodimerization of LZ subdomains relative to bZIP domains observed here into its two major constituent components $\Delta\Delta S_{solv}$ and $\Delta\Delta S_{conf}$. In this analysis, a negative value of $\Delta\Delta S$ indicates that the heterodimerization of LZ subdomains is entropically less favorable relative to heterodimerization of bZIP domains, a negative value of $\Delta\Delta S_{solv}$ indicates that the change in solvent entropy upon the heterodimerization of LZ subdomains is less favorable relative to the heterodimerization of bZIP domains, and a positive value of $\Delta\Delta S_{conf}$ indicates that the change in conformational entropy upon the heterodimerization of LZ subdomains is more favorable relative to the heterodimerization of bZIP domains. That $\Delta\Delta S_{solv}$ is less favorable and $\Delta\Delta S_{conf}$ is more favorable for the heterodimerization of LZ subdomains relative to the heterodimerization of bZIP domains suggests strongly that the basic regions undergo at least partial folding upon heterodimerization of bZIP domains and, in so doing, are likely to modulate the heterodimerization of leucine zippers.

4.5 Concluding remarks

Thermodynamics is a powerful tool to gain insights into the energetic components that define protein-protein interactions relevant to biological function. Despite their discovery over two decades ago, the thermodynamics of heterodimerzation of leucine zippers of Jun and Fos hitherto have not been characterized. Our thermodynamic analysis here shows that the heterodimerization of leucine zippers of Jun and Fos is under enthalpic control and accompanied by entropic penalty at physiological temperatures. We have reasoned herein that the nature of the entropic penalty is likely to be largely due to the restriction in the conformational degrees of freedom of the backbone and sidechain atoms upon heterodimerization. We attribute large negative changes in heat capacity observed upon the heterodimerization of Jun and Fos to the formation of α-helical heterodimers from the corresponding unfolded leucine zippers. One additional key finding of our study is that the basic regions in the bZIP domains modulate the heterodimerization of leucine zippers and may undergo some degree of folding though their complete folding may necessitate the binding of DNA in agreement with previous studies [62, 104]. Taken together, our study provides novel insights into the thermodynamics of a key protein-protein interaction pertinent to cellular transcriptional machinery.

5 Chapter 5: Evidence that the bZIP Domains of the Jun Transcription Factor Bind to DNA as Monomers Prior to Folding and Homodimerization

5.1 Summary

The Jun oncoprotein belongs to the AP1 family of transcription factors that is collectively engaged in diverse cellular processes by virtue of their ability to bind to the promoters of a wide spectrum of genes in a DNA sequence-dependent manner. Here, using isothermal titration calorimetry, we report detailed thermodynamics of the binding of bZIP domain of Jun to synthetic dsDNA oligos containing the TRE and CRE consensus promoter elements. Our data suggest that binding of Jun to both sites occurs with indistinguishable affinities but with distinct thermodynamic signatures comprised of favorable enthalpic contributions accompanied by entropic penalty at physiological temperatures. Furthermore, anomalously large negative heat capacity changes observed provoke a model in which Jun loads onto DNA as unfolded monomers coupled with subsequent folding and homodimerization upon association. Taken together, our data provide novel insights into the energetics of a key protein-DNA interaction pertinent to cellular signaling and cancer. Our study underscores the notion that the folding and dimerization of transcription factors upon association with DNA may be a more general mechanism employed in protein-DNA interactions and that the conventional school of thought may need to be re-evaluated.

5.2 Background

The Jun oncoprotein is a component of the transcription factor AP1 (activator protein 1) that couples extracellular information in the form of growth factors, cytokines, hormones and stress to DNA transcription and, in so doing, orchestrates a diverse array of

cellular processes such as cell growth and proliferation, cell cycle regulation, embryonic development and cancer. Jun unleashes its transcriptional activity by virtue of its ability to recognize the pseudo-palindromic TGACTCA and palindromic TGACGTCA consensus sequences found in the promoters of a multitude of genes such as metallothionein IIa, collagenase, interleukin 2 and cyclin D1 either as a homodimer or alternatively as a heterodimer in complex with a related oncoprotein Fos [4, 5]. In recent years, new members of the AP1 family, such as ATF and Maf, that can also act as dimerization partners for Jun have been discovered [125, 126]. The consensus sequences TGACTCA and TGACGTCA, respectively referred to as the TPA (12-O-tetradecanoylphorbol-13-acetate) response element (TRE) and the cAMP response element (CRE), occur with a high frequency in the human genome [39, 82]. Jun is expressed in a wide variety of tissues and is subject to activation by a diverse array of mitogenic inputs, including up-regulation by MAP kinases [26, 27]. Upon activation, Jun can switch on gene transcription via homodimerization or heterodimerization with one of the members of the AP1 family as well as through its co-operation with other transcription factors in the recruitment of the transcriptional machinery to the site of DNA [3, 5, 39, 83, 84]. However, the ability of Jun to heterodimerize with Fos not only changes its specificity but is also believed to substantially enhance its transcriptional activity as demonstrated through the transforming potential of the Jun-Fos heterodimer in a wide variety of mammalian cells [3, 5, 84].

The ability of Jun to recognize DNA sequences at the promoters of specific genes resides in a region that has come to be known as the basic zipper (bZIP) domain (Figure 5-1a). The bZIP domain can be further dissected into two well-defined functional

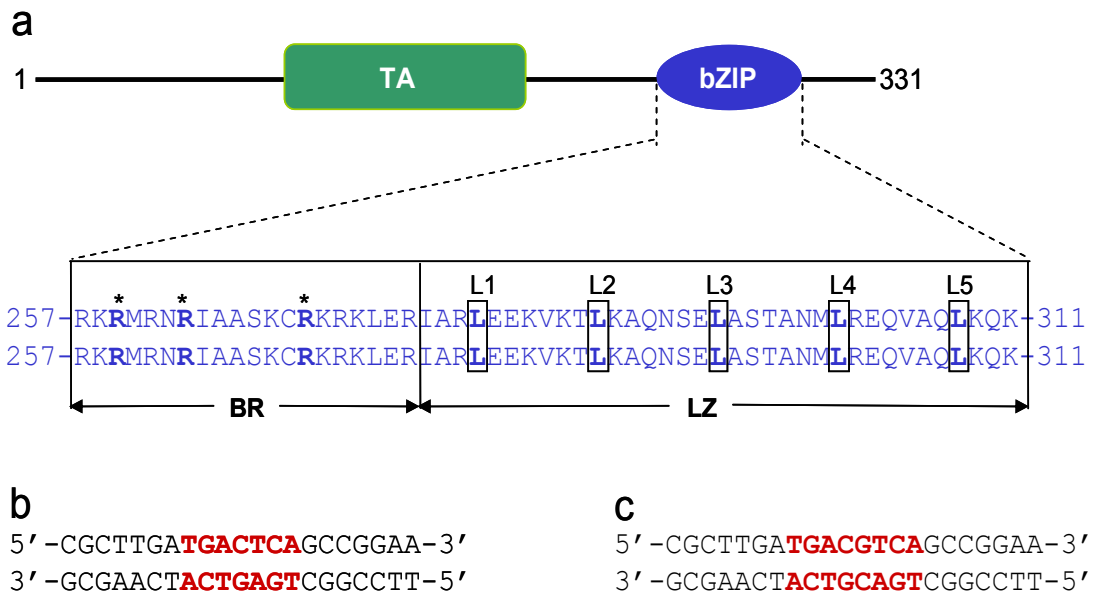

Figure 5-1. Protein and DNA sequences. (a) A schematic showing domain organization of the transcription factor Jun. The basic zipper (bZIP) domain responsible for sequence-specific recognition of DNA at the promoter regions is located C-terminal to the transactivation (TA) domain. The amino acid sequence of the bZIP domain is shown in duplicate to illustrate how two monomers associate to generate a homodimer. The sequence boundaries of the N-terminal basic region (BR) and the C-terminal leucine zipper (LZ) subdomains are also demarcated. The five signature leucines (L1-L5) characteristic of LZ subdomains, spaced exactly six residues apart, are boxed and bold faced. The basic residues in the BR subdomains that contact the DNA bases and the backbone phosphates are indicated by asterisks and bold faced. (b) Nucleotide sequence of 21-mer dsDNA oligo containing the TRE site (bold faced and colored red). (c) Nucleotide sequence of 22-mer dsDNA oligo containing the CRE site (bold faced and colored red).

subdomains termed the basic region (BR) at the N-terminus followed by the leucine zipper (LZ) at the C-terminus. The leucine zipper is a highly conserved protein module found in a wide variety of transcription factors and structural proteins and contains a signature leucine at every seventh position within the five successive heptads of amino acid residues. The leucine zippers adapt continuous α-helical conformations and induce Jun-Jun homodimerization or its heterodimerization with other members of AP1 family by virtue of their ability to wrap around each other in a coiled coil dimer [41, 42]. Such intermolecular arrangement brings the basic regions at the N-termini of bZIP domains into close proximity and thereby enables them to insert into the major grooves of DNA at the promoter regions in an optimal fashion in a manner akin to a pair of forceps. While

the α-helices are held together by numerous inter-helical hydrophobic contacts and salt bridges, hydrogen bonding between the sidechains of basic residues in the basic regions and the sidechains of nucleotides accounts for high affinity binding of bZIP domains to DNA [5, 41, 43]. It is widely believed that the basic regions within the bZIP domains are unstructured in the absence of DNA and undergo folding to adapt α-helical conformations only upon DNA binding [104]. This notion was further confirmed by our recent thermodynamic studies on the binding of Jun-Fos heterodimer to DNA, wherein we reasoned that the heat capacity changes accompanying this interaction could only be explained by the coupling of folding of basic regions upon association with DNA [122]. This previous study also demonstrated that the binding of DNA to Jun-Fos heterodimer is largely driven by favorable enthalpic changes accompanied by entropic penalty at physiological temperatures.

In an effort to further our understanding of the relationship between structure and thermodynamics governing the binding of the members of the AP1 family to DNA, we have employed isothermal titration calorimetry (ITC) to delineate the thermodynamics of the binding of the bZIP domain of Jun to synthetic dsDNA oligos containing the TRE and CRE consensus promoter elements. Our data suggest that binding of Jun to both sites occurs with indistinguishable affinities but with distinct thermodynamic signatures comprised of favorable enthalpic contributions accompanied by entropic penalty at physiological temperatures. Furthermore, anomalously large negative heat capacity changes observed provoke a model in which Jun loads onto DNA as unfolded monomers coupled with subsequent folding and homodimerization upon association. Taken together, our data provide novel insights into the energetics of a key protein-DNA

interaction pertinent to cellular signaling and cancer. Our study underscores the notion that the folding and dimerization of transcription factors upon association with DNA may be a more general mechanism employed in protein-DNA interactions and that the conventional school of thought may need to be re-evaluated.

5.3 Experimental procedures

5.3.1 Sample preparation

bZIP domain of human Jun (residues 251-331) was cloned into pET102 bacterial expression vector, with an N-terminal thioredoxin (Trx)-tag and a C-terminal polyhistidine (His)-tag, using Invitrogen TOPO technology. The recombinant protein was expressed, purified and characterized as described previously [122]. Total monomeric concentration of bZIP domain was determined by the fluorescence-based Quant-It assay (Invitrogen) and spectrophotometrically using a extinction co-efficient of $14{,}230 M^{-1}cm^{-1}$ at 280nm. The extinction co-efficient was calculated using the online software ProtParam at ExPasy Server [86]. Results from both methods were in an excellent agreement. HPLC-grade DNA oligos containing the consensus TRE and CRE sites were commercially obtained from Sigma Genosys. The complete nucleotide sequences of these oligos are presented in Figures 5-1b and 5-1c. Oligo concentrations were determined spectrophotometrically on the basis of their extinction co-efficients derived from their nucleotide sequences using the online software OligoAnalyzer 3.0 (Integrated DNA Technologies) based on the nearest-neighbor model [68]. Double-stranded DNA (dsDNA) annealed oligos were generated as described previously [122].

5.3.2 ITC measurements

Isothermal titration calorimetry (ITC) experiments were performed on Microcal VP-ITC instrument and data were acquired and processed using fully automized features in Microcal ORIGIN software. All measurements were repeated 2-3 times. Briefly, the bZIP domain of Jun and dsDNA oligos were prepared in 50mM Tris, 200mM NaCl, 1mM EDTA and 5mM β-mercaptoethanol at pH 8.0. Because of its low stability as stated earlier [122], the bZIP domain of Jun was concentrated to about 10μM using the Amicon Ultra-15 centrifugal filter units (MWCO 10kD) immediately prior to running the experiments. All samples were de-gassed using the ThermoVac accessory for 10min. The experiments were initiated by injecting 20 x 10μl injections of 50-100μM of dsDNA oligo from the syringe into the calorimetric cell containing 1.8ml of 5-10μM of the dimer-equivalent bZIP domain at a fixed temperature in the narrow range 15-35°C. The change in thermal power as a function of each injection was automatically recorded using Microcal ORIGIN software and the raw data were further processed to yield binding isotherms of heat release per injection as a function of DNA to protein molar ratio. The heats of mixing and dilution were subtracted from the heat of binding per injection by carrying out a control experiment in which the same buffer in the calorimetric cell was titrated against the dsDNA oligos in an identical manner. Control experiments with scrambled dsDNA oligos generated similar thermal power to that obtained for the buffer alone — as did the titration of dsDNA oligos containing TRE and CRE sites against a protein construct containing thioredoxin with a C-terminal His-tag (Trx-His). Titration of concentrated Trx-His protein construct into the calorimetric cell containing the bZIP domain produced no observable signal, implying that neither Trx-tag nor His-tag interact

with the bZIP domain of Jun. To extract the binding affinity (K_d) and the binding enthalpy (ΔH), the binding isotherms were iteratively fit to the following built-in function by non-linear least squares regression analysis using the integrated Microcal ORIGIN software:

$$q(i) = (n\Delta HVP/2)\ \{[1+(L/nP)+(K_d/nP)] - [[1+(L/nP)+(K_d/nP)]^2 - (4L/nP)]^{1/2}\} \qquad [1]$$

where q(i) is the heat release (kcal/mol) for the ith injection, n is the binding stoichiometry, V is the effective volume of protein solution in the calorimetric cell (1.46ml), P is the total protein concentration in the calorimetric cell (μM) and L is the total concentration of DNA added (μM). Given that Jun monomers freely exchange with their dimeric counterparts at equilibrium, P was assumed to be equivalent to the total experimentally measured monomeric concentration of bZIP domain. It should be noted that the above equation is derived from the binding of a ligand to a macromolecule using the law of mass action [74].

5.3.3 SASA calculations

The magnitude of changes in polar and apolar solvent-accessible surface area (SASA) in the bZIP domain of Jun upon binding to dsDNA oligos containing the TRE and CRE consensus sites were calculated from thermodynamic data obtained using ITC and compared with those obtained from structural data based on the 3D structural models (see below). For calculation of changes in polar SASA ($\Delta SASA_{polar}$) and apolar SASA ($\Delta SASA_{apolar}$) upon the binding of dsDNA oligos containing the TRE and CRE consensus sites to bZIP domain of Jun from thermodynamic data, it was assumed that ΔC_p and ΔH at 60°C (ΔH_{60}) are additive and linearly depend on the change in $\Delta SASA_{polar}$ and

$\Delta SASA_{apolar}$ as embodied in the following empirically-derived expressions [70, 109, 127, 128]:

$$\Delta C_p = a[\Delta SASA_{polar}] + b[\Delta SASA_{apolar}] \quad [2]$$

$$\Delta H_{60} = c[\Delta SASA_{polar}] + d[\Delta SASA_{apolar}] \quad [3]$$

where a, b, c and d are empirically-determined co-efficients with values of –0.26 cal/mol/K/Å^2, +0.45 cal/mol/K/Å^2, +31.34 cal/mol/Å^2 and –8.44 cal/mol/Å^2, respectively. The co-efficients a and b are independent of temperature, while c and d are referenced against a temperature of 60°C, which equates to the median melting temperature of the proteins from which these constants are derived [70, 109, 127, 128]. ΔC_p was calculated from the slope of a plot of ΔH versus temperature in the narrow temperature range 15-35°C for the binding of TRE and CRE dsDNA oligos to the bZIP domain of Jun using the ITC instrument. ΔH_{60} was calculated by the extrapolation of a plot of ΔH versus temperature to 60°C for the binding of TRE and CRE dsDNA oligos to the bZIP domain of Jun using the ITC instrument. With ΔC_p and ΔH_{60} experimentally determined using ITC and the knowledge of co-efficients a-d from empirical models [69-73], equations [2] and [3] were simultaneously solved to obtain the magnitude of changes in $\Delta SASA_{polar}$ and $\Delta SASA_{apolar}$ independent of structural information upon the binding of dsDNA oligos to the bZIP domain of Jun.

To determine changes in $\Delta SASA_{polar}$ and $\Delta SASA_{apolar}$ upon the binding of dsDNA oligos containing TRE and CRE sequences to bZIP domain of Jun from structural data, three models of binding were assumed — the Lock-and-Key (LK) model, the Induced Fit (IF) model and the Equilibrium Shift (ES) model. In the LK model, it was assumed that the bZIP domains exist as fully folded homodimers and undergo no conformational

change upon DNA binding — that is the homodimers exist in a pre-formed conformation that best fits the DNA. In the IF model, it was assumed that the bZIP domains exist as partially folded homodimers in which the basic regions are fully unstructured and only become structured upon DNA binding — that is DNA binding induces the folding of basic regions within otherwise pre-formed homodimers. In the ES model, it was assumed that the fully folded and the partially folded bZIP homodimers exist in equilibrium with the fully unfolded bZIP monomers and that DNA only binds to the monomers resulting in their folding and homodimerization — that is the bZIP domains bind to DNA as unfolded monomers such that their folding and homodimerization in association with DNA shifts the equilibrium with fully folded and partially folded homodimers in their direction. Changes in $\Delta SASA_{polar}$ and $\Delta SASA_{apolar}$ upon the binding of dsDNA oligos to bZIP domain of Jun from structural data were calculated using the following relationships:

$$\Delta SASA_{polar} = SASA_{bp} - (SASA_{fp} + SASA_{dp}) \quad [4]$$

$$\Delta SASA_{apolar} = SASA_{ba} - (SASA_{fa} + SASA_{da}) \quad [5]$$

where $SASA_{bp}$ and $SASA_{ba}$ are the polar and apolar SASA of bZIP homodimers bound to DNA, $SASA_{fp}$ and $SASA_{fa}$ are the polar and apolar SASA of fully folded bZIP homodimers alone, or partially folded bZIP homodimers alone, or fully unfolded bZIP monomers alone, and $SASA_{dp}$ and $SASA_{da}$ are the polar and apolar SASA of dsDNA oligos alone. For all three above-mentioned models of binding, $SASA_{bp}$ and $SASA_{ba}$ were calculated from structural models of bZIP homodimer in complex with dsDNA oligos containing atomic coordinates of both the bZIP domains and the corresponding sense and antisense dsDNA oligos, while $SASA_{dp}$ and $SASA_{da}$ were calculated from the same structural models of bZIP homodimer in complex with dsDNA oligos but

containing atomic coordinates of only the corresponding sense and antisense dsDNA oligos. For the LK model, $SASA_{fp}$ and $SASA_{fa}$ were calculated from structural models of bZIP homodimer in complex with dsDNA oligos but containing atomic coordinates of the bZIP domains only. For the IF model, $SASA_{fp}$ and $SASA_{fa}$ were calculated from structural models of bZIP homodimer determined in the absence of DNA with basic regions allowed to adopt unfolded conformations (see below). For the ES model, $SASA_{fp}$ and $SASA_{fa}$ were calculated from structural models of bZIP monomers with compact unfolded conformations. All SASA calculations were performed using the online software GETAREA with a probe radius of 1.4Å [87].

5.3.4 Structural analysis

3D structures of bZIP domains of Jun as fully unfolded monomers alone, as partially folded homodimers alone and as fully folded homodimers bound to dsDNA oligos containing TRE and CRE sites were modeled using the MODELLER software based on homology modeling [76]. The model of bZIP domains of Jun-Jun homodimer in complex with 21-mer dsDNA oligo containing the TRE site was obtained using the crystal structure of bZIP domains of Jun-Jun homodimer in complex with a dsDNA oligo containing the TRE consensus sequence TGACTCA but differing in flanking sequences as a template (with a PDB code of 2H7H). The model of bZIP domains of Jun-Jun homodimer in complex with 22-mer dsDNA oligo containing the CRE site was obtained using the crystal structure of bZIP domains of Jun-Jun homodimer in complex with a dsDNA oligo containing the CRE consensus sequence TGACGTCA but differing in flanking sequences as a template (with a PDB code of 1JNM). The model of bZIP domains of Jun-Jun as a partially folded homodimer in the absence of DNA was obtained

using the crystal structure of leucine zippers of Jun-Jun homodimer alone as a template (with a PDB code of 1JUN) and the residues in the N-terminal basic regions were allowed to adopt an open compact conformation and allowed to reach the energy minima without any restraints. The models of bZIP domains of Jun as fully unfolded monomers were obtained without a template with all residues allowed to adopt an open compact conformation and allowed to reach the energy minima without any restraints. 3D structural models of bZIP domains of Jun-Fos heterodimer in complex with dsDNA oligos containing TRE and CRE sites were modeled as described previously [122]. In each case, a total of 100 structural models were calculated and the structure with the lowest energy, as judged by the MODELLER Objective Function, was selected for further energy minimization in MODELLER prior to analysis. The structures were rendered using RIBBONS [77]. All other calculations were performed on the lowest energy-minimized structural model.

5.4 Results and discussion

5.4.1 Jun-Jun homodimer binds to TRE and CRE with indistinguishable affinities but with distinct thermodynamic signatures

In an attempt to unravel the thermodynamic mechanism of the binding of bZIP domains of Jun-Jun homodimer to dsDNA oligos containing the TRE and CRE sites, we employed the technique of ITC (Figure 5-2). Comparison of the various thermodynamic parameters is presented in Table 5-1. Our data suggest that the bZIP domains of Jun-Jun homodimer bind to TRE and CRE sites with virtually indistinguishable affinities. The notion that protein-ligand interactions cannot be merely understood in terms of their binding affinities could not be more applicable to the system being scrutinized here. Indeed, decomposition of the apparent binding affinities of TRE and CRE to Jun-Jun

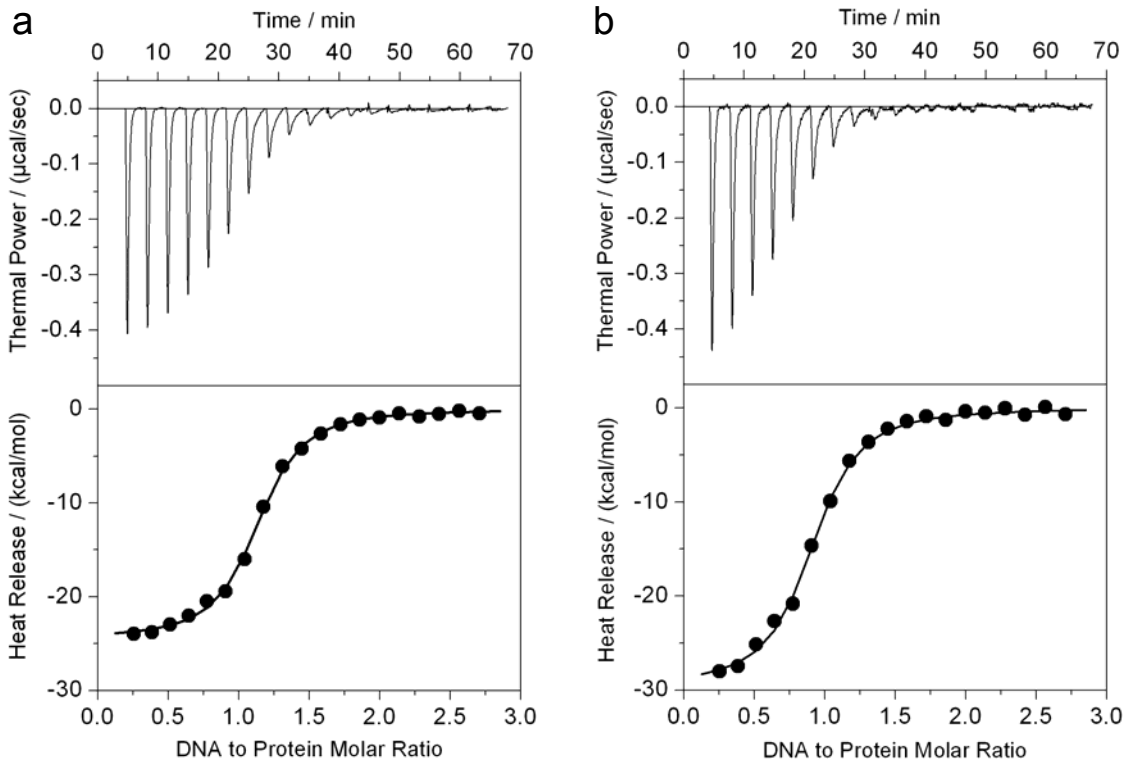

Figure 5-2. ITC analysis of the binding of the bZIP domain of Jun to dsDNA oligos containing TRE (a) and CRE (b) consensus promoter sites. bZIP domain of Jun in the calorimetric cell was titrated with 25 x 10µl injections of dsDNA oligo from the injection syringe at 25°C and pH 8.0. The first injection and the corresponding heat release are not shown due to systematic uncertainties in the measurement. The top panels show the raw ITC data describing the change in thermal power as a function of time upon subsequent injections. The raw data were processed to generate the binding isotherms of heat release per injection as a function of increasing DNA to protein molar ratio as shown in the bottom panels. The solid lines represent the fit of the data in the binding isotherms to the function based on the binding of a ligand to a macromolecule using the Microcal ORIGIN software [74].

homodimer into their corresponding enthalpic (ΔH) and entropic (TΔS) components suggests that while both interactions are under strong enthalpic control accompanied by entropic penalty, these underlying forces contribute non-equally but in an opposing manner to the overall free energy of binding. Thus, while the binding of CRE to Jun-Jun homodimer is enthalpically more favorable by about +6 kcal/mol relative to TRE, binding of the latter is accompanied by roughly an equal but opposing gain of entropy relative to CRE, resulting in no overall differences in the affinity of these two DNA promoter elements to Jun-Jun homodimer. Although various attempts have been made in

Table 5-1. Experimentally determined thermodynamic parameters for the binding of bZIP domain of Jun to dsDNA oligos containing TRE and CRE consensus sequences obtained from ITC measurements at 25°C and pH 8.0

	K_d (μM)	ΔH (kcalmol^{-1})	$T\Delta S$ (kcalmol^{-1})	ΔG (kcalmol^{-1})
TRE	0.06 ± 0.01	−24.23 ± 0.17	−14.30 ± 0.17	−9.92 ± 0.01
CRE	0.07 ± 0.02	−29.73 ± 0.59	−19.98 ± 0.44	−9.74 ± 0.14

The values for the binding affinity (K_d) and the binding enthalpy (ΔH) were obtained from the fit of a function as given in expression [1], based on the binding of a ligand to a macromolecule using the law of mass action assuming a 1:1 binding stoichiometry [74], to the ITC isotherms shown in Figure 5-2. Free energy of binding (ΔG) was calculated from the relationship $\Delta G=RT\ln K_d$, where R is the universal molar gas constant (1.99 cal/mol/K) and T is the absolute temperature in Kelvins. Entropic contribution (TΔS) to binding was calculated from the relationship TΔS=ΔH-ΔG. The binding stoichiometries to the fits agreed to within ±10%. Errors were calculated from 2-3 independent measurements. All errors are given to one standard deviation.

the past to obtain binding constants on the basis of non-continuous and non-quantitative methods [42, 49, 56, 57, 89], this is the first study to not only report quantitative binding data but also detailed energetics governing the interaction of Jun-Jun homodimer to DNA.

What might be the molecular basis of favorable enthalpic change accompanied by entropic penalty for the interaction of TRE and CRE with Jun-Jun homodimer? The favorable enthalpic change is most likely due to the formation of hydrogen bonding, hydrophobic contacts and electrostatic interactions between the bZIP domains of Jun-Jun homodimer and its target DNA duplexes as observed in the 3D structure [43] — see also unpublished structures with PDB codes 1JNM and 2H7H. The more favorable enthalpic change of about +6 kcal/mol for the interaction CRE versus TRE to Jun-Jun homodimer may be attributed to the presence of an extra base pair between the TGA and TCA half-sites in CRE. It should be noted here that enthalpy-driven nature of protein-DNA interactions observed here is neither a rule nor an exception as numerous examples of protein-DNA interactions under enthalpic as well as entropic control have been reported previously [90, 93-100]. Unlike the molecular basis for favorable enthalpic change, the

rationale for the entropic penalty encountered here is less intuitive. It is widely believed that the net entropic change upon molecular associations largely results from interplay between two opposing entropic components — ΔS_{solv} and ΔS_{conf}. The ΔS_{solv} is the favorable entropy change due to enhancement in the degrees of freedom of solvent molecules as a result of their restructuring and displacement, particularly around apolar groups, upon molecular associations. In contrast, the ΔS_{conf} is the unfavorable entropic change that arises from the restriction of conformational degrees of freedom of the backbone and sidechain atoms upon molecular associations. It has been suggested that the basic regions in the bZIP domains are unstructured in the absence of DNA and undergo folding only upon DNA binding [104]. Thus, such restructuring of protein upon DNA binding could further negatively contribute to the ΔS_{conf}. Furthermore, it is believed that DNA also experiences some degree of bending and hence loss in conformational freedom upon binding to bZIP domains [58, 118]. On the basis of the foregoing arguments, we attribute the unfavorable entropic change observed here upon the binding of bZIP domains of Jun-Jun homodimer to DNA largely to the loss of conformational degrees of freedom of backbone and sidechain atoms in both the protein and DNA as embodied in the term ΔS_{conf}. The less favorable entropic change of about -6 kcal/mol for the interaction of CRE versus TRE to Jun-Jun homodimer may be attributable to the presence of an extra base pair between the TGA and TCA half-sites in CRE.

5.4.2 Enthalpy and entropy compensate the effect of temperature on the binding of DNA to Jun-Jun homodimer

In an effort to determine the effect of temperature on the various thermodynamic parameters, we analyzed the binding of the bZIP domains of Jun-Jun homodimer to dsDNA oligos containing the TRE and CRE consensus sites in the temperature range 15-

35°C (Figure 5-3). Our data indicate that both the enthalpic (ΔH) and entropic (TΔS) contributions to the overall free energy of binding (ΔG) show strong temperature-dependence and that both ΔH and TΔS largely compensate for each other to generate ΔG that is virtually independent of temperature. Thus, while ΔH and TΔS experience more than 20 kcal/mol change in their contributions to binding in going from 15°C to 35°C, ΔG varies no more than 1 kcal/mol over the same temperature range. Consistent with this observation is the relatively constant nature of the binding affinity (0.05-0.15μM) over the same temperature range for the interaction of both the TRE and CRE dsDNA oligos with the bZIP domains of Jun-Jun homodimer. The linear and opposing dependence of ΔH and TΔS as a function of temperature, while maintaining a more or less constant ΔG, is a common feature observed in protein folding and binding reactions [70, 108, 109]. This phenomenon gives rise to two key temperature points T_H and T_S — the temperatures where enthalpic (ΔH) and entropic (TΔS) contributions to the free energy of binding change sign, respectively. Thus, in the case of the binding of DNA to Jun-Jun homodimer, ΔH will become negative and hence thermodynamically favorable above T_H, while TΔS will become negative and hence thermodynamically unfavorable above T_S. Table 5-2 provides the values for T_H and T_S accompanying the binding of DNA to the bZIP domains of Jun-Jun homodimer. As evidenced in Table 5-2, both T_H and T_S fall well below the physiological temperature of 37°C, implying that the binding of Jun-Jun homodimer to DNA will be largely under enthalpic control accompanied by entropic penalty at physiological temperatures.

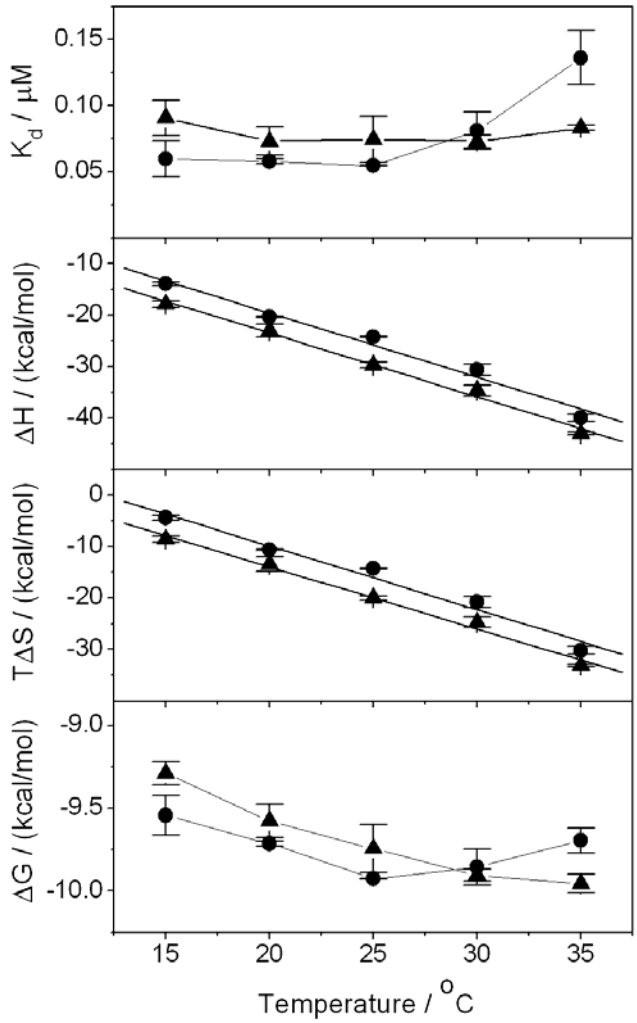

Figure 5-3. Dependence of thermodynamic parameters K_d, ΔH, TΔS and ΔG on temperature for the binding of bZIP domain of Jun to dsDNA oligos containing TRE (●) and CRE (▲) sites. bZIP domain of Jun in the calorimetric cell was titrated with 25 x 10μl injections of dsDNA oligo from the injection syringe at various temperatures in the range 15-35°C at pH 8.0. To determine the various thermodynamic parameters, the binding isotherms were fit to the function based on the binding of a ligand to a macromolecule using the Microcal ORIGIN software [74]. Each data point is the arithmetic mean of 2-3 experiments. All error bars are given to one standard deviation. The solid lines for the ΔH and TΔS plots show linear fits to the data, while the solid lines for the K_d and ΔG plots show straight lines connecting the data points for clarity.

The temperature-dependence of ΔH is related to heat capacity of binding (ΔC_p) by Kirchhoff's relationship $\Delta C_p = d(\Delta H)/dT$. In other words, the slope of a plot of ΔH versus temperature equates to ΔC_p. Heat capacity is an important thermodynamic parameter in that it is related to the extent of the burial and dehydration of molecular surfaces from

Table 5-2. Experimentally determined thermodynamic parameters for the binding of bZIP domain of Jun to dsDNA oligos containing TRE and CRE consensus sequences obtained from ITC measurements at various temperatures in the range 15-35°C and pH 8.0

	T_H (°C)	T_S (°C)	ΔH_{60} (kcalmol^{-1})	ΔC_p (kcalmol^{-1}K^{-1})
TRE	−4.27 ± 0.28	+12.01 ± 0.01	−69.39 ± 1.71	−1.25 ± 0.04
CRE	−1.03 ± 1.50	+8.41 ± 1.16	−73.08 ± 1.82	−1.24 ± 0.06

The values for the various parameters shown were obtained as follows. The values for T_H, the temperature at which ΔH is zero, were obtained from the extrapolation of linear fits to the ΔH versus temperature plots (Figure 5-3). The values for T_S, the temperature at which TΔS is zero, were obtained from the extrapolation of linear fits to the TΔS versus temperature plots (Figure 5-3). The values for ΔH_{60}, the enthalpy at 60°C, were obtained from the extrapolation of linear fits to the ΔH versus temperature plots (Figure 5-3). The values for ΔC_p, the heat capacity change, were obtained from the slopes of linear fits to the ΔH versus temperature plots (Figure 5-3). Errors were calculated from 2-3 independent measurements. All errors are given to one standard deviation.

surrounding solvent molecules upon intermolecular association. This concept has come to be referred to as the change in solvent-accessible surface area (ΔSASA) [70, 94, 110, 111]. In an attempt to understand how the binding of Jun-Jun homodimer to DNA affects SASA, we calculated ΔC_p values hovering around –1200 cal/mol/K from the slopes of ΔH versus temperature plots for the binding of bZIP domains of Jun-Jun homodimer to TRE and CRE sites (Figure 5-3 and Table 5-2). What might be the significance of such large negative values of ΔC_p observed here? A positive value of ΔC_p implies that the occlusion of polar surfaces dominates the intermolecular association over apolar surfaces [70, 112, 113]. The fact that ΔC_p is accompanied by large negative changes suggests strongly that the binding of bZIP domains of Jun-Jun homodimer to DNA involves substantial burial of hydrophobic residues with little contributions from polar residues. It should be noted here that protein-ligand interactions typically result in the magnitude of ΔC_p of less than –1000 cal/mol/K, while values of ΔC_p in the range –1000 to -2000 cal/mol/K are characteristic of proteins undergoing folding due to burial of a large number of apolar groups as a result of hydrophobic effect. Could the rather large negative

values of ΔC_p observed here reflect the plausible scenario that folding and homodimerization of bZIP domains of Jun may be coupled to DNA binding?

5.4.3 Jun-Jun homodimer binds to DNA with higher affinity than Jun-Fos heterodimer but the latter harbors more favorable enthalpic change

Our thermodynamic data reported here for the binding of Jun-Jun homodimer to DNA are in sharp contrast to the thermodynamic data reported previously for the binding of Jun-Fos heterodimer to DNA [122]. While Jun-Jun homodimer is observed to bind to TRE and CRE sites with an affinity of between 0.06-0.07μM here, the Jun-Fos heterodimer was observed to do so with an affinity of between 0.15-0.21μM in the previous study [122]. The binding of Jun-Jun homodimer to DNA has been widely believed to be weaker than that of Jun-Fos heterodimer on the basis of non-continuous and non-quantitative measurements [5, 42, 49, 56, 57, 89]. Our present study however suggests that this is not the case but, on the contrary, it is the Jun-Jun homodimer that binds to DNA with an affinity that is over 2-fold greater than that of Jun-Fos heterodimer. Comparison of the differential binding affinities of Jun-Jun homodimer versus the Jun-Fos heterodimer to their cognate DNA sequences only provides a glimpse of the complete contrast in the thermodynamic picture of these key protein-DNA interactions.

In Figure 5-4a, we present the differential thermodynamic signatures for the binding of TRE and CRE to Jun-Jun homodimer relative to Jun-Fos heterodimer on the basis of the data presented here and those reported earlier [122]. In this plot, a positive value of $\Delta\Delta H$ implies that the enthalpy change is less favorable for the binding of DNA to Jun-Jun homodimer relative to Jun-Fos heterodimer, while a positive value of $T\Delta\Delta S$ is indicative of favorable gain of entropy for the binding of DNA to Jun-Jun homodimer relative to Jun-Fos heterodimer. Thus, as evidenced, the binding of Jun-Fos heterodimer

to DNA is enthalpically more favorable by about +7 kcal/mol relative to Jun-Jun homodimer. However, this is slightly more than offset by a gain of about +8 kcal/mol of favorable entropic change for the binding of Jun-Jun homodimer to DNA relative to Jun-Fos heterodimer resulting in an overall enhanced binding of the former transcription factor. Assuming that the overall entropic change results from two major opposing entropic forces, namely ΔS_{solv} and ΔS_{conf}, we further decomposed the overall favorable entropic gain of about +8 kcal/mol for the binding of Jun-Jun homodimer to DNA relative to Jun-Fos heterodimer into its constituent components to generate a plot of differential entropic signatures (Figure 5-4b). In this plot, a positive value of $\Delta\Delta S$ implies that the entropy is more favorable for the binding of DNA to Jun-Jun homodimer relative to Jun-Fos heterodimer, a positive value of $\Delta\Delta S_{solv}$ indicates that the change in solvent entropy is more favorable for the binding of DNA to Jun-Jun homodimer relative to Jun-Fos heterodimer, and a negative value of $\Delta\Delta S_{conf}$ demonstrates that the change in conformational entropy harbors greater entropic penalty for the binding of DNA to Jun-Jun homodimer relative to Jun-Fos heterodimer. As shown, the binding of DNA to Jun-Jun homodimer leads to a favorable change in the solvent entropy of about 100 cal/mol/K but this is largely offset by a negative contribution to change in the conformational entropy of about –75 cal/mol/K. That this is so suggests strongly that the binding of the Jun-Jun homodimer to DNA is accompanied by a large conformational change in the protein relative to Jun-Fos heterodimer. We will elaborate on the molecular nature of this conformational change in the latter part of this study but, for now, we turn our attention

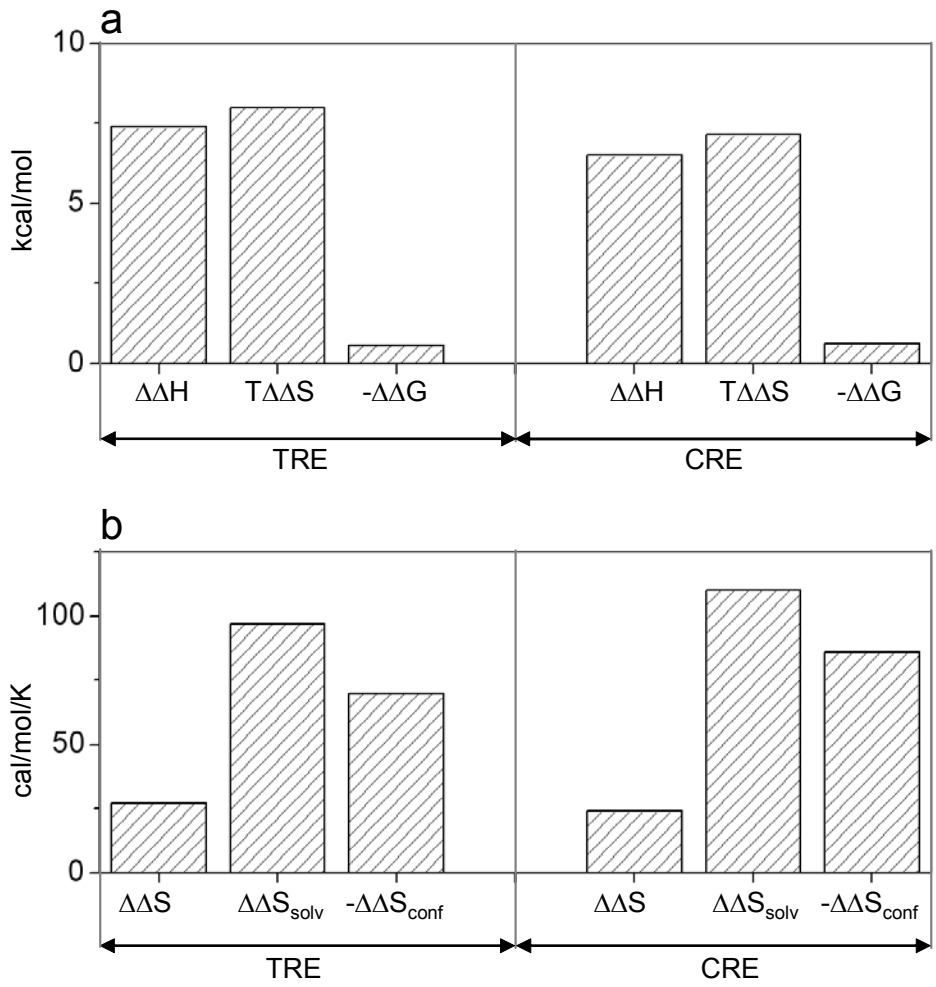

Figure 5-4. Differential energetics for the binding of TRE and CRE dsDNA oligos to Jun-Jun homodimer versus Jun-Fos heterodimer. (a) Differential thermodynamic signatures for the binding of DNA to Jun-Jun homodimer relative to Jun-Fos heterodimer. ΔΔH, TΔΔS and ΔΔG were calculated from the relationships $\Delta\Delta H=\Delta H_{jj}-\Delta H_{jf}$, $T\Delta\Delta S=T\Delta S_{jj}-T\Delta S_{jf}$ and $\Delta\Delta G=\Delta G_{jj}-\Delta G_{jf}$, where the subscripts jj and jf denote the corresponding thermodynamic parameters for the binding of DNA to Jun-Jun homodimer and Jun-Fos heterodimer, respectively. (b) Differential entropic signatures for the binding of DNA to Jun-Jun homodimer relative to Jun-Fos heterodimer. ΔΔS, $\Delta\Delta S_{solv}$ and $\Delta\Delta S_{conf}$ were calculated from the relationships $\Delta\Delta S=\Delta S_{jj}-\Delta S_{jf}$, $\Delta\Delta S_{solv}=\Delta S_{solv(jj)}-\Delta S_{solv(jf)}$ and $\Delta\Delta S_{conf}=\Delta S_{conf(jj)}-\Delta S_{conf(jf)}$, where the subscripts jj and jf denote the corresponding thermodynamic parameters for the binding of DNA to Jun-Jun homodimer and Jun-Fos heterodimer, respectively. ΔS_{solv} was calculated from the relationship $\Delta S_{solv}=\Delta C_p \ln[298/385]$ and ΔS_{conf} from the relationship $\Delta S_{conf}=\Delta S-\Delta S_{solv}$ for the binding of DNA to Jun-Jun homodimer or Jun-Fos heterodimer with the ΔS and ΔC_p being the corresponding thermodynamic parameters. Thermodynamic parameters for the binding of DNA to Jun-Jun homodimer are reported here, while those for the binding of DNA to Jun-Fos heterodimer were reported in the earlier study [122].

to delineating the molecular basis of the more favorable enthalpic change of about +7 kcal/mol observed here for the binding of DNA to Jun-Fos heterodimer relative to Jun-Jun homodimer (Figure 5-4a).

Given that the enthalpic change largely results from the formation of hydrogen bonding, hydrophobic contacts and electrostatic interactions between molecular surfaces, we reasoned that the enthalpically more favorable binding of DNA to the Jun-Fos heterodimer relative to Jun-Jun homodimer may be a manifestation of the differential burial of nucleotides against the amino acid residues in the basic regions of Jun-Fos heterodimer versus the Jun-Jun homodimer due to the variation in the amino acid sequence of the two transcription factors. In Figure 5-5, we present the differential burial of the nucleotides corresponding to the consensus sequences TGACTCA and TGACGTCA for TRE and CRE, respectively, upon the binding of DNA to Jun-Jun homodimer relative to Jun-Fos heterdimer. In these plots, a positive value of ΔSASA for a given nucleotide implies that it is buried more in association with Jun-Fos heterodimer relative to Jun-Jun homodimer, while a negative value of ΔSASA for a given nucleotide is indicative of greater burial in association with Jun-Jun homodimer relative to Jun-Fos heterodimer. Thus, for example, adenosine in the TGA half-site within the TRE sense strand buries more surface area in contact with Jun-Jun homodimer relative to Jun-Fos heterodimer, while cytidine in the TCA half-site within the TRE sense strand buries more surface area in contact with Jun-Fos heterodimer relative to Jun-Jun homodimer. Further differences in the extent to which nucleotides are buried upon interaction with Jun-Jun homodimer versus the Jun-Fos heterodimer can be found throughout the consensus sites of TRE and CRE. In short, the differential enthalpic changes observed upon the binding of DNA to Jun-Jun homodimer relative to Jun-Fos heterodimer may be attributable to the differential burial of nucleotides and amino acid residues upon their association.

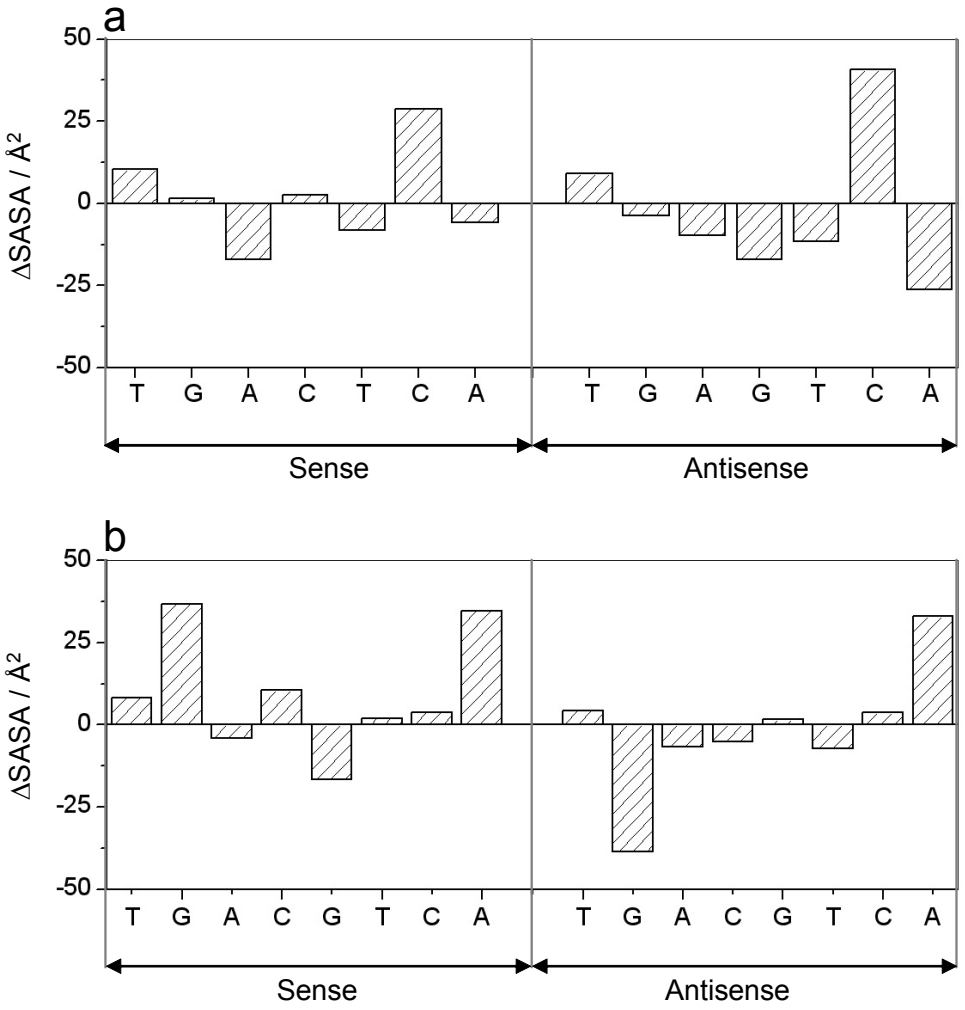

Figure 5-5. Differential changes in SASA for the binding of TRE and CRE dsDNA oligos to Jun-Jun homodimer versus Jun-Fos heterodimer. (a) Differential changes in SASA observed for each nucleotide in the sense and antisense strands of TRE consensus site TGACTCA. ΔSASA for each nucleotide was calculated using the relationship $\Delta SASA=SASA_{jj}-SASA_{jf}$, where the subscripts jj and jf denote the SASA observed for each nucleotide for the binding of TRE to the bZIP domains of Jun-Jun homodimer and Jun-Fos heterodimer, respectively. (b) Differential changes in SASA observed for each nucleotide in the sense and antisense strands of CRE consensus site TGACGTCA. ΔSASA for each nucleotide was calculated using the relationship $\Delta SASA=SASA_{jj}-SASA_{jf}$, where the subscripts jj and jf denote the SASA observed for each nucleotide for the binding of CRE to the bZIP domains of Jun-Jun homodimer and Jun-Fos heterodimer, respectively.

5.4.4 Jun binds to DNA as a monomer with coupled folding and homodimerization of bZIP domains upon association

Experimental determination of values of ΔC_p combined with ΔH_{60} (enthalpy change at 60°C) have been widely used to quantitatively calculate changes in polar SASA ($\Delta SASA_{polar}$), apolar SASA ($\Delta SASA_{apolar}$) and total SASA ($\Delta SASA_{total}$) upon

intermolecular association [69-73] (Table 5-2). Such changes in SASA upon the binding of bZIP domains of Jun-Jun homodimer to DNA from our thermodynamic measurements are reported in Table 5-3. To rationalize what these numbers mean in terms of the mechanism of the protein-DNA interaction under scrutiny here, we also determined changes in SASA upon the binding of the bZIP domains of Jun to DNA from structural data independent of our thermodynamic measurements. To calculate such changes in SASA from structural data, we assumed three models of binding — the Lock-and-Key (LK) model, the Induced Fit (IF) model and the Equilibrium Shift (ES) model (Figure 5-6). In the LK model, it was assumed that the bZIP domains exist as fully folded homodimers and undergo no conformational change upon DNA binding — that is the homodimers exist in a pre-formed conformation that best fits the DNA. Being the simplest and the classical model of protein-ligand interactions, the logic for the consideration of LK model needs no further light. In the IF model, it was assumed that the bZIP domains exist as partially folded homodimers in which the basic regions are fully unstructured and only become structured upon DNA binding — that is DNA binding induces the folding of basic regions within otherwise pre-formed homodimers. The justification for the IF model arises from the salient observation that the basic regions in bZIP domains are largely unstructured in the absence of DNA and undergo folding only upon binding to DNA [62, 63, 104, 114-116]. In the ES model, it was assumed that the fully folded and the partially folded bZIP homodimers exist in equilibrium with the fully unfolded bZIP monomers and that DNA only binds to the monomers resulting in their folding and homodimerization — that is the bZIP domains

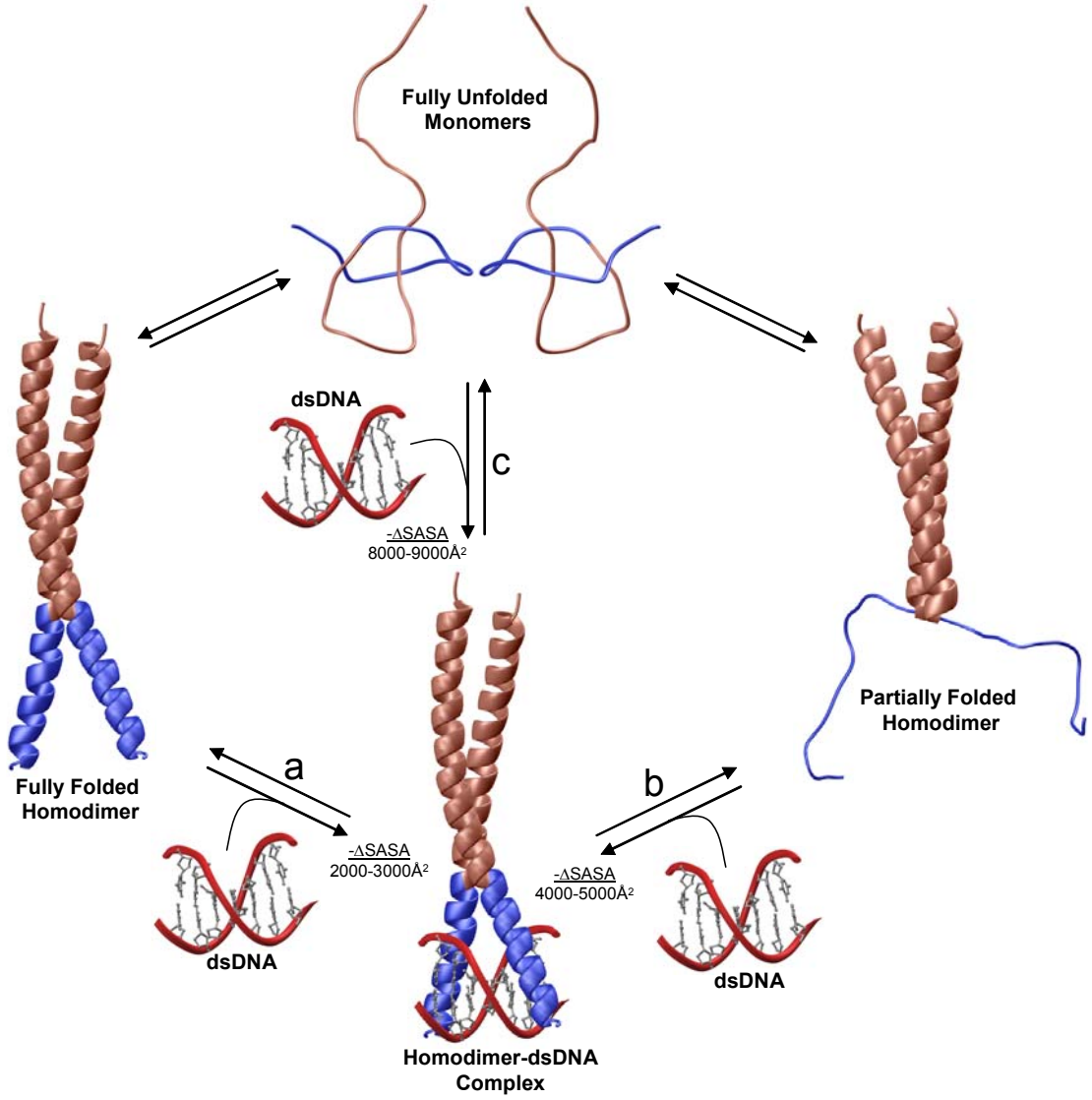

Figure 5-6. Plausible pathways for the binding of bZIP domain of Jun to dsDNA oligo containing the consensus sequence TGACTCA via Lock-and-Key (LK), Induced Fit (IF) and Equilibrium Shift (ES) models. (a) In the LK model, the bZIP domains are envisaged to exist as fully folded homodimers and undergo no conformational change upon DNA binding — that is the homodimers exist in a pre-formed conformation that best fits the DNA. The LK model is expected to result in the total burial of SASA of between 2000-3000Å^2. (b) In the IF model, the bZIP domains are presumed to exist as partially folded homodimers in which the basic regions are fully unstructured and only become structured upon DNA binding — that is DNA binding induces the folding of basic regions within otherwise pre-formed homodimers. The IF model is expected to result in the total burial of SASA of between 4000-5000Å^2. (c) In the ES model, the fully folded and the partially folded bZIP homodimers are hypothesized to exist in equilibrium with the fully unfolded bZIP monomers but DNA only binds to the monomers resulting in their folding and homodimerization — that is the bZIP domains bind to DNA as unfolded monomers such that their folding and homodimerization in association with DNA shifts the equilibrium with fully folded and partially folded homodimers in their direction. The ES model is expected to result in the total burial of SASA of between 8000-9000Å^2. 3D structures of bZIP domains alone and in complex with dsDNA oligo containing the TRE and CRE consensus sequences were determined using the MODELLER software. The dsDNA oligo shown contains the TRE consensus sequence TGACTCA with the DNA phosphate backbone depicted in red, while the sidechains of nucleotide bases are colored gray. For the bZIP domains shown, the leucine zippers are colored brown and the basic regions are in blue

bind to DNA as unfolded monomers such that their folding and homodimerization in association with DNA shifts the equilibrium with fully folded and partially folded homodimers in their direction. The ES model conjures support from the kinetic observation that the rate of dimerization of bZIP domains is significantly enhanced in the presence of DNA, implying that the bZIP domains associate with DNA as monomers coupled with their subsequent folding and dimerization [64]. The necessity for the fully unfolded monomers to be in equilibrium with the fully folded and the partially folded bZIP homodimers, as conjectured in the ES model, is due to the knowledge that the bZIP homodimers of Jun dissociate into monomers with a dissociation constant in the low micromolar range [47, 124, 129]. In light of this fact, it is thus logical to assume that the bZIP homodimers of Jun are likely to exist in equilibrium with monomers at the micromolar protein concentrations used in the calorimetric measurements recorded here. It is also of worthy note that the various states of Jun encompassing the fully unfolded and partially folded homodimer do not correspond to distinct conformations but rather should be considered as being comprised of an ensemble of conformations in agreement with previous studies [130, 131].

Table 5-3 summarizes and compares values for $\Delta SASA_{polar}$, $\Delta SASA_{apolar}$ and $\Delta SASA_{total}$ upon the interaction of the bZIP domains of Jun to TRE and CRE sites from our thermodynamic and structural data. Our analysis shows that there are significant conflicts between the ΔSASA values calculated from thermodynamic data versus the LK and IF models of binding described above. In contrast, the values determined from thermodynamic data agree par excellence with those calculated from the ES model. While ΔSASA values determined from thermodynamic data are between 2-3 fold greater

Table 5-3. Changes in polar SASA ($\Delta SASA_{polar}$), apolar SASA ($\Delta SASA_{apolar}$) and total SASA ($\Delta SASA_{total}$) upon the binding of bZIP domain of Jun to dsDNA oligos containing TRE and CRE sites obtained from thermodynamic and structural data.

DNA site →	TRE				CRE			
Method →	Thermodynamic	Structural			Thermodynamic	Structural		
Model →	MI	LK	IF	ES	MI	LK	IF	ES
$\Delta SASA_{polar}$ (Å^2)	−3508	−1385	−2201	−3404	−3640	−1157	−1835	−3038
$\Delta SASA_{apolar}$ (Å^2)	−4805	−1343	−2168	−4701	−4859	−1360	−2453	−4986
$\Delta SASA_{total}$ (Å^2)	−8313	−2728	−4369	−8105	−8499	−2517	−4288	−8024

ΔSASA values based on thermodynamic data were obtained from the measurement of ΔC_p and ΔH_{60} for the binding of the bZIP domain of Jun to dsDNA oligos containing TRE and CRE sites (Figure 5-3 and Table 5-2) using expressions [2] and [3], while ΔSASA values based on structural data were derived from 3D structural models of the bZIP domains of Jun alone and in complex with dsDNA oligos containing TRE and CRE sites (Figure 6) using expressions [4] and [5]. For ΔSASA values calculated from structural data, three models of binding were assumed — the Lock-and-Key (LK) model, the Induced Fit (IF) model and the Equilibrium Shift (ES) model (Figure 6). In the LK model, it was assumed that the bZIP domains exist as fully folded homodimers and undergo no conformational change upon DNA binding — that is the homodimers exist in a pre-formed conformation that best fits the DNA. In the IF model, it was assumed that the bZIP domains exist as partially folded homodimers in which the basic regions are fully unstructured and only become structured upon DNA binding — that is DNA binding induces the folding of basic regions within otherwise pre-formed homodimers. In the ES model, it was assumed that the fully folded and the partially folded bZIP homodimers exist in equilibrium with the fully unfolded bZIP monomers and that DNA only binds to the monomers resulting in their folding and homodimerization — that is the bZIP domains bind to DNA as unfolded monomers such that their folding and homodimerization in association with DNA shifts the equilibrium with fully folded and partially folded homodimers in their direction. ΔSASA values calculated from thermodynamic data make no assumptions and are thus model-independent (MI).

than those determined from structural data assuming the LK and IF models, these values agree within about 5-10% to those calculated from structural data assuming the ES model. The small anomalies in ΔSASA values between those obtained from thermodynamic data versus those calculated from ES model are likely due to errors in the atomic coordinates of the structural models. On the same token, the semi-empirical expressions [2] and [3] used to calculate ΔSASA values from thermodynamic data are by no means ideal and their poor parametrization may have also contributed to the small anomalies observed here between ΔSASA values obtained from thermodynamic data versus those calculated from the ES model of protein-DNA interaction. An alternative explanation for such anomalies may also be due to the assumption that DNA experiences

no conformational change upon interaction with the protein in spite of the evidence that it undergoes bending upon binding [50, 58, 117, 118]. Nonetheless, this latter assumption is an excellent approximation in our a priori calculations of ΔSASA from structural data due to negligible occlusion of molecular surface in DNA upon bending compared to rather large surface area buried upon protein-DNA contacts coupled with protein folding. It is thus not surprising that, despite small anomalies, the ΔSASA values observed upon protein-DNA interaction calculated from thermodynamic data versus those calculated from the ES model show remarkable consistencies. In sum, our heat capacity measurements reported here strongly support a model whereby the bZIP domains of Jun load onto DNA as monomers such that association with DNA triggers their folding and homodimerization. That is to say that the DNA-binding is coupled to the folding and homodimerization of bZIP domains of Jun.

5.5 Concluding remarks

Despite the knowledge of the significance of protein-DNA interactions to life for more than half a century, our understanding of the transient sequence of events leading up to the recognition of DNA by its protein counterparts hitherto remains abysmal. The classical picture based upon the notion of two rigid bodies coming together continues to strike a chord with most scientific literature and textbooks dealing with protein-DNA interactions. That this is so underlies the difficulties associated with unraveling the precise pathways by which transcription factors recognize specific response elements within the promoters of genes.

In an attempt to further our understanding of the mechanisms of protein-DNA interactions, we have reported herein thermodynamics of the binding of bZIP domains of

the transcription factor Jun to its cognate TRE and CRE sites within DNA. Our study shows that while both TRE and CRE bind to bZIP domains with virtually indistinguishable affinities, the nature of underlying thermodynamic forces is quite different. Furthermore, in comparison with our previous study [122], the Jun-Jun homodimer binds to DNA with an affinity that is over 2-fold greater than that observed for the binding of Jun-Fos heterodimer. This key finding is in stark contrast to a number of previous studies whereby the binding of Jun-Fos heterodimer to DNA has been suggested to be stronger than that of Jun-Jun homodimer and epitomizes the power of ITC as a quantitative tool for the analysis of protein-DNA interactions [5, 42, 49, 56, 57, 89]. The conventional view that the Jun-Fos heterodimer bound to DNA much stronger than Jun-Jun homodimer was in part resurrected due to the observation that the latter is a much weaker dimer than the former and that such differential dimer stability may be directly correlated with their binding potential [47, 124, 129]. On the contrary, the decreased stability of dimeric transcription factors may be a recipe for their enhanced binding potential to DNA through pathways that are kinetically more favorable as we have exquisitely shown here. It is also widely believed that the Jun-Fos heterodimer is a more potent activator of mitogenic transcription than the Jun-Jun homodimer and that such differential potency is largely due to the higher DNA-binding affinity of Jun-Fos heterodimer relative to Jun-Jun homodimer [5, 42, 49, 56, 57, 89]. Our data presented here refute this long-held claim and suggest that differential transcriptional activities may be attributable to differential energetics in lieu of differential binding affinities. In light of this view, understanding the spatial and temporal specificity of transcription factors may

require complete understanding of the underlying thermodynamic forces rather than mere analysis of their relative binding affinities.

Our heat capacity changes accompanying the Jun-DNA interaction are best accounted for by a model in which Jun monomers load onto DNA as monomers such that association with DNA triggers their folding and homodimerization. Because this model does not necessitate the requirement of a pre-formed Jun-Jun homodimer that best fits the DNA and given that Jun may largely exist as a monomer under physiological conditions due to its relatively weak dimer dissociation constant in the low micromolar range [47, 124, 129], it may also be the most favorable pathway under physiological conditions. However, such a mechanism does not mutually exclude other models in which Jun may bind to DNA as a fully folded or partially folded homodimer and only direct kinetic analysis can provide information on the most preferred pathway under non-equilibrium conditions. Nonetheless, kinetic studies have revealed that the bZIP domains of Jun and Fos associate with DNA as monomers coupled with their subsequent folding and heterodimerization [64]. In these studies, the heterodimerization of bZIP domains of Jun and Fos was best described by a single slow kinetic phase in the absence of DNA on the basis of fluorescence resonance energy transfer (FRET) between the two protein molecules. However, a second fast kinetic phase was observed when such heterodimerization was analyzed in the presence of DNA. Because no FRET was observed upon the interaction of pre-formed heterodimers with DNA, it was argued that the second fast kinetic phase must arise from the heterodimerization of Jun and Fos on DNA. This observation was further corroborated by the dependence of both the rate constant and the amplitude of the second fast kinetic phase upon DNA concentration. It is

clearly evident that direct kinetic analysis of the binding of Jun to DNA using the above-mentioned FRET is not possible. However, we are pursuing a number of alternative strategies to decipher the preferred kinetic pathway by which Jun binds to DNA.

In conclusion, our thermodynamic analysis of Jun-DNA interaction suggests that the binding of transcription factors to DNA as monomers coupled with their subsequent folding and dimerization may be a more common mechanism employed in protein-DNA interactions and that the conventional school of thought may need to be re-evaluated. Our future studies will be directed toward obtaining further evidence in support of this model. Nevertheless, our present study promises to break new ground and provokes further research on elucidating the precise kinetic pathways by which protein-DNA interactions ensue.

6 Chapter 6: Single Nucleotide Variants of the TGACTCA Motif Modulate Energetics and Orientation of Binding of the Jun-Fos Heterodimeric Transcription Factor

6.1 Summary

The Jun-Fos heterodimeric transcription factor is the terminal link between the transfer of extracellular information in the form of growth factors and cytokines to the site of DNA transcription within the nucleus in a wide variety of cellular processes central to health and disease. Here, using isothermal titration calorimetry, we report detailed thermodynamics of the binding of bZIP domains of Jun-Fos heterodimer to synthetic dsDNA oligos containing the TGACTCA cis-element and all possible single nucleotide variants thereof encountered widely within the promoters of a diverse array of genes. Our data show that Jun-Fos heterodimer tolerates single nucleotide substitutions and binds to TGACTCA variants with affinities in the physiologically relevant micromolar-submicromolar range. The energetics of binding are richly favored by enthalpic forces and opposed by entropic changes across the entire spectrum of TGACTCA variants in agreement with the notion that protein-DNA interactions are largely driven by electrostatic interactions and intermolecular hydrogen bonding. Of particular interest is the observation that the Jun-Fos heterodimer binds to specific TGACTCA variants in a preferred orientation. Our 3D atomic models reveal that such orientational preference results from asymmetric binding and may in part be attributable to chemically distinct but structurally equivalent residues R263 and K148 located within the basic regions of Jun and Fos, respectively. Taken together, our data suggest that the single nucleotide variants of the TGACTCA motif modulate energetics and orientation of binding of the Jun-Fos heterodimer and that such behavior may be a critical determinant

of differential regulation of specific genes under the control of this transcription factor. Our study also bears important consequences for the occurrence of single nucleotide polymorphisms within the TGACTCA cis-element at specific gene promoters between different individuals.

6.2 Background

The transcription factor AP1 (activator protein 1), comprised largely of constituent proteins Jun and Fos, executes the terminal stage of many critical signaling cascades that initiate at the cell surface and reach their climax in the nucleus [5, 39]. Upon activation by MAP kinases, AP1 binds to the TGACTCA consensus motif and many closely related sequences within the promoters of a multitude of genes as Jun-Jun homodimer or Jun-Fos heterodimer. In so doing, Jun and Fos recruit the transcriptional machinery to the site of DNA and switch on expression of genes involved in a diverse array of cellular processes such as cell growth and proliferation, cell cycle regulation, embryonic development and cancer [3, 83, 84]. Jun and Fos recognize the TGACTCA and related sequences at the promoters of specific genes through their so-called basic zipper (bZIP) domains (Figure 6-1a). The bZIP domain can be further dissected into two well-defined functional subdomains termed the basic region (BR) at the N-terminus followed by the leucine zipper (LZ) at the C-terminus. The leucine zipper is a highly conserved protein module found in a wide variety of cellular proteins and usually contains a signature leucine at every seventh position within the five successive heptads of amino acid residues. The leucine zippers adapt continuous α-helices in the context of Jun-Jun homodimer or Jun-Fos heterodimer by virtue of their ability to wrap around each other in a coiled coil dimer [41-43]. Such intermolecular arrangement brings the basic

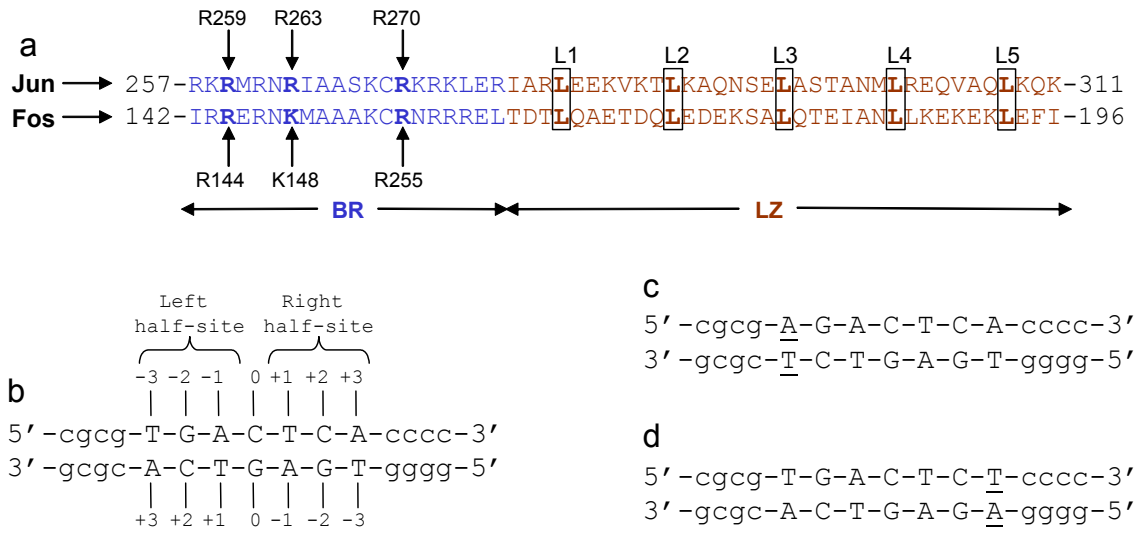

Figure 6-1. Protein and DNA sequences. (a) Subdivision of bZIP domain into its respective N-terminal basic region (BR) and the C-terminal leucine zipper (LZ) for Jun and Fos transcription factors. The BR and LZ subdomains are colored blue and brown, respectively. The five signature leucines (L1-L5) characteristic of LZ subdomains are boxed and bold faced. The basic residues within the BR subdomains that hydrogen bond with specific DNA bases are labeled by vertical arrows. (b) Nucleotide sequence of 15-mer dsDNA oligo containing the TGACTCA motif. The TGACTCA motif is capitalized whilst the flanking nucleotides are shown in small letters. The numbering of various nucleotides relative to the central C/G base pair (assumed to be at zero position) in both strands is indicated. The TGA and TCA half-sites within this motif are also marked. (c) Nucleotide sequence of dsDNA oligo containing the T→A mutation at –3 position and herein referred to as the A-3 oligo. The variant nucleotides in both strands are underlined. (d) Nucleotide sequence of dsDNA oligo containing the A→T mutation at +3 position and herein referred to as the T+3 oligo. The variant nucleotides in both strands are underlined. Note that the A-3 and T+3 oligos shown in (c) and (d) are examples of a pair of symmetrically related dsDNA oligos in that they contain identical half-sites in opposite directions— these half-sites are indistinguishable upon the rotation of the variant motif by 180° in the plane of the paper (two-fold symmetry).

regions at the N-termini of bZIP domains into close proximity and thereby enables them to insert into the major grooves of DNA at the promoter regions in an optimal fashion in a manner akin to a pair of forceps [43]. The basic regions are believed to be unstructured and fold into α-helices only upon association with DNA in a coupled folding-binding manner [62, 63, 104, 114, 116]. While the α-helices are held together by numerous inter-helical hydrophobic contacts and salt bridges, hydrogen bonding between the sidechains of basic residues in the basic regions and the DNA bases accounts for high affinity binding of bZIP domains to DNA.

X-ray crystal structure analysis shows that the bZIP domains of Jun-Fos heterodimer bind to the TGACTCA sequence in a non-preferred orientation [43], implying that this promoter element does not dictate the orientation of the Jun-Fos heterodimer. Given that the AP1 transcription factor cooperates and acts in concert with a diverse array of other transcription factors, including many steroid hormone receptors, within the transcription initiation complex in regulating gene expression, the orientational binding of Jun-Fos heterodimer to gene promoters could be a key determinant of its transcriptional potency. But how would this be achieved? Since the discovery of TGACTCA motif as an essential cis-element for recognizing the AP1 transcription factor within the promoter of metallothionein 2A gene over two decades ago [48], there has been a growing number of studies demonstrating the key role of single nucleotide variants (SNVs) of this element within a diverse spectrum of genes for recruiting AP1 to the site of transcriptional machinery [132-144]. This exciting episode of spicing up the TGACTCA motif with genetic variation across a diverse array of gene promoters took yet another twist recently upon the demonstration that this element may not only be subject to SNVs but that single nucleotide polymorphisms (SNPs) may also feature heavily within this element across different individuals, especially those from distinct ethnic groups [145]. Thus, while SNVs may allow differential regulation of specific AP1-responsive genes, SNPs may determine the phenotypic makeup of an individual by virtue of their ability to differentially modulate the binding of transcription factors to identical promoter regions between different individuals and thus could account for differential response of individuals to specific diseases. A SNP within the promoter of MDM2

ubiquitin ligase gene, a negative regulator of p53 tumor suppressor, has indeed been implicated in cancer [146].

Although the binding of AP1 to promoter elements containing the TGACTCA sequence has been extensively explored in biophysical terms over the past two decades or so, the effect of genetic variations within this sequence upon protein-DNA interaction remains hitherto poorly understood. In an attempt to analyze the effect of such genetic variations within the TGACTCA sequence on AP1-DNA interactions and whether such variations could dictate their relative orientation, we have employed here isothermal titration calorimetry (ITC) to study detailed thermodynamics of the binding of bZIP domains of Jun-Fos heterodimer to synthetic dsDNA oligos containing the TGACTCA cis-element and all possible SNVs thereof encountered widely within the promoters of a diverse array of genes. Our data suggest that TGACTCA variants modulate energetics and orientation of binding of the Jun-Fos heterodimer and that such behavior may be a critical determinant of differential regulation of specific genes under the control of this transcription factor. Our study also bears important consequences for the occurrence of SNPs within the TGACTCA cis-element at the promoters of identical genes in different individuals.

6.3 Materials and methods

6.3.1 Protein preparation

bZIP domains of human Jun and Fos were cloned and expressed as described previously [122]. Briefly, the proteins were cloned into pET102 bacterial expression vector, with an N-terminal thioredoxin (Trx)-tag and a C-terminal polyhistidine (His)-tag, using Invitrogen TOPO technology. Additionally, thrombin protease sites were

introduced at both the N- and C-termini of the proteins to aid in the removal of tags after purification. Proteins were subsequently expressed in Escherichia coli Rosetta2(DE3) bacterial strain (Novagen) and purified on Ni-NTA affinity column using standard procedures. Further treatment of bZIP domains of Jun and Fos on MonoQ ion-exchange column coupled to GE Akta FPLC system led to purification of recombinant domains to apparent homogeneity as judged by SDS-PAGE analysis. The identity of recombinant proteins was confirmed by MALDI-TOF mass spectrometry analysis. Final yields were typically between 10-20mg protein of apparent homogeneity per liter of bacterial culture. As noted previously [122], the treatment of recombinant proteins with thrombin protease significantly destabilized the bZIP domains of both Jun and Fos and both domains appeared to be proteolytically unstable. For this reason, all experiments reported herein were carried out on recombinant fusion bZIP domains of Jun and Fos containing a Trx-tag at the N-terminus and a His-tag at the C-terminus. The tags were found to have no effect on the binding of these domains to DNA under all conditions used here. Protein concentrations were determined as described earlier [122]. Jun-Fos bZIP heterodimers were generated by mixing equimolar amounts of the purified bZIP domains of Jun and Fos. The efficiency of bZIP heterodimerization was close to 100% as judged by Native-PAGE and size exclusion chromatography (SEC) analysis using a Hiload Superdex 200 column.

6.3.2 DNA synthesis

HPLC-grade 15-mer DNA oligos containing the TGACTCA consensus motif and all possible single nucleotide variants thereof were commercially obtained from Sigma Genosys. The flanking nucleotides were appropriately chosen to prevent self-annealing of

sense and antisense strands. The design of such oligos and the numbering of various nucleotides relative to the central C/G base pair are depicted in Figures 5-1b-d. Oligo concentrations were determined spectrophotometrically on the basis of their extinction co-efficients derived from their nucleotide sequences using the online software OligoAnalyzer 3.0 (Integrated DNA Technologies) based on the nearest-neighbor model [68]. Double-stranded DNA (dsDNA) oligos were generated as described earlier [122].

6.3.3 ITC measurements

Isothermal titration calorimetry (ITC) experiments were performed on Microcal VP-ITC instrument and data were acquired and processed using Microcal ORIGIN software. All measurements were repeated 2-3 times. Briefly, the bZIP domains of Jun-Fos heterodimer and dsDNA oligos were prepared in 50mM Tris, 200mM NaCl, 1mM EDTA and 5mM β-mercaptoethanol at pH 8.0. The experiments were initiated by injecting 20 x 10μl aliquots of 50-100μM of dsDNA oligo from the syringe into the calorimetric cell containing 1.8ml of 5-10μM of the bZIP domains of Jun-Fos heterodimer at 25°C. The data were fit to a 1-site model derived from the binding of a ligand to a macromolecule using the law of mass action to extract the various thermodynamic parameters as described previously [122]. Although Tris buffer is not ideally suited for ITC analysis due to its high ionization enthalpy, enthalpy of binding of Jun-Fos heterodimer to dsDNA oligos was identical in both Tris and phosphate buffers, implying that the values of enthalpy being reported here solely arise from the binding process with no contributions from coupled equilibria due to protonation/deprotonation.

6.3.4 Structural modeling

3D structures of bZIP domains of Jun-Fos heterodimer in complex with dsDNA oligos containing the TGACGCA variant motif in two possible orientations I and II were modeled using the crystal structure of bZIP domains of Jun-Fos heterodimer in complex with a dsDNA oligo containing the TGACTCA sequence as a template (with PDB code of 1FOS) in MODELLER [76]. Additionally, hydrogen bonding restraints were introduced. For orientation I, hydrogen bonds were added between the NH1 atom of R263 in Jun and O6 atom of G+1 in the sense strand of TGACGCA motif, between the NH2 atom of R263 in Jun and O6 atom of G+1 in the sense strand of TGACGCA motif, and between the NZ atom of K148 in Fos and O4 atom of T+1 in the antisense strand of TGACGCA motif. For orientation II, hydrogen bonds were added between the NH1 atom of R263 in Jun and O4 atom of T+1 in the antisense strand of TGACGCA motif, between the NH2 atom of R263 in Jun and O4 atom of T+1 in the antisense strand of TGACGCA motif, and between the NZ atom of K148 in Fos and O6 atom of G+1 in the sense strand of TGACGCA motif. In each case, a total of 100 structural models were calculated and the structure with the lowest energy, as judged by the MODELLER Objective Function, was selected for further energy minimization in MODELLER prior to analysis. The structures were rendered using RIBBONS [77].

6.4 Results and discussion

6.4.1 Jun-Fos heterodimer tolerates single nucleotide substitutions at all positions within the TGACTCA motif

A previous study based on an analysis of highly qualitative nature demonstrated that single nucleotide substitutions at specific positions within the TGACTCA motif

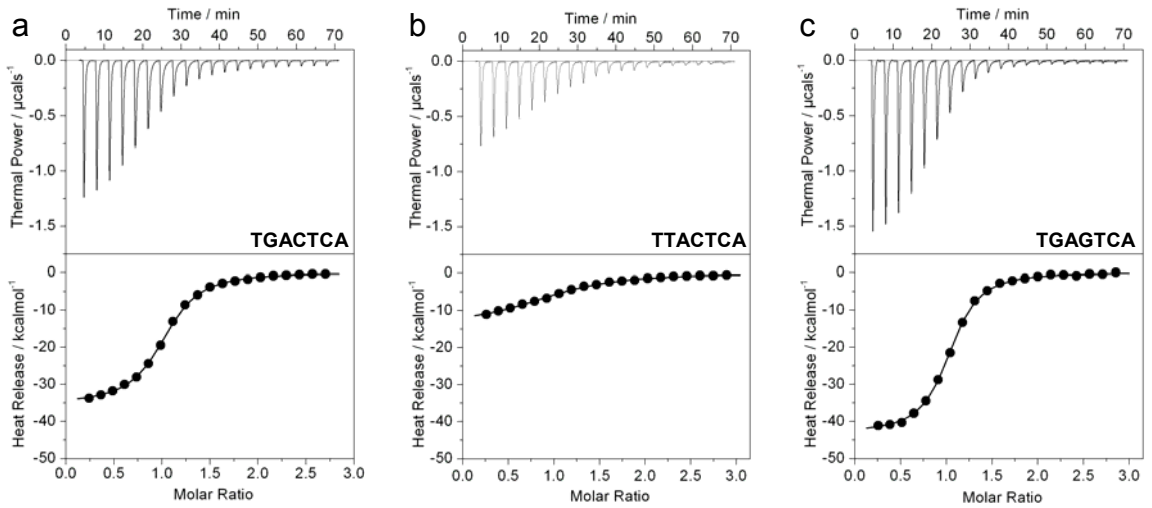

Figure 6-2. Representative ITC isotherms for the binding of bZIP domains of Jun-Fos heterodimer to dsDNA oligos containing the promoter sites TGACTCA (a), TTACTCA (b) and TGAGTCA (c). Note that the variant motif TGAGTCA is related to the wildtype TGACTCA motif by 2-fold symmetry. The position of the variant nucleotide in each of the sites relative to the consensus sequence TGACTCA is underlined. The solid lines represent the fit of the data points in the lower panels to a function based on the binding of a ligand to a macromolecule using the Microcal ORIGIN software [74].

completely abrogated binding by the bZIP domains of Jun-Fos heterodimer [51]. This conclusion is highly surprising given that all possible single nucleotide variants of the TGACTCA motif are encountered within the promoter regions of a diverse array of genes whereby they act in concert with other promoter elements to regulate the action of AP1 transcription factor in a differential manner [132-144]. In light of this argument coupled with the fact that the overall binding energy results from the summation of all interactions between protein residues and DNA bases in a cooperative manner, we reasoned that single substitutions within the TGACTCA motif may result in the reduction of free energy associated with such protein-DNA interactions but are unlikely to lead to complete abrogation within the physiological context. To test our hypothesis, we introduced all possible single substitutions within the TGACTCA motif and measured their effect on the binding of bZIP domains of Jun-Fos heterodimer using the powerful and highly quantitative technique of ITC. Figure 6-2 shows representative isotherms

Table 6-1. Experimentally determined thermodynamic parameters for the binding of bZIP domains of Jun-Fos heterodimer to dsDNA oligos containing the wild-type (WT) consensus motif TGACTCA and all possible single nucleotide variants thereof obtained from ITC measurements at 25°C and pH 8.0

motif	sequence	n	K_d/μM	ΔH/kcal mol^{-1}	$T\Delta S$/kcal mol^{-1}	ΔG/kcal mol^{-1}	gene promoter
WT	TGACTCA	1.02 ± 0.01	0.21 ± 0.01	−35.08 ± 0.22	−25.94 ± 0.19	−9.12 ± 0.03	metallothionein 2A
A−3	AGACTCA	0.98 ± 0.09	3.55 ± 0.76	−31.02 ± 0.59	−23.56 ± 0.48	−7.45 ± 0.13	aflatoxin B1 aldehyde reductase AKR7
C−3	CGACTCA	0.97 ± 0.04	5.96 ± 2.28	−19.70 ± 0.56	−12.53 ± 0.32	−7.16 ± 0.23	transcriptional activator ORF50
G−3	GGACTCA	1.03 ± 0.06	2.82 ± 0.17	−31.36 ± 0.36	−23.80 ± 0.36	−7.58 ± 0.04	mu-opioid receptor MOR1
A−2	TAACTCA	0.98 ± 0.01	6.31 ± 0.08	−26.48 ± 0.28	−19.37 ± 0.30	−7.10 ± 0.01	cytochrome P450 enzyme CYP27B1
C−2	TCACTCA	1.05 ± 0.04	8.10 ± 1.06	−30.74 ± 0.20	−23.77 ± 0.27	−6.95 ± 0.08	monocyte chemotactic protein MCP1
T−2	TTACTCA	1.00 ± 0.02	2.08 ± 0.05	−13.98 ± 0.14	−6.23 ± 0.17	−7.76 ± 0.01	transcriptional regulator JE/MCAF
C−1	TGCCTCA	0.92 ± 0.03	4.13 ± 0.59	−29.89 ± 0.21	−22.53 ± 0.30	−7.35 ± 0.08	pro-opiomelanocortin hormone POMC
G−1	TGGCTCA	1.06 ± 0.01	3.90 ± 0.50	−33.61 ± 0.13	−26.21 ± 0.06	−7.39 ± 0.08	bone protein osteocalcin
T−1	TGTCTCA	0.98 ± 0.01	3.62 ± 0.30	−36.89 ± 0.06	−29.46 ± 0.11	−7.43 ± 0.05	transforming growth factor TGFB1
A0	TGAATCA	0.95 ± 0.04	0.58 ± 0.04	−39.88 ± 0.10	−31.36 ± 0.06	−8.52 ± 0.04	mitochondrial proton carrier UCP1
G0	TGAGTCA	1.05 ± 0.04	0.14 ± 0.01	−42.87 ± 0.16	−33.52 ± 0.21	−9.34 ± 0.03	cell cycle regulator cyclin D1
T0	TGATTCA	1.04 ± 0.04	0.50 ± 0.02	−38.85 ± 0.21	−30.25 ± 0.21	−8.60 ± 0.02	tyrosine hydroxylase TH
A+1	TGACACA	1.00 ± 0.03	0.92 ± 0.03	−37.06 ± 0.14	−28.82 ± 0.17	−8.24 ± 0.02	glucocorticoid receptor GR
C+1	TGACCCA	0.95 ± 0.06	4.45 ± 0.40	−31.95 ± 0.09	−24.63 ± 0.02	−7.31 ± 0.05	interferon IFN-γ
G+1	TGACGCA	0.95 ± 0.01	0.64 ± 0.01	−32.78 ± 0.13	−24.33 ± 0.15	−8.46 ± 0.01	ECM glycoprotein fibronectin 1
A+2	TGACTAA	1.04 ± 0.04	2.00 ± 0.05	−20.90 ± 0.09	−13.11 ± 0.13	−7.78 ± 0.01	MAP kinase phosphatase MKP3
G+2	TGACTGA	0.92 ± 0.03	5.92 ± 1.05	−28.16 ± 0.19	−21.01 ± 0.08	−7.14 ± 0.11	steroidogenic acute regulator STAR
T+2	TGACTTA	1.07 ± 0.03	5.16 ± 0.15	−27.88 ± 0.11	−20.65 ± 0.08	−7.22 ± 0.02	nerve growth factor NGF
C+3	TGACTCC	1.04 ± 0.02	3.07 ± 0.29	−31.90 ± 0.11	−24.35 ± 0.04	−7.53 ± 0.06	chloride channel CLC5
G+3	TGACTCG	0.97 ± 0.02	1.80 ± 0.17	−33.33 ± 0.06	−25.46 ± 0.02	−7.85 ± 0.06	amyloid precursor protein β-APP
T+3	TGACTCT	1.02 ± 0.03	3.30 ± 0.15	−42.46 ± 0.13	−34.97 ± 0.10	−7.49 ± 0.03	tumor suppressor p53

Note that the DNA sequence shown for the TGACTCA motif and its single nucleotide variants corresponds to the sense strand only and nucleotides flanking these motifs have been omitted for clarity (see Figures 1b-d). The substituted nucleotide relative to the TGACTCA motif is underlined. One example of a gene promoter that contains a particular TGACTCA variant for recruiting the AP1 transcription factor is provided for physiological relevance [132-144]. The values for the stoichiometry (n), affinity (K_d) and enthalpy change (ΔH) accompanying the binding of bZIP domains of Jun-Fos heterodimer to dsDNA oligos were obtained from the fit of a function, based on the binding of a ligand to a macromolecule using the law of mass action [74], to the ITC isotherms. Free energy of binding (ΔG) was calculated from the relationship $\Delta G=RT\ln K_d$, where R is the universal molar gas constant (1.99 cal/mol/K) and T is the absolute temperature (K). Entropic contribution (TΔS) to binding was calculated from the relationship $T\Delta S=\Delta H-\Delta G$. Errors were calculated from 2-3 independent measurements. All errors are given to one standard deviation.

obtained from such experiments, while detailed analysis of all the associated thermodynamic parameters is presented in Table 6-1. Our data indeed support our hypothesis that no single nucleotide substitution at any position within the TGACTCA motif is sufficient per se to completely abolish the binding of Jun-Fos heterodimer. The binding affinities observed are in the physiologically relevant range and vary by up to 60-fold from a value of 0.14μM to 8.10μM, implying that genetic variations within the TGACTCA motif at distinct gene promoters may be critical determinants of tightly modulating the transcriptional potency of the Jun-Fos heterodimer.

To further rationalize the effect of such genetic variations on the transcriptional output of Jun-Fos heterodimer, we categorized the spectrum of binding affinities observed between bZIP domains and variants of TGACTCA motif into three major classes: those substitutions that bind with an affinity similar to the wildtype (WT) motif containing the TGACTCA sequence, those that bind with submicromolar affinity and those that bind with micromolar affinity. The only motif that binds to bZIP domains with an affinity similar to the WT motif is the one containing G at 0 position (G0) within the TGACTCA sequence. That this is so is expected in light of the fact that the TGAGTCA motif is related to the TGACTCA motif by a two-fold symmetry with identical TGA and TCA half-sites in both the sense and antisense strands but in opposite orientations. There are four single nucleotide substitutions within the TGACTCA sequence that reduce the binding affinity of bZIP domains only moderately by up to about four-fold to submicromolar levels. These include the substitutions A at 0 position (A0), T at 0 position (T0), A at +1 position (A+1) and G at +1 position (G+1). All other single nucleotide substitutions within the TGACTCA motif reduce the binding affinity of bZIP domains by at least an order of magnitude to micromolar levels. In summary, we report here for the first time that no single nucleotide substitutions at any given position within the TGACTCA motif abrogate the binding of bZIP domains of Jun-Fos heterodimer. Our new findings suggest that single nucleotide substitutions within the cis-acting promoter elements may have evolved as a subtle mechanism to differentially regulate transcriptional activity of AP1 at distinct promoters through differential binding. It is likely that the effect of such single nucleotide substitutions on the energetics of binding

directly correlates with the transcriptional activity of AP1 under physiological context. This hypothesis will be tested in our future studies.

6.4.2 Enthalpy-entropy compensation buffers the binding of Jun-Fos heterodimer to single nucleotide variants of the TGACTCA motif

It is evident from our data that the binding of bZIP domains to DNA is largely driven by favorable enthalpic contributions accompanied by entropic penalty at physiological temperatures regardless of the position of nucleotide substitution within the TGACTCA motif (Table 6-1). This is not at all surprising given that an extensive network of hydrogen bonding, electrostatic interactions and hydrophobic contacts between residues in the basic regions of bZIP domains and DNA have to be established. On the other side, the entropic penalty largely results from the overall difference between two major opposing entropic forces. The favorable entropic force is the increase in the degrees of freedom of water molecules while the unfavorable entropic force is the decrease in the degrees of freedom of backbone and sidechain atoms within the protein and DNA molecules upon complexation. The substantial entropic penalty observed here thus suggests that the protein and DNA molecules experience a greater loss of entropy than that gained by water molecules upon intermolecular association.

It should be particularly noted here that enthalpy and entropy are not necessarily opposing forces but they often act in an antagonistic manner in biological systems to maintain more or less constant energetics of binding in response to external factors such as temperature and internal changes such as mutations. Such thermodynamic single nucleotide substitutions. As exquisitely illustrated in Figure 6-3, various pairs of TGACTCA variants undergo enthalpy-entropy compensation with little or negligible

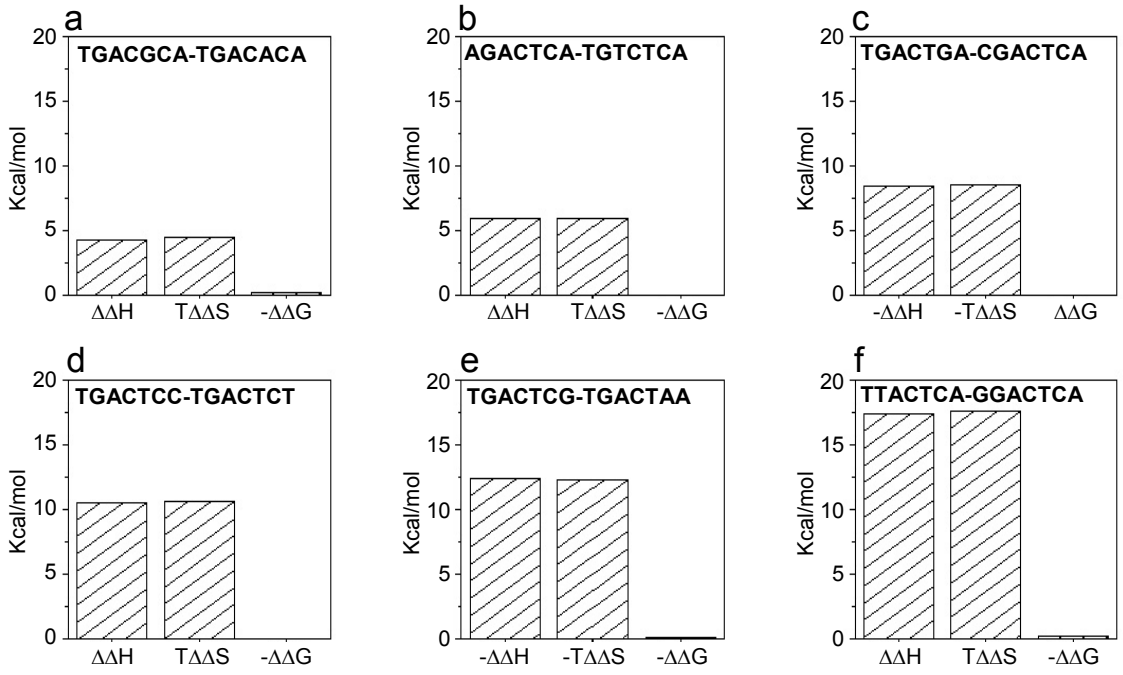

Figure 6-3. Differential thermodynamic signatures for the binding of bZIP domains of Jun-Fos heterodimer to various pairs of dsDNA oligos containing TGACTCA variants with similar affinities: (a) TGAC<u>G</u>CA relative to TGAC<u>A</u>CA; (b) <u>A</u>GACTCA relative to TG<u>T</u>CTCA; (c) TGACT<u>G</u>A relative to <u>C</u>GACTCA; (d) TGACTC<u>C</u> relative to TGACTC<u>T</u>; (e) TGACTC<u>G</u> relative to TGACT<u>A</u>A; and (f) T<u>T</u>ACTCA relative to <u>G</u>GACTCA. The position of the variant nucleotide in each of the sites relative to the consensus sequence TGACTCA is underlined. ΔΔH, TΔΔS and ΔΔG were calculated from the relationships $\Delta\Delta H=\Delta H_x-\Delta H_y$, $T\Delta\Delta S=T\Delta S_x-T\Delta S_y$ and $\Delta\Delta G=\Delta G_x-\Delta G_y$, where the subscripts x and y denote the corresponding thermodynamic parameters for the binding of bZIP domains of Jun-Fos heterodimer to dsDNA oligo x relative to oligo y, respectively (Table 6-1).

effect on the overall binding energetics. These data thus underscore how subtle genetic variations within cis-elements may not necessarily translate into loss of binding energetics and that such a feat may be accomplished through an underlying enthalpy-entropy compensatory mechanism intrinsic to biological systems. Homeostasis, or thermostasis as a portmanteau, also appears to be a hallmark of the binding of bZIP domains of Jun-Fos heterodimer to TGACTCA motif containing various single nucleotide substitutions. As exquisitely illustrated in Figure 6-3, various pairs of TGACTCA variants undergo enthalpy-entropy compensation with little or negligible effect on the overall binding energetics. These data thus underscore how subtle genetic variations within cis-elements may not necessarily translate into loss of binding

energetics and that such a feat may be accomplished through an underlying enthalpy-entropy compensatory mechanism intrinsic to biological systems.

6.4.3 Jun-Fos heterodimer binds to specific variants of the TGACTCA motif in a preferred orientation

The TGA and TCA half-sites within the TGACTCA motif are related by a two-fold symmetry — 180° rotation of either half-site about the central C/G base pair within the context of dsDNA generates the other (Figures 6-1b-d). However, the TGACTCA motif is not a perfect palindrome and therefore Jun and Fos may have a preference for one half-site over the other due to non-identical contacts with the central C/G base pair. Such a scenario may blossom into the binding of Jun and Fos to TGACTCA motif with a preferred orientation. If this is indeed true, one would expect differential energetics of binding of Jun and Fos to dsDNA oligos containing symmetrically related TGACTCA variants since non-identical contacts with TGA and TCA half-sites would almost certainly be expected to result in varying binding affinities. On the other hand, if Jun and Fos engage in identical contacts with TGA and TCA half-sites, this would not be expected to result in differential energetics of binding to dsDNA oligos containing symmetrically related TGACTCA variants. This latter scenario would be indicative of non-preferred orientation of Jun-Fos heterodimer in association with DNA due to the fact that equivalent energetics of binding would allow the two monomers to freely exchange with each half-site. In an attempt to analyze the extent to which Jun and Fos may exhibit such orientational preference and the extent to which it may be modulated by genetic variations, we plotted relative binding affinities of symmetrically related pairs of dsDNA

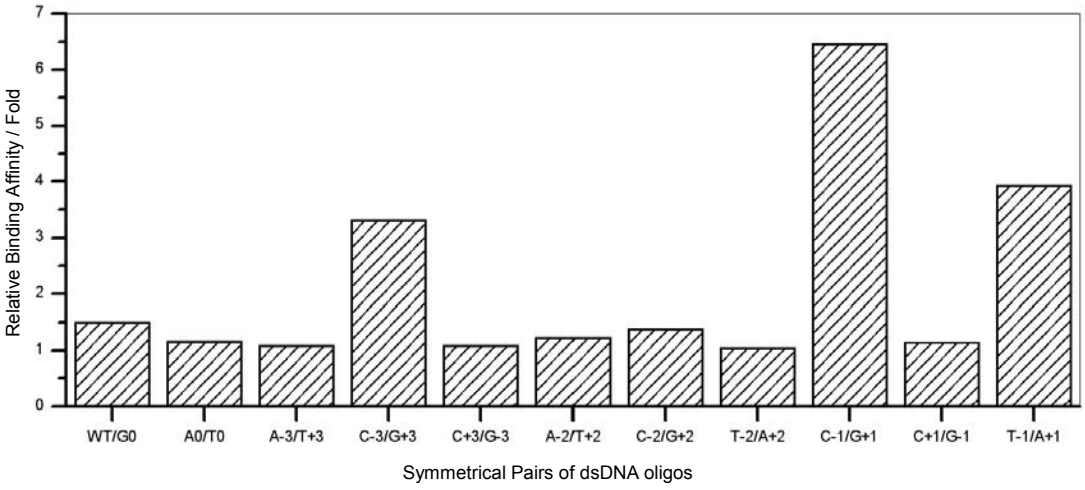

Figure 6-4. Analysis of relative binding affinities of symmetrically related pairs of dsDNA oligos containing TGACTCA variants to bZIP domains of Jun-Fos heterodimer. Relative binding affinity is defined as the ratio of the binding affinity of one dsDNA oligo to Jun-Fos heterodimer over that of the other. The nomenclature for the various dsDNA oligos shown is the same as that indicated in Table 6-1.

oligos containing TGACTCA variants to bZIP domains of Jun-Fos heterodimer (Figure 6-4).

It is clearly evident from our data that the vast majority of symmetrically related pairs of dsDNA oligos bind to Jun-Fos heterodimer with relative binding affinities of close to unity, implying that they most likely bind without a preferred orientation due to little or negligible differences in the energetics of binding. Among these symmetrically related pairs of dsDNA oligos that appear to assume non-preferred orientation are the TGACTCA (WT) and TGAGTCA (G0) motifs — an observation that is consistent with X-ray structural analysis of Jun-Fos heterodimer in complex with dsDNA oligo containing the TGACTCA motif and a number of other studies reported previously [43, 58, 118, 147]. Although the finding that Jun and Fos bind to TGACTCA motif in a non-preferred orientation was expected, the fact that they employ quite a distinct interplay between underlying enthalpic and entropic forces is being reported here for the first time (Table 6-1). Such enthalpy-entropy compensation to smooth out any differentiation

between the overall binding energetics to symmetrically related WT and G0 dsDNA oligos cannot be accounted for by structural data and points to the need for further understanding of protein-DNA interactions in biophysical terms. We believe that such differences in the underlying thermodynamic parameters for the binding of Jun-Fos heterodimer to symmetrically related WT and G0 dsDNA oligos are most likely due to the differences in the flanking nucleotides. In order to avoid self-annealing of sense and antisense strands, it has not been possible at this stage to completely eliminate the contributions of flanking nucleotides on our thermodynamic measurements reported herein. However, it should be noted that such contributions should be minimal given that the flanking nucleotides make no discernable contact with Jun-Fos heterodimer [43]. We will fully explore the effect of flanking nucleotides on the thermodynamics of binding of Jun-Fos heterodimer to DNA in our future studies. Interestingly, our analysis also reveals that Jun-Fos heterodimer binds in a non-preferred orientation to symmetrically related dsDNA oligos containing the TGAATCA (A0) and TGATTCA (T0) motifs, indicating that the central C/G base pair can be substituted by A/T base pair without any effect on the orientation of Jun-Fos heterodimer. However, unlike the central C/G base pair, the underlying interplay between entropic and enthalpic forces appears to be very similar in the case of A/T base pair (Table 6-1). Perhaps even more striking is the observation that the Jun-Fos heterodimer binds without preferred orientation to symmetrically related motifs A-3/T+3, C+3/G-3, A-2/T+2, C-2/G+2, T-2/A+2 and C+1/G+1, which all contain a non-central single nucleotide substitution.

The preference for orientation however does not completely escape the Jun-Fos heterodimer in its quest to bind to single nucleotide variants of the TGACTCA motif. The

symmetrically related motifs C-3/G+3, T-1/A+1 and C-1/G+1 clearly exhibit differential energetics of binding to Jun-Fos heterodimer as demonstrated through their relative binding affinities of over 3-fold, 4-fold and 6-fold, respectively. That this is so implies strongly that the Jun-Fos heterodimer binds to these pairs of symmetrically related motifs in a preferred orientation due to formation of non-identical contacts with DNA in the two possible orientations allowed.

6.4.4 3D atomic models provide structural basis of the binding of Jun-Fos heterodimer to the TGACGCA variant in a preferred orientation

X-ray crystallography analysis provided the structural basis for the non-preferred orientation of Jun-Fos heterodimer in complex with dsDNA oligo containing the TGACTCA motif [43]. This seminal work revealed that the lack of such preference for orientation was not due to the symmetric binding of Jun and Fos to the central C/G base pair but, on the contrary, asymmetric interactions were observed between R270 in Jun and R155 in Fos with DNA bases. Thus, while R270 sidechain was observed to hydrogen bond with the central C (C0) in one strand, R155 sidechain engaged in hydrogen bonding with the central G (G0) in the other strand. However, such binding asymmetry should not be expected to translate into preferred orientation for Jun-Fos heterodimer due to the fact that R270 and R155 respectively occupy structurally equivalent positions within the α-helical basic regions of Jun and Fos, and are therefore able to exchange freely with each other as exquisitely demonstrated in the crystal structure [43]. In other words, the R270-C0 and R155-G0 contacts are energetically equivalent due to the involvement of an identical hydrogen bonding partner, guanidino moiety, in each case. Thus, although the lack of orientational preference upon the interaction of Jun-Fos heterodimer with TGACTCA motif can be rationalized in structural terms, the question as to how certain

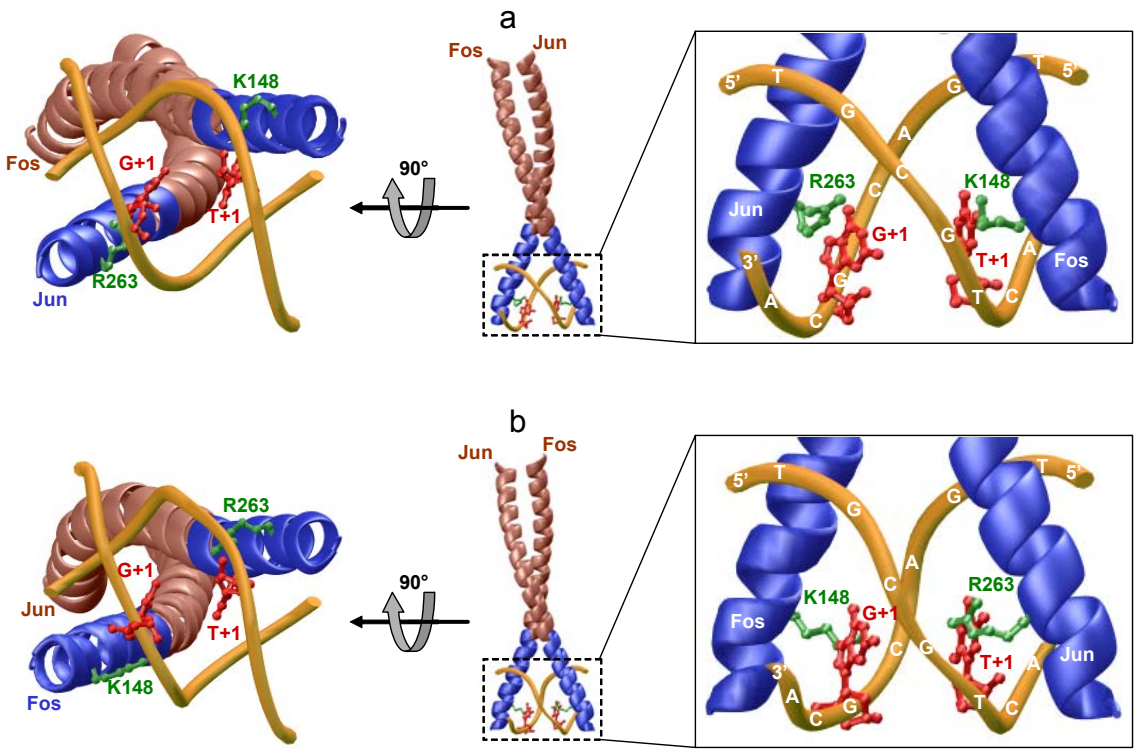

Figure 6-5. 3D structural models of bZIP domains of Jun-Fos heterodimer in complex with dsDNA oligos containing the TGACGCA motif in two possible orientations I (a) and II (b). The two orientations are related by a 180°-rotation of the Jun-Fos heterodimer about the dyad axis of symmetry. The backbones of LZ and BR subdomains of the bZIP heterodimer are colored brown and blue, respectively. The backbone of dsDNA is shown in yellow. The sidechains of R263 in Jun and K148 in Fos are colored green. The guanine at position +1 (G+1) in the sense strand TGACGCA and thymine at position +1 (T+1) in the antisense strand TGCGTCA are colored red. Insets show close-up views of contacts between specific bZIP residues with DNA bases.

single nucleotide substitutions within this motif might confer oriented binding raises further curiosity.

Our data presented here suggest strongly that the Jun-Fos heterodimer binds to CGACTCA, TGACTCG, TGTCTCA, TGACACA, TGCCTCA and TGACGCA variants of the TGACTCA motif in a preferred orientation (Figure 6-4). In an effort to gain insights into the structural basis of such preference for oriented binding, we modeled 3D structures of the bZIP domains of Jun-Fos heterodimer in complex with dsDNA oligos containing the TGACGCA motif in two possible orientations, hereinafter referred to as orientations I and II (Figure 6-5). Given that the T→G substitution at +1 position within

the TGACTCA motif to generate the TGACGCA motif would destroy the symmetric location of T at +1 position (T+1) within both strands, it might be reasonable to suspect that the oriented binding could result from the differential protein-DNA contacts at G+1 in the sense strand and its topologically equivalent counterpart T+1 in the antisense strand within the TGACGCA motif. Our structural analysis indeed reveals that G+1 in the sense strand and its topological equivalent T+1 in the antisense strand within the TGACGCA motif make non-equivalent contacts with Jun-Fos heterodimer. Thus, while G+1 in the sense strand hydrogen bonds with the guanidino moiety of R263 in Jun, T+1 in the antisense strand is involved in hydrogen bonding with the ε-amino group of K148 in orientation I (Figure 6-5a). These contacts are exquisitely mirrored by R263 and K148 in orientation II through the free exchange of Jun and Fos due to occupation of structurally equivalent positions within the basic regions (Figure 6-5b). Given the unique chemistry of guanidino moiety relative to ε-amino group, R263-G+1 and K148-T+1 contacts are likely to be energetically non-equivalent and thereby such energetic difference could favor the binding of Jun-Fos heterodimer to TGACGCA motif in only one of two possible orientations with important consequences on its transcriptional role. It is of worthy note that the TGACGCA element is found within the promoters of genes such as fibronectin 1 and sodium/iodide symporter [139, 140]. In the latter case, it is immediately flanked between the cis-elements for the TTF-1 and Pax-8 transcription factors [148-150]. Given such topological arrangement, it is conceivable that the precise orientation of Jun-Fos heterodimer in association with the TGACGCA element within the promoter of sodium/iodide symporter gene may be a critical determinant of the nature of other interacting cellular partners and hence gene expression.

Although our 3D structural models provide a compelling rationale for the binding of Jun-Fos heterodimer to the TGACGCA motif with a preferred orientation, it should be borne in mind that only structural analysis through experimental means can confirm the accuracy of such models and which one of the two possible orientations may actually be the preferred one. We also emphasize that for the binding of Jun-Fos heterodimer to other variants of the TGACTCA motif with a preferred orientation, R263 and K148 may not necessarily be responsible for differential protein-DNA contacts but rather the role of additional residues within the basic regions may have to be invoked. It has indeed been previously shown that amino acid substitutions within the basic regions of bZIP domains can result in oriented binding of Jun-Fos heterodimer to DNA [58]. In short, our study warrants extensive X-ray crystallographic analysis of Jun-Fos heterodimer in complex with specific TGACTCA variants to fully unravel the structural basis of the preference for oriented binding and, needless to say, our future efforts will be directed along these lines of curiosity.

6.5 Concluding remarks

Transcription factors do not act alone but rather in concert in a cooperative manner within the transcriptional initiation complex responsible for switching on gene expression. The architecture of such cooperative machinery parallels the design of composite sites within gene promoters for the recognition of a diverse array of transcription factors that often arrive in droves and in association with each other. The orientation of transcription factors relative to each other as well as DNA coupled with the sequence of specific cis-elements thus play a key role in gauging the transcriptional output in a spatial and temporal manner. Because of such inter-dependence and added

versatility, the cis-elements within a composite site may not necessarily contain an optimal sequence for binding to a corresponding transcription factor. Understanding how subtle nucleotide changes within such cis-elements may affect their binding energetics to trans-factors is thus of paramount importance to not only unraveling the regulatory mechanisms of transcriptional machinery but may also offer insights into the role of genetic variations, such as single nucleotide polymorphisms, in specific promoter regions between different individuals. It is this curiosity that led to the design of our current study.

Here, we report how single nucleotide variants of the TGACTCA consensus motif modulate energetics and orientation of binding of the AP1 transcription factor. The major conclusions of our study are that the single nucleotide substitutions within the TGACTCA motif do not abrogate the binding of Jun-Fos heterodimer. On the contrary, the Jun-Fos heterodimer binds to the TGACTCA variants with affinities in the physiologically relevant micromolar-submicromolar range. Given that TGACTCA variants are widely encountered within the gene promoters [132-144], we believe that such differential energetics of binding may have emerged as an evolutionary mechanism for the differential regulation of AP1-responsive genes. Our data also indicate that certain single nucleotide variants of the TGACTCA motif may also dictate the orientation of the Jun-Fos heterodimer in complex with DNA and that such a feat may be attributable to chemically distinct amino acid residues located in structurally equivalent positions within the basic regions of Jun and Fos. Supporting this corollary is the salient observation that amino acid substitutions within the basic regions of bZIP domains can result in oriented binding of Jun-Fos heterodimer to DNA [58]. Our present study thus mirrors this

previous finding in that specific nucleotide substitutions within the TGACTCA motif can also accomplish a similar fate on the Jun-Fos heterodimer.

In concluding, the demonstration that single nucleotide variants within the TGACTCA motif can modulate energetics and orientation of binding of Jun-Fos heterodimer suggests that it may be a general feature of other cis-elements acting within the gene promoters. Our study thus clearly provides a precedent for guiding the design of future experiments to expand our understanding of the regulatory mechanisms underlying the operation of the transcriptional machinery.

7 Chapter 7: Conclusion

Although the critical role of Jun-Fos heterodimeric transcription factor in cellular signaling was reported over two decades ago [4, 41, 48, 56, 78-81], thermodynamics of this key protein-DNA interaction have hitherto not been investigated. Knowledge of thermodynamics is central to the understanding of intrinsic forces that determine the structure and stability of protein-protein and protein-DNA interaction. Through this thesis, I have reported on our extensive thermodynamic analysis towards the goal of elucidating these significant biophysical forces.

Initially, we examined the role of the various thermodynamic forces at play in the binding of Jun-Fos heterodimer to DNA, as detailed analysis, using ITC, has never been performed on this interaction. Our study shows for the first time that binding of Jun-Fos heterodimer to TRE and CRE consensus elements is under enthalpic control and that this process is accompanied by an unfavorable loss of entropy at physiological temperatures. Furthermore, changes in SASA upon protein-DNA interaction determined from ITC measurements suggest strongly that the basic regions in the bZIP domains of Jun-Fos heterodimer are partially unstructured and become structured only upon interaction with DNA in a coupled folding and binding manner. This finding corroborates the notion that the coupled folding and DNA-binding is a general feature of the bZIP family of transcription factors [62, 63, 114-116].

It is well documented that variants of the TRE consensus motif exist in the promoters of genes under the control of Jun-Fos [132-144], yet poorly understood how they modulate binding. Only a single study, before ours, examined the effect of single nucleotide variants and through analysis of highly qualitative nature, demonstrated that

single nucleotide substitutions at specific positions within the TRE motif completely abrogated binding by the bZIP domains of Jun-Fos heterodimer [51]. In light of the fact that the overall binding energy results from the summation of all interactions between protein residues and DNA bases in a cooperative manner, we reasoned that single substitutions within the TRE consensus motif may result in the reduction of free energy associated with such protein-DNA interactions but are unlikely to lead to complete abrogation within the physiological context. Therefore, we examined how single nucleotide variants of the TRE consensus motif modulate energetics and orientation of binding of the AP1 transcription factor using the quantitative technique of ITC. The major conclusions from this section of the dissertation include the finding that, contrary to the previous study, single nucleotide substitutions within the TRE consensus motif do not abrogate the binding of Jun-Fos heterodimer. In fact, Jun-Fos heterodimer binds to the TRE variants with affinities in the physiologically relevant micromolar-submicromolar range. Given that such variants are widely encountered within the gene promoters [132-144], we believe that such differential energetics of binding may have emerged as an evolutionary mechanism for the differential regulation of AP1-responsive genes.

This notion is further supported by data that indicate that certain single nucleotide variants of the TRE consensus motif may also dictate the orientation of the Jun-Fos heterodimer in complex with DNA. Such a feat we attribute to chemically distinct amino acid residues located in structurally equivalent positions within the basic regions of Jun and Fos (Figure 6-5). The ability of single nucleotide variants to modulate orientation may have significant implications for transcriptional complexes involving Jun-Fos. Many

transcription factors bind asymmetrically to Jun and Fos [39]. In particular, the protein nFAT, as shown in an X-ray crystal structure, can only interact with Jun-Fos in one orientation [61]. This would suggest that some variants may optimally position or disallow interaction with nFAT. However, our conclusions are primarily drawn from data derived from 3D structural modeling. To achieve direct confirmation of our model we must solve the 3D structure, utilizing X-ray crystallography or NMR, of Jun-Fos in complex with the TGACGCA variant to determine if it indeed binds in a predominant orientation. In addition to this future work, our study of the interaction of Jun and Fos with single nucleotide variants introduces other questions. In particular, in contrast to the TRE consensus sequence, the extra central nucleotide within the CRE consensus sequence causes it to be fully palindromic, thus would single nucleotide variants of the CRE consensus sequence confer orientation? In general, how would such variants of the CRE consensus sequence modulate binding to Jun-Fos heterodimer? For either TRE or CRE, do flanking nucleotides play any role in modulated binding? In summary, our demonstration that single nucleotide variants within the TGACTCA motif modulate energetics and orientation of binding of Jun-Fos heterodimer suggests this phenomenon may be a general feature for cis-elements acting within gene promoters. It is also of worthy note that this study marks the first time anyone has performed detailed thermodynamic analysis of *any* transcription factor binding to single nucleotide variants of its consensus sequence. Thus, our study clearly provides a precedent for guiding the design of future experiments to expand our understanding of the regulatory mechanisms underlying the operation of transcriptional machinery.

Following our detailed thermodynamic analysis of Jun-Fos heterodimer binding to its DNA response elements, TRE and CRE, we further explored how binding energetics differ in the context of the biologically relevant Jun-Jun homodimer. Our data show that although the energetics are comparable between the two, unexpectedly, the Jun-Jun homodimer binds more tightly than the Jun-Fos heterodimer (the implications of which are discussed in chapter 6). Even more surprising is that our data support a model by which unfolded Jun monomers load onto DNA as monomers such that association with DNA triggers their folding and homodimerization. Given that Jun may largely exist as a monomer under physiological conditions due to its relatively weak dimer dissociation constant in the low micromolar range [47, 124, 129], this binding model may be the favorable pathway under physiological conditions. However, such a mechanism does not mutually exclude other models by which Jun may bind to DNA as a fully folded or partially folded homodimer and only direct kinetic analysis can provide information on the most preferred pathway under non-equilibrium conditions.

Our attempts to perform kinetic analysis, unfortunately, have met with difficulty arising from interference of circular dichroism signals due to our thioredoxin stability tags and subsequent proteolytic instability following our attempt to remove these tags. These difficulties have limited our use of kinetic analysis, yet there remains many areas for future work. The binding of single nucleotide variants of the TRE or CRE consensus sequence to Jun-Jun homodimer has never been explored. Such analysis would improve our understanding of the ability of Cis-acting DNA response elements to modulate binding given their interaction with two identical basic regions. Jun is also known to interact with 13 other human bZIP transcription factors [38], yet the thermodynamics of

these interactions have never been examined. Furthermore, our observations of the differences by which Jun-Jun homodimer loads onto DNA in contrast to the Jun-Fos heterodimer raise questions as to what would be the preferred model in the context of all possible configurations of the various bZIP heterodimers – could the ability to load onto DNA as monomers or dimers be the general mechanism of regulation? Jun-Fos loads onto DNA as a dimer, thus does Fos sequester Jun and prevent it from loading via the monomer pathway where it can interact with other bZIP proteins?

Finally, despite their discovery over two decades ago, the thermodynamics of heterodimerzation of Jun and Fos hitherto have not been characterized. For the first time ever, our thermodynamic analysis shows that the heterodimerization of Jun and Fos is under enthalpic control and accompanied by entropic penalty at physiological temperatures. Additionally, our data is highly suggestive of a model for the heterodimerization of Jun and Fos in which leucine zipper monomers are unfolded and form α-helices upon dimerization. Such observations are consistent with the notion that transcription factors exist in monomeric form before higher order assembly [64]. In so doing, transcription factors increase their capture radius, a measure of probability of interaction. They can also capitalize on the instability of the transition state intermediate, thus allowing transcription factors to discriminate between various binding sites [64]. Taken together, our study involving the heterodimerization of Jun-Fos provides novel insights into the thermodynamics of a key protein-protein interaction pertinent to cellular transcriptional machinery.

In conclusion, Jun and Fos are relevant proteins in many cellular pathways and have an established oncogenic potential. Our finding that Jun-Fos can tolerate single

nucleotide variants of its consensus sequence expands the list of potential targets genes and therefore pathways under the control of this transcription factor. For these reasons these proteins offer an exciting target for drug therapy. Current strategies for the design of drugs that can inhibit the oncogenic action of Jun-Fos heterodimer on cellular machinery are based on molecules that either interfere with the heterodimerization, or alternatively, compete with TRE and CRE sites for binding to Jun-Fos heterodimer. Through an extensive biophysical analysis of the various interactions, involving Jun and Fos, we have shown evidence suggesting large conformational changes are involved in dimerization and DNA binding. Taken together, the findings described within this dissertation may offer novel opportunities for the design of drugs that may lock Jun or Fos in a partially unstructured state so as to completely abrogate its DNA binding or dimerization ability and ultimately its oncogenic potential.

REFERENCES

1. Mahner, S., Baasch, C., Schwarz, J., Hein, S., Wolber, L., Janicke, F., and Milde-Langosch, K. (2008). c-Fos expression is a molecular predictor of progression and survival in epithelial ovarian carcinoma. Br J Cancer *99*, 1269-1275.

2. Platanias, L.C. (2003). Map kinase signaling pathways and hematologic malignancies. Blood *101*, 4667-4679.

3. Milde-Langosch, K. (2005). The Fos family of transcription factors and their role in tumourigenesis. Eur J Cancer *41*, 2449-2461.

4. Lee, W., Mitchell, P., and Tjian, R. (1987). Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. Cell *49*, 741-752.

5. Angel, P., and Karin, M. (1991). The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. Biochim Biophys Acta *1072*, 129-157.

6. Villard, J. (2004). Transcription regulation and human diseases. Swiss Med Wkly *134*, 571-579.

7. Zhang, Y., Pu, X., Shi, M., Chen, L., Song, Y., Qian, L., Yuan, G., Zhang, H., Yu, M., Hu, M., Shen, B., and Guo, N. (2007). Critical role of c-Jun overexpression in liver metastasis of human breast cancer xenograft model. BMC Cancer *7*, 145.

8. Shen, Q., Uray, I.P., Li, Y., Krisko, T.I., Strecker, T.E., Kim, H.T., and Brown, P.H. (2008). The AP-1 transcription factor regulates breast cancer cell growth via cyclins and E2F factors. Oncogene *27*, 366-377.

9. Prusty, B.K., and Das, B.C. (2005). Constitutive activation of transcription factor AP-1 in cervical cancer and suppression of human papillomavirus (HPV) transcription and AP-1 activity in HeLa cells by curcumin. Int J Cancer *113*, 951-960.

10. Divya, C.S., and Pillai, M.R. (2006). Antitumor action of curcumin in human papillomavirus associated cells involves downregulation of viral oncogenes, prevention of NFkB and AP-1 translocation, and modulation of apoptosis. Mol Carcinog *45*, 320-332.

11. Ivshina, A.V., George, J., Senko, O., Mow, B., Putti, T.C., Smeds, J., Lindahl, T., Pawitan, Y., Hall, P., Nordgren, H., Wong, J.E., Liu, E.T., Bergh, J., Kuznetsov, V.A., and Miller, L.D. (2006). Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. Cancer Res *66*, 10292-10301.

12. Sotiriou, C., Wirapati, P., Loi, S., Harris, A., Fox, S., Smeds, J., Nordgren, H., Farmer, P., Praz, V., Haibe-Kains, B., Desmedt, C., Larsimont, D., Cardoso, F., Peterse, H., Nuyten, D., Buyse, M., Van de Vijver, M.J., Bergh, J., Piccart, M., and Delorenzi, M. (2006). Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst *98*, 262-272.

13. Sorlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Thorsen, T., Quist, H., Matese, J.C., Brown, P.O., Botstein, D., Eystein Lonning, P., and Borresen-Dale, A.L. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A *98*, 10869-10874.

14. Finak, G., Bertos, N., Pepin, F., Sadekova, S., Souleimanova, M., Zhao, H., Chen, H., Omeroglu, G., Meterissian, S., Omeroglu, A., Hallett, M., and Park, M. (2008). Stromal gene expression predicts clinical outcome in breast cancer. Nat Med *14*, 518-527.

15. Dairkee, S.H., Ji, Y., Ben, Y., Moore, D.H., Meng, Z., and Jeffrey, S.S. (2004). A molecular 'signature' of primary breast cancer cultures; patterns resembling tumor tissue. BMC Genomics *5*, 47.

16. Smith, L.M., Wise, S.C., Hendricks, D.T., Sabichi, A.L., Bos, T., Reddy, P., Brown, P.H., and Birrer, M.J. (1999). c-Jun overexpression in MCF-7 breast cancer cells produces a tumorigenic, invasive and hormone resistant phenotype. Oncogene *18*, 6063-6070.

17. Briggs, J., Chamboredon, S., Castellazzi, M., Kerry, J.A., and Bos, T.J. (2002). Transcriptional upregulation of SPARC, in response to c-Jun overexpression, contributes to increased motility and invasion of MCF7 breast cancer cells. Oncogene *21*, 7077-7091.

18. Graham, J.D., Bain, D.L., Richer, J.K., Jackson, T.A., Tung, L., and Horwitz, K.B. (2000). Thoughts on tamoxifen resistant breast cancer. Are coregulators the answer or just a red herring? J Steroid Biochem Mol Biol *74*, 255-259.

19. Johnston, S.R., Lu, B., Scott, G.K., Kushner, P.J., Smith, I.E., Dowsett, M., and Benz, C.C. (1999). Increased activator protein-1 DNA binding and c-Jun NH2-terminal kinase activity in human breast tumors with acquired tamoxifen resistance. Clin Cancer Res *5*, 251-256.

20. Teyssier, C., Belguise, K., Galtier, F., and Chalbos, D. (2001). Characterization of the physical interaction between estrogen receptor alpha and Jun proteins. J Biol Chem *276*, 36361-36369.

21. Hazzalin, C.A., and Mahadevan, L.C. (2002). MAPK-regulated transcription: a continuously variable gene switch? Nat Rev Mol Cell Biol *3*, 30-40.

22. Mechta-Grigoriou, F., Gerald, D., and Yaniv, M. (2001). The mammalian Jun proteins: redundancy and specificity. Oncogene *20*, 2378-2389.

23. Turpaev, K.T. (2006). Role of transcription factor AP-1 in integration of cellular signaling systems. Mol Biol (Mosk) *40*, 945-961.

24. Vanhoutte, P., Barnier, J.V., Guibert, B., Pages, C., Besson, M.J., Hipskind, R.A., and Caboche, J. (1999). Glutamate induces phosphorylation of Elk-1 and CREB, along with c-fos activation, via an extracellular signal-regulated kinase-dependent pathway in brain slices. Mol Cell Biol *19*, 136-146.

25. Shaulian, E., and Karin, M. (2001). AP-1 in cell proliferation and survival. Oncogene *20*, 2390-2400.

26. Whitmarsh, A.J., and Davis, R.J. (1996). Transcription factor AP-1 regulation by mitogen-activated protein kinase signal transduction pathways. J Mol Med *74*, 589-607.

27. Pramanik, R., Qi, X., Borowicz, S., Choubey, D., Schultz, R.M., Han, J., and Chen, G. (2003). p38 isoforms have opposite effects on AP-1-dependent transcription through regulation of c-Jun. The determinant roles of the isoforms in the p38 MAPK signal specificity. J Biol Chem *278*, 4831-4839.

28. Franklin, C.C., Sanchez, V., Wagner, F., Woodgett, J.R., and Kraft, A.S. (1992). Phorbol ester-induced amino-terminal phosphorylation of human Jun but not JunB regulates transcriptional activation. Proc Natl Acad Sci U S A *89*, 7247-7251.

29. Coffey, E.T., Smiciene, G., Hongisto, V., Cao, J., Brecht, S., Herdegen, T., and Courtney, M.J. (2002). c-Jun N-terminal protein kinase (JNK) 2/3 is specifically activated by stress, mediating c-Jun activation, in the presence of constitutive JNK1 activity in cerebellar neurons. J Neurosci *22*, 4335-4345.

30. Morton, S., Davis, R.J., McLaren, A., and Cohen, P. (2003). A reinvestigation of the multisite phosphorylation of the transcription factor c-Jun. Embo J *22*, 3876-3886.

31. Vinciguerra, M., Esposito, I., Salzano, S., Madeo, A., Nagel, G., Maggiolini, M., Gallo, A., and Musti, A.M. (2008). Negative charged threonine 95 of c-Jun is essential for c-Jun N-terminal kinase-dependent phosphorylation of threonine 91/93 and stress-induced c-Jun biological activity. Int J Biochem Cell Biol *40*, 307-316.

32. Bannister, A.J., Brown, H.J., Sutherland, J.A., and Kouzarides, T. (1994). Phosphorylation of the c-Fos and c-Jun HOB1 motif stimulates its activation capacity. Nucleic Acids Res *22*, 5173-5176.

33. Tanos, T., Marinissen, M.J., Leskow, F.C., Hochbaum, D., Martinetto, H., Gutkind, J.S., and Coso, O.A. (2005). Phosphorylation of c-Fos by members of the p38 MAPK family. Role in the AP-1 response to UV light. J Biol Chem *280*, 18842-18852.

34. Lin, A., Frost, J., Deng, T., Smeal, T., al-Alawi, N., Kikkawa, U., Hunter, T., Brenner, D., and Karin, M. (1992). Casein kinase II is a negative regulator of c-Jun DNA binding and AP-1 activity. Cell *70*, 777-789.

35. Huang, C.C., Wang, J.M., Kikkawa, U., Mukai, H., Shen, M.R., Morita, I., Chen, B.K., and Chang, W.C. (2008). Calcineurin-mediated dephosphorylation of c-Jun Ser-243 is required for c-Jun protein stability and cell transformation. Oncogene *27*, 2422-2429.

36. Sasaki, T., Kojima, H., Kishimoto, R., Ikeda, A., Kunimoto, H., and Nakajima, K. (2006). Spatiotemporal regulation of c-Fos by ERK5 and the E3 ubiquitin ligase UBR1, and its biological role. Mol Cell *24*, 63-75.

37. Bossis, G., Malnou, C.E., Farras, R., Andermarcher, E., Hipskind, R., Rodriguez, M., Schmidt, D., Muller, S., Jariel-Encontre, I., and Piechaczyk, M. (2005). Down-regulation of c-Fos/c-Jun AP-1 dimer activity by sumoylation. Mol Cell Biol *25*, 6964-6979.

38. Newman, J.R., and Keating, A.E. (2003). Comprehensive identification of human bZIP interactions with coiled-coil arrays. Science *300*, 2097-2101.

39. Chinenov, Y., and Kerppola, T.K. (2001). Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity. Oncogene *20*, 2438-2452.

40. Abate, C., Luk, D., and Curran, T. (1991). Transcriptional regulation by Fos and Jun in vitro: interaction among multiple activator and regulatory domains. Mol Cell Biol *11*, 3624-3632.

41. Curran, T., and Franza, B.R., Jr. (1988). Fos and Jun: the AP-1 connection. Cell *55*, 395-397.

42. Halazonetis, T.D., Georgopoulos, K., Greenberg, M.E., and Leder, P. (1988). c-Jun dimerizes with itself and with c-Fos, forming complexes of different DNA binding affinities. Cell *55*, 917-924.

43. Glover, J.N., and Harrison, S.C. (1995). Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. Nature *373*, 257-261.

44. Mason, J.M., and Arndt, K.M. (2004). Coiled coil domains: stability, specificity, and biological implications. Chembiochem *5*, 170-176.

45. Graddis, T.J., Myszka, D.G., and Chaiken, I.M. (1993). Controlled formation of model homo- and heterodimer coiled coil polypeptides. Biochemistry *32*, 12664-12671.

46. Kohn, W.D., Kay, C.M., and Hodges, R.S. (1995). Protein destabilization by electrostatic repulsions in the two-stranded alpha-helical coiled-coil/leucine zipper. Protein Sci *4*, 237-250.

47. O'Shea, E.K., Rutkowski, R., and Kim, P.S. (1992). Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. Cell *68*, 699-708.

48. Lee, W., Haslinger, A., Karin, M., and Tjian, R. (1987). Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40. Nature *325*, 368-372.

49. Kwon, H., Park, S., Lee, S., Lee, D.K., and Yang, C.H. (2001). Determination of binding constant of transcription factor AP-1 and DNA. Application of inhibitors. Eur J Biochem *268*, 565-572.

50. John, M., Leppik, R., Busch, S.J., Granger-Schnarr, M., and Schnarr, M. (1996). DNA binding of Jun and Fos bZIP domains: homodimers and heterodimers induce a DNA conformational change in solution. Nucleic Acids Res *24*, 4487-4494.

51. Risse, G., Jooss, K., Neuberg, M., Bruller, H.J., and Muller, R. (1989). Asymmetrical recognition of the palindromic AP1 binding site (TRE) by Fos protein complexes. Embo J *8*, 3825-3832.

52. Kim, J., Tzamarias, D., Ellenberger, T., Harrison, S.C., and Struhl, K. (1993). Adaptability at the protein-DNA interface is an important aspect of sequence recognition by bZIP proteins. Proc Natl Acad Sci U S A *90*, 4513-4517.

53. Kim, H., Pennie, W.D., Sun, Y., and Colburn, N.H. (1997). Differential functional significance of AP-1 binding sites in the promoter of the gene encoding mouse tissue inhibitor of metalloproteinases-3. Biochem J *324 (Pt 2)*, 547-553.

54. Montminy, M.R., Sevarino, K.A., Wagner, J.A., Mandel, G., and Goodman, R.H. (1986). Identification of a cyclic-AMP-responsive element within the rat somatostatin gene. Proc Natl Acad Sci U S A *83*, 6682-6686.

55. Nakabeppu, Y., Ryder, K., and Nathans, D. (1988). DNA binding activities of three murine Jun proteins: stimulation by Fos. Cell *55*, 907-915.

56. Rauscher, F.J., 3rd, Voulalas, P.J., Franza, B.R., Jr., and Curran, T. (1988). Fos and Jun bind cooperatively to the AP-1 site: reconstitution in vitro. Genes Dev *2*, 1687-1699.

57. Nakabeppu, Y., and Nathans, D. (1989). The basic region of Fos mediates specific DNA binding. Embo J *8*, 3833-3841.

58. Leonard, D.A., Rajaram, N., and Kerppola, T.K. (1997). Structural basis of DNA bending and oriented heterodimer binding by the basic leucine zipper domains of Fos and Jun. Proc Natl Acad Sci U S A *94*, 4913-4918.

59. O'Shea, E.K., Klemm, J.D., Kim, P.S., and Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science *254*, 539-544.

60. Junius, F.K., O'Donoghue, S.I., Nilges, M., Weiss, A.S., and King, G.F. (1996). High resolution NMR solution structure of the leucine zipper domain of the c-Jun homodimer. J Biol Chem *271*, 13663-13667.

61. Chen, L., Glover, J.N., Hogan, P.G., Rao, A., and Harrison, S.C. (1998). Structure of the DNA-binding domains from NFAT, Fos and Jun bound specifically to DNA. Nature *392*, 42-48.

62. Weiss, M.A., Ellenberger, T., Wobbe, C.R., Lee, J.P., Harrison, S.C., and Struhl, K. (1990). Folding transition in the DNA-binding domain of GCN4 on specific binding to DNA. Nature *347*, 575-578.

63. Thompson, K.S., Vinson, C.R., and Freire, E. (1993). Thermodynamic characterization of the structural stability of the coiled-coil region of the bZIP transcription factor GCN4. Biochemistry *32*, 5491-5496.

64. Kohler, J.J., and Schepartz, A. (2001). Kinetic studies of Fos.Jun.DNA complex formation: DNA binding prior to dimerization. Biochemistry *40*, 130-142.

65. Berger, C., Piubelli, L., Haditsch, U., and Bosshard, H.R. (1998). Diffusion-controlled DNA recognition by an unfolded, monomeric bZIP transcription factor. FEBS Lett *425*, 14-18.

66. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res *31*, 3784-3788.

67. Dautrevaux, M., and Boulanger, Y. (1967). [Myoglobins]. Bull Soc Chim Biol (Paris) *49*, 949-983.

68. Cantor, C.R., Warshaw, M.M., and Shapiro, H. (1970). Oligonucleotide interactions. 3. Circular dichroism studies of the conformation of deoxyoligonucleotides. Biopolymers *9*, 1059-1077.

69. Edgcomb, S.P., and Murphy, K.P. (2000). Structural energetics of protein folding and binding. Curr Opin Biotechnol *11*, 62-66.

70. Murphy, K.P., and Freire, E. (1992). Thermodynamics of structural stability and cooperative folding behavior in proteins. Adv Protein Chem *43*, 313-361.

71. Xie, D., and Freire, E. (1994). Molecular basis of cooperativity in protein folding. V. Thermodynamic and structural conditions for the stabilization of compact denatured states. Proteins *19*, 291-301.

72. Murphy, K.P. (1999). Predicting binding energetics from structure: looking beyond DeltaG degrees. Med Res Rev *19*, 333-339.

73. Freire, E. (1993). Structural thermodynamics: prediction of protein stability and protein binding affinities. Arch Biochem Biophys *303*, 181-184.

74. Wiseman, T., Williston, S., Brandts, J.F., and Lin, L.N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. Anal Biochem *179*, 131-137.

75. Baker, B.M., and Murphy, K.P. (1996). Evaluation of linked protonation effects in protein binding reactions using isothermal titration calorimetry. Biophys J *71*, 2049-2055.

76. Marti-Renom, M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F., and Sali, A. (2000). Comparative Protein Structure Modeling of Genes and Genomes. Annu. Rev. Biophys. Biomol. Struct. *29*, 291-325.

77. Carson, M. (1991). Ribbons 2.0. J. Appl. Crystallogr. *24*, 958-961.

78. Bohmann, D., Bos, T.J., Admon, A., Nishimura, T., Vogt, P.K., and Tjian, R. (1987). Human proto-oncogene c-Jun encodes a DNA binding protein with structural and functional properties of transcription factor AP-1. Science *238*, 1386-1392.

79. Angel, P., Allegretto, E.A., Okino, S.T., Hattori, K., Boyle, W.J., Hunter, T., and Karin, M. (1988). Oncogene Jun encodes a sequence-specific trans-activator similar to AP-1. Nature *332*, 166-171.

80. Franza, B.R., Jr., Rauscher, F.J., 3rd, Josephs, S.F., and Curran, T. (1988). The Fos complex and Fos-related antigens recognize sequence elements that contain AP-1 binding sites. Science *239*, 1150-1153.

81. Rauscher, F.J., 3rd, Cohen, D.R., Curran, T., Bos, T.J., Vogt, P.K., Bohmann, D., Tjian, R., and Franza, B.R., Jr. (1988). Fos-associated protein p39 is the product of the Jun proto-oncogene. Science *240*, 1010-1016.

82. Zhou, H., Zarubin, T., Ji, Z., Min, Z., Zhu, W., Downey, J.S., Lin, S., and Han, J. (2005). Frequency and distribution of AP-1 sites in the human genome. DNA Res *12*, 139-150.

83. Baxevanis, A.D., and Vinson, C.R. (1993). Interactions of coiled coils in transcription factors: where is the specificity? Curr Opin Genet Dev *3*, 278-285.

84. Raivich, G., and Behrens, A. (2006). Role of the AP-1 transcription factor c-Jun in developing, adult and injured brain. Prog Neurobiol *78*, 347-363.

85. Alani, R., Brown, P., Binetruy, B., Dosaka, H., Rosenberg, R.K., Angel, P., Karin, M., and Birrer, M.J. (1991). The transactivating domain of the c-Jun proto-oncoprotein is required for cotransformation of rat embryo cells. Mol Cell Biol *11*, 6286-6295.

86. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In The Proteomics Protocols Handbook, J.M. Walker, ed. (Humana Press), pp. 571-607.

87. Fraczkiewicz, R., and Braun, W. (1998). Exact and efficient analytical calculation of the accessible surface area and their gradients for macromolecules. J. Comp. Chem. *19*, 319-333.

88. Koradi, R., Billeter, M., and Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph *14*, 51-55, 29-32.

89. Ryseck, R.P., and Bravo, R. (1991). c-Jun, Jun B, and Jun D differ in their binding affinities to AP-1 and CRE consensus sequences: effect of Fos proteins. Oncogene *6*, 533-542.

90. Berger, C., Jelesarov, I., and Bosshard, H.R. (1996). Coupled folding and site-specific binding of the GCN4-bZIP transcription factor to the AP-1 and ATF/CREB DNA sites studied by microcalorimetry. Biochemistry *35*, 14984-14991.

91. Dragan, A.I., Frank, L., Liu, Y., Makeyeva, E.N., Crane-Robinson, C., and Privalov, P.L. (2004). Thermodynamic signature of GCN4-bZIP binding to DNA indicates the role of water in discriminating between the AP-1 and ATF/CREB sites. J Mol Biol *343*, 865-878.

92. Wiseman, T., Williston, S., Brandts, J.F., and Lin, L.N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. Anal. Biochem. *179*, 131-137.

93. Ha, J.H., Spolar, R.S., and Record, M.T., Jr. (1989). Role of the hydrophobic effect in stability of site-specific protein-DNA complexes. J Mol Biol *209*, 801-816.

94. Ladbury, J.E., Wright, J.G., Sturtevant, J.M., and Sigler, P.B. (1994). A thermodynamic study of the trp repressor-operator interaction. J Mol Biol *238*, 669-681.

95. Foguel, D., and Silva, J.L. (1994). Cold denaturation of a repressor-operator complex: the role of entropy in protein-DNA recognition. Proc Natl Acad Sci U S A *91*, 8244-8247.

96. Petri, V., Hsieh, M., and Brenowitz, M. (1995). Thermodynamic and kinetic characterization of the binding of the TATA binding protein to the adenovirus E4 promoter. Biochemistry *34*, 9977-9984.

97. Merabet, E., and Ackers, G.K. (1995). Calorimetric analysis of lambda cI repressor binding to DNA operator sites. Biochemistry *34*, 8554-8563.

98. Sieber, M., and Allemann, R.K. (2000). Thermodynamics of DNA binding of MM17, a 'single chain dimer' of transcription factor MASH-1. Nucleic Acids Res *28*, 2122-2127.

99. Datta, K., and LiCata, V.J. (2003). Thermodynamics of the binding of Thermus aquaticus DNA polymerase to primed-template DNA. Nucleic Acids Res *31*, 5590-5597.

100. Datta, K., Wowor, A.J., Richard, A.J., and LiCata, V.J. (2006). Temperature dependence and thermodynamics of Klenow polymerase binding to primed-template DNA. Biophys J *90*, 1739-1751.

101. Wang, X., Cao, W., Cao, A., and Lai, L. (2003). Thermodynamic characterization of the folding coupled DNA binding by the monomeric transcription activator GCN4 peptide. Biophys J *84*, 1867-1875.

102. Milev, S., Bosshard, H.R., and Jelesarov, I. (2005). Enthalpic and entropic effects of salt and polyol osmolytes on site-specific protein-DNA association: the integrase Tn916-DNA complex. Biochemistry *44*, 285-293.

103. Milev, S., Gorfe, A.A., Karshikoff, A., Clubb, R.T., Bosshard, H.R., and Jelesarov, I. (2003). Energetics of sequence-specific protein-DNA association: binding of integrase Tn916 to its target DNA. Biochemistry *42*, 3481-3491.

104. Patel, L., Abate, C., and Curran, T. (1990). Altered protein conformation on DNA binding by Fos and Jun. Nature *347*, 572-575.

105. Siebert, X., and Amzel, L.M. (2004). Loss of translational entropy in molecular associations. Proteins *54*, 104-115.

106. Murphy, K.P., Xie, D., Thompson, K.S., Amzel, L.M., and Freire, E. (1994). Entropy in biological binding processes: estimation of translational entropy loss. Proteins *18*, 63-67.

107. Tamura, A., and Privalov, P.L. (1997). The Entropy cost of protein association. J. Mol. Biol. *273*, 1048-1060.

108. Murphy, K.P., Bhakuni, V., Xie, D., and Freire, E. (1992). Molecular basis of co-operativity in protein folding. III. Structural identification of cooperative folding units and folding intermediates. J Mol Biol *227*, 293-306.

109. Spolar, R.S., and M.T. Record, J. (1994). Coupling of local folding to site-specific binding of proteins to DNA. Science *263*, 777-784.

110. Privalov, P.L., and Gill, S.J. (1988). Stability of protein structure and hydrophobic interaction. Adv Protein Chem *39*, 191-234.

111. Spolar, R.S., Ha, J.H., and Record, M.T., Jr. (1989). Hydrophobic effect in protein folding and other noncovalent processes involving proteins. Proc Natl Acad Sci U S A *86*, 8382-8385.

112. Privalov, P.L., and Makhatadze, G.I. (1992). Contribution of hydration and non-covalent interactions to the heat capacity effect on protein unfolding. J Mol Biol *224*, 715-723.

113. Spolar, R.S., Livingstone, J.R., and Record, M.T., Jr. (1992). Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. Biochemistry *31*, 3947-3955.

114. Weiss, M.A. (1990). Thermal unfolding studies of a leucine zipper domain and its specific DNA complex: implications for scissor's grip recognition. Biochemistry *29*, 8020-8024.

115. Bosshard, H.R., Durr, E., Hitz, T., and Jelesarov, I. (2001). Energetics of coiled coil folding: the nature of the transition states. Biochemistry *40*, 3544-3552.

116. Saudek, V., Pastore, A., Castiglione Morelli, M.A., Frank, R., Gausepohl, H., Gibson, T., Weih, F., and Roesch, P. (1990). Solution structure of the DNA-binding domain of the yeast transcriptional activator protein GCN4. Protein Eng *4*, 3-10.

117. Diebold, R.J., Rajaram, N., Leonard, D.A., and Kerppola, T.K. (1998). Molecular basis of cooperative DNA bending and oriented heterodimer binding in the NFAT1-Fos-Jun-ARRE2 complex. Proc Natl Acad Sci U S A *95*, 7915-7920.

118. Leonard, D.A., and Kerppola, T.K. (1998). DNA bending determines Fos-Jun heterodimer orientation. Nat Struct Biol *5*, 877-881.

119. Ransone, L.J., Visvader, J., Wamsley, P., and Verma, I.M. (1990). Trans-dominant negative mutants of Fos and Jun. Proc Natl Acad Sci U S A *87*, 3806-3810.

120. Kim, J., and Struhl, K. (1995). Determinants of half-site spacing preferences that distinguish AP-1 and ATF/CREB bZIP domains. Nucleic Acids Res *23*, 2531-2537.

121. Konig, P., and Richmond, T.J. (1993). The X-ray structure of the GCN4-bZIP bound to ATF/CREB site DNA shows the complex depends on DNA flexibility. J Mol Biol *233*, 139-154.

122. Seldeen, K.L., McDonald, C.B., Deegan, B.J., and Farooq, A. (2008). Coupling of folding and DNA-binding in the bZIP domains of Jun-Fos heterodimeric transcription factor. Arch Biochem Biophys *473*, 48-60.

123. Pernelle, C., Clerc, F.F., Dureuil, C., Bracco, L., and Tocque, B. (1993). An efficient screening assay for the rapid and precise determination of affinities between leucine zipper domains. Biochemistry *32*, 11682-11687.

124. Patel, L.R., Curran, T., and Kerppola, T.K. (1994). Energy transfer analysis of Fos-Jun dimerization and DNA binding. Proc Natl Acad Sci U S A *91*, 7360-7364.

125. van Dam, H., and Castellazzi, M. (2001). Distinct roles of Jun:Fos and Jun:ATF dimers in oncogenesis. Oncogene *20*, 2453-2464.

126. Kataoka, K., Noda, M., and Nishizawa, M. (1994). Maf nuclear oncoprotein recognizes sequences related to an AP-1 site and forms heterodimers with both Fos and Jun. Mol Cell Biol *14*, 700-712.

127. Gomez, J., Hilser, V.J., Xie, D., and Freire, E. (1995). The heat capacity of proteins. Proteins *22*, 404-412.

128. Uedaira, H., Kono, H., Ponraj, P., Kitajima, K., and Sarai, A. (2003). Structure-Thermodynamic relationship in protein-DNA binding: Heat capacity change. Genome Informatics *14*, 510-511.

129. O'Shea, E.K., Rutkowski, R., Stafford, W.F., 3rd, and Kim, P.S. (1989). Preferential heterodimer formation by isolated leucine zippers from fos and jun. Science *245*, 646-648.

130. Marti, D.N., Jelesarov, I., and Bosshard, H.R. (2000). Interhelical ion pairing in coiled coils: solution structure of a heterodimeric leucine zipper and determination of pKa values of Glu side chains. Biochemistry *39*, 12804-12818.

131. Marti, D.N., and Bosshard, H.R. (2004). Inverse electrostatic effect: electrostatic repulsion in the unfolded state stabilizes a leucine zipper. Biochemistry *43*, 12436-12447.

132. Hengerer, B., Lindholm, D., Heumann, R., Ruther, U., Wagner, E.F., and Thoenen, H. (1990). Lesion-induced increase in nerve growth factor mRNA is mediated by c-fos. Proc Natl Acad Sci U S A *87*, 3899-3903.

133. Therrien, M., and Drouin, J. (1991). Pituitary pro-opiomelanocortin gene expression requires synergistic interactions of several regulatory elements. Mol Cell Biol *11*, 3492-3503.

134. Koike, M., Kuroki, T., and Nose, K. (1993). Common target for 12-O-tetradecanoylphorbol-13-acetate and ras in the transcriptional enhancer of the growth factor-inducible JE gene. Mol Carcinog *8*, 105-111.

135. Fabre, S., Manin, M., Pailhoux, E., Veyssiere, G., and Jean, C. (1994). Identification of a functional androgen response element in the promoter of the gene for the androgen-regulated aldose reductase-like protein specific to the mouse vas deferens. J Biol Chem *269*, 5857-5864.

136. Shyy, J.Y., Lin, M.C., Han, J., Lu, Y., Petrime, M., and Chien, S. (1995). The cis-acting phorbol ester "12-O-tetradecanoylphorbol 13-acetate"-responsive element is involved in shear stress-induced monocyte chemotactic protein 1 gene expression. Proc Natl Acad Sci U S A *92*, 8069-8073.

137. Larose, M., Cassard-Doulcier, A.M., Fleury, C., Serra, F., Champigny, O., Bouillaud, F., and Ricquier, D. (1996). Essential cis-acting elements in rat uncoupling protein gene are in an enhancer containing a complex retinoic acid response domain. J Biol Chem *271*, 31533-31542.

138. Goldberg, D., Polly, P., Eisman, J.A., and Morrison, N.A. (1996). Identification of an osteocalcin gene promoter sequence that binds AP1. J Cell Biochem *60*, 447-457.

139. Moriguchi, Y., Matsubara, H., Mori, Y., Murasawa, S., Masaki, H., Maruyama, K., Tsutsumi, Y., Shibasaki, Y., Tanaka, Y., Nakajima, T., Oda, K., and Iwasaka, T. (1999). Angiotensin II-induced transactivation of epidermal growth factor receptor regulates fibronectin and transforming growth factor-beta synthesis via transcriptional and posttranscriptional mechanisms. Circ Res *84*, 1073-1084.

140. Chun, J.T., Di Dato, V., D'Andrea, B., Zannini, M., and Di Lauro, R. (2004). The CRE-like element inside the 5'-upstream region of the rat sodium/iodide symporter gene interacts with diverse classes of b-Zip molecules that regulate transcriptional activities through strong synergy with Pax-8. Mol Endocrinol *18*, 2817-2829.

141. Hayama, A., Uchida, S., Sasaki, S., and Marumo, F. (2000). Isolation and characterization of the human CLC-5 chloride channel gene promoter. Gene *261*, 355-364.

142. Borner, C., Hollt, V., and Kraus, J. (2002). Involvement of activator protein-1 in transcriptional regulation of the human mu-opioid receptor gene. Mol Pharmacol *61*, 800-805.

143. Wang, S.E., Wu, F.Y., Chen, H., Shamay, M., Zheng, Q., and Hayward, G.S. (2004). Early activation of the Kaposi's sarcoma-associated herpesvirus RTA, RAP, and MTA promoters by the tetradecanoyl phorbol acetate-induced AP1 pathway. J Virol *78*, 4248-4267.

144. Dwivedi, P.P., Anderson, P.H., Omdahl, J.L., Grimes, H.L., Morris, H.A., and May, B.K. (2005). Identification of growth factor independent-1 (GFI1) as a repressor of 25-hydroxyvitamin D 1-alpha hydroxylase (CYP27B1) gene expression in human prostate cancer cells. Endocr Relat Cancer *12*, 351-365.

145. Bream, J.H., Ping, A., Zhang, X., Winkler, C., and Young, H.A. (2002). A single nucleotide polymorphism in the proximal IFN-gamma promoter alters control of gene transcription. Genes Immun *3*, 165-169.

146. Bond, G.L., Hu, W., Bond, E.E., Robins, H., Lutzker, S.G., Arva, N.C., Bargonetti, J., Bartel, F., Taubert, H., Wuerl, P., Onel, K., Yip, L., Hwang, S.J., Strong, L.C., Lozano, G., and Levine, A.J. (2004). A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. Cell *119*, 591-602.

147. Chytil, M., Peterson, B.R., Erlanson, D.A., and Verdine, G.L. (1998). The orientation of the AP-1 heterodimer on DNA strongly affects transcriptional potency. Proc Natl Acad Sci U S A *95*, 14076-14081.

148. Taki, K., Kogai, T., Kanamoto, Y., Hershman, J.M., and Brent, G.A. (2002). A thyroid-specific far-upstream enhancer in the human sodium/iodide symporter gene requires Pax-8 binding and cyclic adenosine 3',5'-monophosphate response element-like sequence binding proteins for full activity and is differentially regulated in normal and thyroid cancer cells. Mol Endocrinol *16*, 2266-2282.

149. Schmitt, T.L., Espinoza, C.R., and Loos, U. (2002). Characterization of a thyroid-specific and cyclic adenosine monophosphate-responsive enhancer far upstream from the human sodium iodide symporter gene. Thyroid *12*, 273-279.

150. Ohno, M., Zannini, M., Levy, O., Carrasco, N., and di Lauro, R. (1999). The paired-domain transcription factor Pax8 binds to the upstream enhancer of the rat sodium/iodide symporter gene and participates in both thyroid-specific and cyclic-AMP-dependent transcription. Mol Cell Biol *19*, 2051-2060.