2011-05-31

Biophysical Studies of the Binding of ERα Nuclear Receptor to DNA

Brian J. Deegan
University of Miami, bdeeganphd@gmail.com

UNIVERSITY OF MIAMI

BIOPHYSICAL STUDIES OF THE BINDING OF ERα NUCLEAR RECEPTOR TO DNA

By

Brian J. Deegan

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 2011

UNIVERSITY OF MIAMI

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

BIOPHYSICAL STUDIES OF THE BINDING OF ERα NUCLEAR RECEPTOR TO DNA

Brian J. Deegan

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. Associate Professor of Biochemistry & Molecular Biology	Vineet Gupta, Ph.D. Assistant Professor of Medicine and Biochemistry & Molecular Biology
Alan Pollack, M.D., Ph.D. Chairman and Professor of Radiation Oncology	Mansoor M. Ahmed, Ph.D. Associate Professor of Radiation Oncology

DEEGAN, BRIAN J. (Ph.D., Cancer Biology)
Biophysical Studies of the Binding of ERα Nuclear Receptor to DNA (June 2011)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Associate Professor Amjad Farooq.
No. of pages in text. (124)

Estrogen receptor α (ERα) is a member of a family of ligand-modulated transcription factors that have come to be known as nuclear receptors. ERα mediates the action of estrogens and plays an integral role in a wide range of physiological processes ranging from embryonic development and morphogenesis to reproduction to cardiovascular health. Not surprisingly, malfunction of the estrogen system is associated with a host of pathological conditions such as osteoporosis, heart disease and most notably breast cancer. Essential to its functioning as a transcription factor are specific protein-DNA interactions which are mediated by the binding of the DNA-binding (DB) domain of ERα to particular DNA sequences located within target gene promoters called estrogen response elements (EREs). Here, using a diverse array of biophysical techniques, including in particular isothermal titration calorimetry coupled with molecular modeling and semi-empirical analysis, I provide new insights into the ERα-DNA interaction in thermodynamic and structural terms.

My data show that the binding of the DB domain of ERα to DNA is coupled to protonation at two specific amino acids, H196 and E203. Protonation of these residues is non-trivial and is required for high affinity binding. Amino acid sequence alignment of

the DB domains of the NR family suggests that this may be a hallmark feature common to the functioning of all nuclear receptors. Furthermore, I demonstrate that the DB domain can tolerate all single nucleotide substitutions within the ERE and bind in the physiologically relevant nanomolar to micromolar range. Comparative thermodynamic analysis reveals that the DB domain binds to these ERE sequences utilizing a considerable range of energetic signatures such that any one thermodynamic component of binding is not predictive of associated affinity. In addition, it is shown that nucleotide substitution results in significant changes in secondary and three-dimensional features of the oligonucleotides and may impact binding affinity. Finally, I demonstrate that the zinc-finger of the DB domain of ERα is relatively promiscuous and can accommodate several heavy-metal divalent cations. Other than zinc, only DB domains reconstituted with cobalt, cadmium and mercury were capable of binding DNA. Incorporation of the metals resulted in a wide range of CD spectroscopic features which were found not to be predictive of DNA binding capacity. Thus, isostructure does not equate to isofunction in the case of metal reconstituted DB domain of ERα. This analysis suggests that metal coordination is not likely to be required for domain folding, but rather is required to bind DNA.

Taken together, this thesis provides novel insights into the physicochemical basis of a key protein-DNA interaction essential to human health and disease. My studies bear the potential to impact the development of novel therapies harboring greater efficacy coupled with lower toxicity for the treatment of disease.

DEDICATION

To my best friend and wife Jocelyn, for her continuous love, support and inspiration.

ACKNOWLEDGEMENTS

It was Theodore Roosevelt who boldly suggested that nothing in life was worth doing less it mean effort, difficulty and pain. The work of a doctoral thesis certainly encompasses those aforementioned plights but in an unrelenting and prolonged manner. It would then stand to reason that this Ph.D. is really "worth it." As with any seemingly insurmountable task, besting such a chore is only made possible by those around you.

I would like to express my deepest gratitude to Dr. Amjad Farooq not only for his precise and skilled mentorship, but also for his belief in me. Amjad's vision to provide an enriching and efficient scientific environment proved critical to the success of my work. Perhaps what's more is the infectious passion with which he approaches science that provided the motivational "fire" to tackle scientific problems on a daily basis. It has also been a tremendous privilege to work with Dr. Ken Seldeen, Dr. Caleb McDonald, Vikas Bhat, David Mikles, Anna Bona and Samantha McIntosh. I am deeply indebted to each of you for your unwavering help and support. It is because of you that I look back at the countless hours in lab with great fondness. I am truly grateful for your friendships. The amount of fun and laughter we shared is what made my time in the lab truly unforgettable.

I am also greatly appreciative of my thesis committee members Dr. Thomas K. Harris, Dr. Alan Pollack, Dr. Vineet Gupta and Dr Mansoor Ahmed. Their experience, expertise, and vision have helped shape my work. I have benefited tremendously as direct result of their efforts and I am grateful for their time, participation, and help throughout this process.

I would like to acknowledge the National Institutes of Health, the American Heart Association, and the Braman Family Breast Cancer Institute for helping to fund my work. Without funding from these organizations, my work would not have been possible.

I would also like to thank my parents, John and Ann, my brother, Justin, and his wife, Amie. This process would have been much more difficult without their loving words and support. For this, I am appreciative.

Most of all, I am truly indebted to my ever-loving wife Jocelyn for the kindness, understanding, patience, encouragement, inspiration and hope she gives me each and every day.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction

1.1 ERα modulates gene expression via several pathways

The adult human is comprised of 10^{13}-10^{14} cells which form the mature cellular layers and specialized tissues and organs [*1*]. The ability to successfully execute a complex body-plan is exquisitely dependent on cellular language and ability to function cohesively. Therefore it is unsurprising that many biomolecules act as signals. At its core, intercellular communication involves the collection, integration and execution of a program based upon extracellular inputs. How such signals are propagated into the cellular milieu involves a variety of signal transduction modalities and multicellular eukaryotic organisms have developed several approaches to facilitate cellular communication at both the local/regional and distant/organismal levels. Some cell signaling molecules, such as cytokines and growth factors, are hydrophilic and act upon cell surface receptors which typically initiate a cascade of protein-protein interactions to affect cellular response. Other biomolecules are hydrophobic and are permitted to diffuse across the cellular membrane where they encounter cognate intracellular receptors typically exerting their effects directly on gene output. Examples of these lipophilic messengers are steroid hormones, vitamin-D_3, thyroxine, retinoic acid and fatty acid derivatives such as the leukotrienes and eicosanoids. Despite the chemical diversity of the aforementioned molecules, they act via strikingly similar molecular mechanisms involving the nuclear receptor (NR) superfamily.

Estrogens (estradiol, estriol, estrone) are steroid hormones whose biological effects are mediated by the estrogen receptor alpha (ERα). The predominant endogenous estrogen is 17β-estradiol and is what is typically implied by the term estrogen. ERα is a

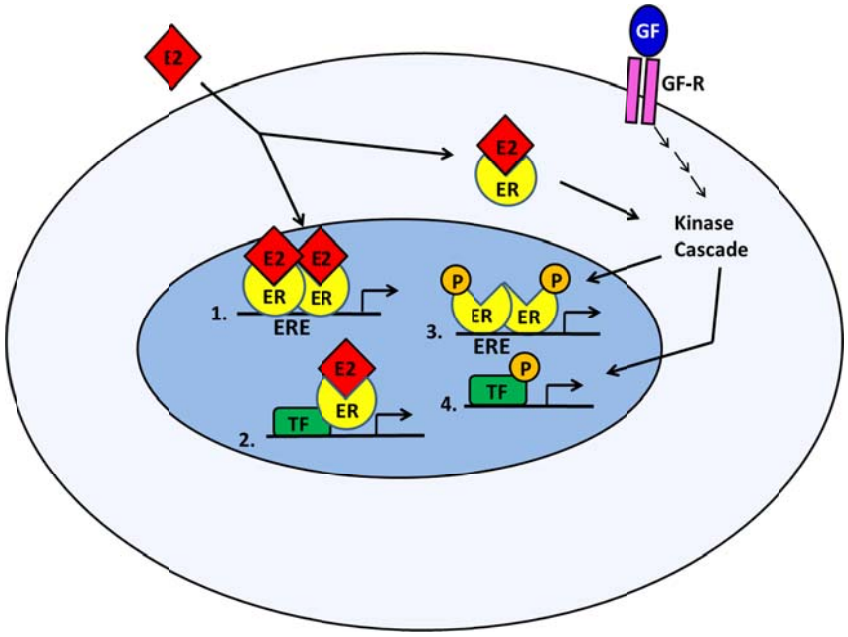

Figure 1-1: A schematic illustrating the role of ERα in cellular signaling through various mechanisms: (1) Classical mechanism of ER action. Upon its diffusion through the cell membreane, estradiol (E2) binds to ERα and induces dimerization. The resulting E2–ERα dimer in turn binds directly to EREs at target gene promoters. (2) ERE-independent genomic action. E2-complexed ERα can engage in "tethering" through protein-protein interactions thereby binding a transcription factor (TF) engaging its promoter. (3) Ligand-independent genomic actions. Growth factors (GF) bind to GF-receptors (GF-R) and activate protein-kinase cascades. Activated kinases can then phosphorylate (P) and activate nuclear ER at EREs. (4) Nongenomic actions. Membrane and cytosolic E2-ERα complexes can also activate protein-kinase cascades, thereby modulating protein function, *e.g.* activation of eNOS or regulation of gene expression through phosphorylation (P) and activation of a TF. Adapted from [*2*]

member of the ligand-modulated transcription factors collectively known as the NR superfamily [*3-6*]. Recent studies indicate ERα regulates biological processes by utilizing a variety of signaling pathways (Figure 1-1) [*7, 8*]. Classically, ERα is activated by binding estradiol followed by receptor dimerization and subsequent engagement of estrogen response elements (EREs) located within the promoters of target genes thereby influencing gene expression [*9-11*]. It should be noted binding of estradiol induces conformational change within the ligand binding (LB) domain of ERα which allows for the recruitment of transcriptional co-activators and co-repressors to ERE-containing genes [*12*]. The cellular context of co-activators and co-repressors in which ERα operates is one mechanism by which ligand-binding can produce differential responses in

different tissues. Additionally, ERα can influence genes not containing an ERE by act as a transcriptional co-regulator via protein-protein interactions with other transcription factor complexes in a ligand-dependent manner. Around one-third of genes which show estrogen-dependence do not harbor an ERE-like promoter motif [*13*] which suggests this indirect genomic function of ERα can have significant biological impact. Typical examples of transcription factors influenced by liganded-ERα include the activator protein-1 (AP1, Jun/Fos) [*14*] and specificity protein-1 (SP-1) [*15*]. Genes under AP1 control upregulated in response to estradiol include Cyclin D1 [*16*], insulin-like growth factor-1 (IGF-1) [*17*], and collagenase [*18*]. An example of an SP-1 dependent gene influenced by ERα is the low-density lipoprotein receptor (LDL-R) [*19*].

Estrogens also induce rapid cellular responses which occur on the time scale of seconds to minutes after estrogen exposure. These effects are much too fast to involve transcription. Non-genomic effects often employ cellular kinases, phosphatases, ion flux and secondary messengers to elicit a timely response [*2, 8*]. Such rapid cellular phenomena are commonly elicited by steroid hormones [*2, 20, 21*]. Several cell lines have been studied demonstrating mitogen activated protein kinase (MAPK) activation in response to estradiol [*22-25*]. Additional actions of ERα reported include activation of adenylyl cyclase leading to increased levels of cyclic andnoside monophosphate (cAMP) [*26*] as well as intracellular calcium ion release [*27*].

Finally, as opposed to the above estrogen-dependent mechanisms, ERα can also be activated in response to cellular kinases which phosphorylate ERα itself and/or one or more of its coregulators thereby affecting gene expression [*8, 28, 29*]. This adds another

layer of complexity likely supplementing classical estrogen-dependent gene expression and permits ERα to carry out its biological actions even in the absence of ligand.

It is clear that the ability of ERα to impact gene expression and cellular physiology is an integrated process consisting of both genomic and nongenomic events which often intersect. The eventual cellular response is therefore a dependent upon the cellular context in which ERα operates. The amounts of receptor, transcription factors, co-activators and co-repressors, the presence of growth factor stimuli and elicited kinase activity all play a role in determining the cellular reaction to estrogens. Given that the molecular mechanisms in response to estrogen are multifactorial it should come as no surprise the wide array of biology estrogen and ERα influence.

1.2 ERα plays a significant role in health and many diseases

Estrogen is a hormone that is produced mainly by the ovaries but exerts its effects systemically on distant tissues. It has been long acknowledged the role estrogen plays in the development of female secondary sexual characteristics and reproductive capacity. However, the efforts of gene targeting and transgenic mouse models have drastically expanded our knowledge of ERα biology. Estrogen and its receptor are now implicated in several physiological processes including cardiovascular health, bone homeostasis, and neurological function [*30*]. Given its broad physiological scope, estrogen acting via ERα is central to the initiation or progression of several pathological processes including but not limited to multiple types of cancers, coronary artery disease, osteoporosis, neurodegenerative disease, autoimmune conditions, and obesity [*30, 31*].

1.2.1 Reproductive Health

The results of the estrogen receptor knockout (ERKO) studies serve as an important lesson in the definition of estrogen biology [*32*]. ERα knockout, not surprisingly, results in female infertility as well as non-development of breast tissue [*32, 33*]. Murine females are born with a rudimentary epithelial ductal system embedded in stromal tissue. During puberty, rises in ovarian steroids induce maturation and elongation of the ductal epithelium leading to the development of secondary sexual characteristics. When the ERα is knocked out, proper breast tissue fails to develop and remains rudimentary through life [*34, 35*]. Function of ERα is also requisite for proper regulation of the menstrual cycle. In particular, ERα is critical to a negative feedback loop which down-regulates pituitary-derived luteinizing hormone leading to ovulation. Knockout of ERα leads to a disruption of the balance of the so-called hypothalamo-pituitary-gonadal axis leading to an anovulatory setting. Estrogen is also critical to growth and maintenance in preparation for implantation of an embryo and disruption of ERα leads to the inability for implantation to occur [*34*]. These results highlight the vital nature of functional estrogen-ERα signaling to female reproductive

In addition to the aforementioned reproductive effects observed in females, it turns out that ERα is also critical to male fertility [*36*]. This was a surprising observation as it was previously thought that ERα only played a role in female fertility. Testicular tissue of the knockout mice undergoes atrophy over time and sperm is lost due to fluid accumulation within the seminiferous tubules [*32*]. ERα is required by the somatic cells within the testes for the proper maturation of sperm and is not an inherent property of sperm lacking ERα [*37, 38*].

1.2.2 Cardiovascular Health

The incidence of cardiovascular events in premenopausal women is low but rises to levels seen in men in the postmenopausal setting. This suggests a potential role of ERα in carioprotection of women. Clinical data supports this picture as reduced levels of ERα is associated with the development of coronary artery disease in female patients [*39*]. Furthermore, methylation of the ERα gene is observed in coronary atherosclerotic plaques compared to controls [*40*]. Estrogen is also associated with favorable changes in lipid profiles. Estrogen has been associated with increased high-density lipoprotein (HDL) and lower low-density lipoprotein (LDL) levels [*41, 42*]. Thus, cardioprotective effects of estrogen are likely to be largely a product of the circulating lipid profiles established by estrogen and ERα.

1.2.3 Osteoporosis

Osteoporosis is the condition of a decline in bone strength or mechanical integrity marked by a decline in bone mineral density. It has long been viewed as a condition predominantly of the postmenopausal population, suggesting a plausible role of estrogen in the maintenance of bone mineral density in premenopausal women. The Women's Health Initiative demonstrated hormone replacement therapy was clinically efficacious for the prevention of osteoporosis in postmenopausal women [*43*]. Likely molecular mechanisms behind these observations have been established. Estrogen prevents bone resorption through stimulation of bone-forming osteoblast activity while reducing the bone-resorbing activity of osteoclasts [*44-47*]. This is accomplished through ERα-dependent expression of osteoprotegrin by osteoblasts which is a soluble scavenger receptor for the molecule receptor activator of nuclear factor kappa-B ligand (RANKL)

which is key to osteoclast development and activity [*47*]. Thus, osteoprotegrin effectively lowers the amount of the RANKL available to osteoclasts which is critical to inducing their activity. In addition to its effect on bone in females, estrogen has been demonstrated to play a direct role in bone maintenance in males with the description of an estrogen-insensitive osteoporotic male patient [*31, 48*]. Today the selective estrogen receptor raloxifene (which acts as an ERα agonist in bone) is used in the prevention of postmenopausal osteopenia in women [*30*].

1.2.4 Breast Cancer

Breast cancer is the leading cause of cancer (excluding skin malignancies) and second leading cause of cancer associated death in females [*49*]. Perhaps the most familiar risk factor for development of breast cancer is prior familial history of the disease, and in particular the heritable BRCA1/2 mutations. Inheriting such a mutation is associated with upwards of an 80% lifetime risk of developing breast cancer. Despite such a high likelihood of developing cancer in patients possessing BRCA mutations, such heritable forms of breast cancer only represent five to ten percent of total breast cancer cases. Clearly risk factors other than heritable factors contribute to the development of the majority cases.

Breast cancer is one of the hormone-dependent cancers. Epidemiological evidence demonstrates a linear increase in age-related incidence with a decrease in the slop of the curve at the time of menopause. In addition, three factors which have a dramatic impact on breast cancer risk are age of menarche, nulliparity and age at menopause [*50*]. Obesity has also been demonstrated to increase the risk of breast cancer, likely through additional estrogen synthesis by adipose tissue [*51*]. Furthermore,

prolonged hormone replacement therapy in postmenopausal women has also been associated with an increased breast cancer risk [*52*]. Taken together, such evidence supports the notion of cumulative lifetime estrogen exposure as a factor contributing to breast carcinogenesis.

Two prevailing hypotheses exist attempting to explain the mechanism of estrogen tumorigenesis [*34*]. The first suggests that increased estrogen exposure leads to an increased number of cell divisions with each division carrying a certain risk of mutation. Over time, such mutations become more likely and accumulate. These genetic lesions then lead to tumor formation. The second employs genotoxic byproducts of oxidative estrogen metabolism via the catechol pathway. This pathway utilizes the cytochrome P450 family of enzymes and one such enzyme, cytochrome P450 1B1 is constitutively expressed in breast [*50*]. Cytochrome P4501 1B1 hydroxylates estradiol forming 4-hydroxyestradiol which can be further converted into estradiol-3,4-quinone by peroxidases [*53, 54*]. Estradiol-3,4-quinone can then react with adenine and guanine bases in DNA forming unstable adducts leading to depurination [*50, 53*]. Through both mechanisms, in addition to those yet to be discovered, estrogen proves to be a capable tumor inciting factor.

Initial clinical suggestion for a role of estrogen in breast carcinogenecis came over a century ago with a report from British surgeon Gorge Beatson. In 1896 Beatson demonstrated estrogen-withdrawal through oophorectomy in young women with advanced breast cancer resulted in regression of tumor size and improved survival [*55*]. The strongest evidence linking estrogen to breast cancer and its outcomes comes from decades of clinical experience with the SERM tamoxifen and improvement of patient

survival as well as breast cancer prevention. Clinical trials consistently demonstrate a survival advantage in tamoxifen treated patients irrespective of invasive status in early stage breast cancer as well as recurrence prevention in the contralateral breast [*56*]. One meta-analysis of multiple clinical trials found tamoxifen reduced breast cancer risk by 38 percent in high risk women [*57*]. This risk reduction applies only to ER-positive tumors further indicating tamoxifen clinical effects are limited to ER-dependent pathways. The success of tamoxifen has spawned the development of additional SERMs such as raloxifene as well as the estrogen synthesis blocking aromatase inhibitors (AIs) which are used in the clinic today. The clinical utility of the multiple pharmaceuticals targeting estrogen-dependent signaling via differing means (ligand competition in the case of SERMs and ligand depletion in the case of AIs) exemplifies the central role of estrogen in breast carcinogenesis.

1.3 ERα shares a modular domain organization common to all nuclear receptors

ERα is a member of a family of ligand-modulated transcription factors that have come to be known as nuclear receptors (NRs) [*3-6*]. In humans, there are currently 48 known members of the NR family and all share a core modular architecture comprised of a central DNA-binding domain flanked between an N-terminual trans-activation (TA) domain and a C-terminal ligand-binding (LB) domain (Figure 1-2) [*58-60*] . Intuitively, the DB domain is the region of the molecule which confers DNA-binding capacity to the nuclear receptors. The DB domains of NRs are comprised of conserved tandem zinc-fingers and bind to DNA motifs known as hormone response elements [*61*]. The ligand binding function of the LB domain is specific for its cognate ligand which ensures specificity of the desired physiological response [*61*]. In addition to ligand binding, the

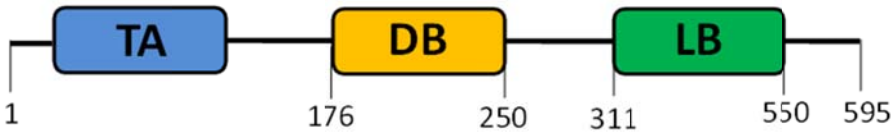

Figure 1-2: Schematic of the domain organization of ERα which typifies nuclear receptors. A central DNA-binding domain (DB) is flanked by an N-terminal Transactivation (TA) domain and a C-terminal Ligand Binding (LB) domain. The numbers represent the domain boundaries used in this work.

LB domain additionally serves as a platform for the recruitment of a multitude of cellular proteins, such as transcription factors, co-activators and co-repressors, to the site of DNA transcription and thereby allowing nuclear receptors to exert their action at genomic level in a concerted fashion [*62, 63*]. While the trans-activation function of the LB domain is ligand-dependent, the TA domain operates in an autonomous manner and it is believed to be responsive to growth factors acting through MAPK signaling and may further synergize the action of various co-activators and co-repressors recruited by the LB domain at the site of DNA transcription [*28, 64*]. In this manner, nuclear receptors orchestrate a diverse array of cellular functions, and the success of coupling a DNA-binding module to a ligand-binding one has allowed nuclear receptors to become the primary players controlling organ-specific physiology [*61*].

1.4 The ERα-ERE interaction is governed by its DB domain

Vital to the physiological actions of ERα is its ability to function as a transcription factor. The defining feature of transcription factors is their capacity to directly bind DNA through protein-DNA interactions. The DB domain of ERα binds DNA as a homodimer to the AGGTCAnnnTGACCT consensus motif, termed estrogen response element (ERE), located within the promoters of target genes [*65*]. The crystal structure revealed that DNA-binding is accomplished through a pair of tandem C4-type Zinc fingers, with each finger containing a Zn^{2+} ion coordinated in a tetrahedral arrangement by four highly

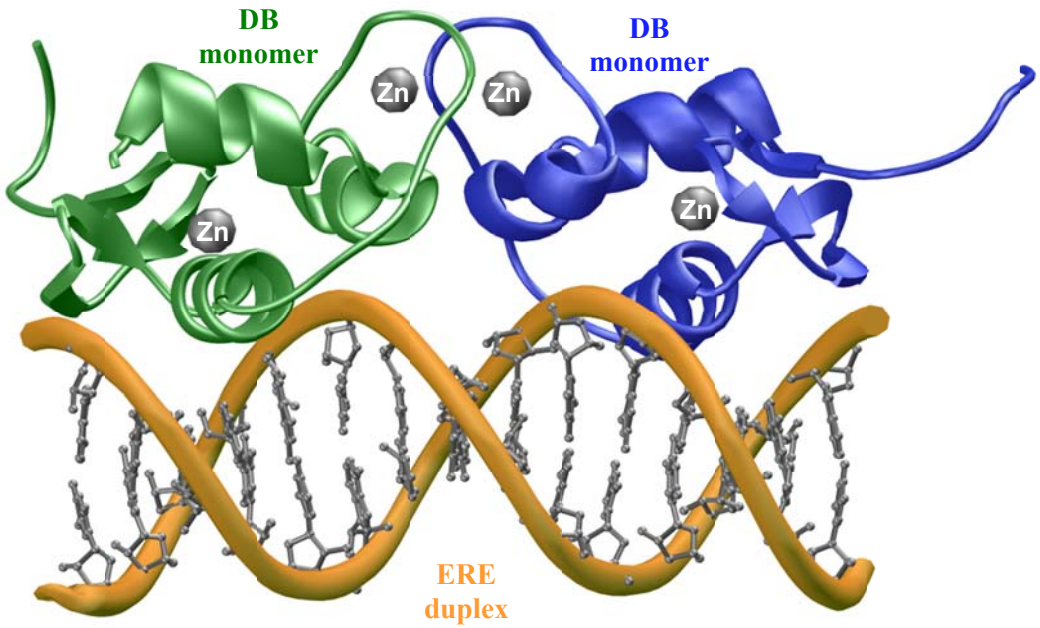

Figure 1-3: 3D crystal structure of the DB domain of ERα in complex with ERE. DB as a homodimer in complex with ERE duplex containing the AGGTCAcagTGACCT consensus sequence as determined by Rhodes and co-workers [*66*]. One monomer of the DB domain is shown in green and the other in blue. The Zn^{2+} divalent ions are depicted as gray spheres. The DNA backbone is shown in yellow and the bases are colored gray for clarity.

conserved cysteine residues [*66, 67*]. The first Zinc finger (ZF-I) within each monomer of DB domain recognizes the hexanucleotide sequence 5'-AGGTCA-3' within the major groove at each end of the ERE duplex, whilst the second Zinc finger (ZF-II) is responsible for the homodimerization of DB domain (Figure1-3). DNA sequence recognition has been demonstrated to be exclusively a property of the DB domain through construction of a protein chimera in which the DB domain of glucocorticoid receptor was swapped into ERα resulting in a receptor which responded to estradiol but bound only to a glucocorticoid response element [*68*]. Furthermore, it was demonstrated that this DNA sequence specificity of the DB domains is largely due to three amino acids within ZF-I of the DB domains [*69*].

From the DNA-side, it has remained unclear which nucleotides within the ERE are the most critical or make the strongest contact with the DB domain in large part due

to differences in laboratory technique in the studies [*70, 71*]. Several lines of evidence suggest that genetic variations within the cognate response elements play a key role in modulating the affinity and specificity of binding of androgen, glucocorticoid and progesterone nuclear receptors [*72-75*]. However, determination of the impact of nucleotide variation on ERα binding affinity has remained elusive, a surprising fact given the consensus motif was determined in 1989 [*65*]. This is an important consideration as most estrogen-regulated genes contain imperfect ERE sequences [*65*].

1.5 Zn-coordination impacts structure and function of the DB domain

The DB domain of ERα contains a pair of tandem C4-type zinc fingers. This tandem configuration of C4 zinc fingers is unique to nuclear receptors and is sometimes referred to as the hormone receptor DNA binding module [*76*]. The divalent zinc provides a structural scaffold of sorts allowing for proper configuration within the domain, thus imparting DNA binding capability. Zinc is coordinated in a tetrahedral geometry by the cysteine sulfurs. It would stand to reason that any metal with the capacity to adopt such tetrahedral coordination geometry would also be able to accommodate the DB domain. Zinc, however, is ideally suited to play such a role in biological systems as it is abundant and relatively nontoxic. In solution zinc exists as a divalent cation (Zn^{2+}) and is a closed-shell ion (filled d-shell, d^{10}) [*76, 77*]. Thus, Zn^{2+} can freely accommodate multiple coordination geometries without ligand-field stabilization energy (LFSE) penalty. Zn^{2+} is also regarded as a mild acid in terms of acid-base hardness theory and can readily interact with a wide variety of coordinating partners such as cysteine sulfur, histidine nitrogen, and oxygen from aspartate and glutamate [*76*]. Other transition metals do not possess such a favorable profile of

chemical properties. This is likely the physical basis of biology's preference for Zn^{2+}over other metals in many settings.

Naturally, given the toxicity of many heavy metals one might wonder to what extent these metals might be able to interfere with zinc in a protein thereby altering or disrupting its function. Such metal replacement studies have been conducted in the case of the DB domain of ERα [*78-80*]. Together these studies reveal that cadmium, cobalt, and iron are able to replace zinc while retaining DNA binding capacity. In contrast, nickel and copper were able to replace zinc within the fingers but render the DB domain incompetent to bind DNA. These studies provide insight into the potential for various metals to compete with zinc to occupy the fingers within the DB domain and indicate a potential mechanism for toxicity of these metals. One limitation of these studies is that binding constants were obtained under non-solution and non-equilibrium conditions using semiquntitative techniques. Also, the studies do not provide any structural data or describe any differences in the physical properties of the reconstituted domains which is critical to rationalization of the observed functional results. Thus, the structural and physical consequences of metal substitutions have yet to be fully explored and may shed light onto many important aspects of metal-protein biology.

1.6 Significance of these studies

ERα is a nuclear receptor which plays an essential role in a wide array of biological and pathological processes. The wealth of knowledge amassed on ERα ranges from clinical data and associated phenotypes to intricacies of its signaling pathways all the way down to its aesthetically pleasing 3D atomic structure. However, essentially no information concerning the ubiquitous and invisible thermodynamic forces which govern

the fundamental interactions central to the functioning of this critical nuclear receptor has been elucidated. Unraveling the molecular mechanisms surrounding ERα-DNA interactions will uncover new insights into the inner workings of this key molecular complex. By utilizing biophysical techniques, this thesis hopes to further our understanding of the biophysical basis of ERα-DNA interactions. Such knowledge may harbor the potential for the development of novel therapies possessing lower toxicity coupled with higher efficacy for the treatment of pathological conditions such as breast cancer. In an attempt to enhance our understanding of ERα-DNA interactions, this thesis will elucidate the physicochemical underpinnings of this key protein-DNA complex.

Chapter 2: Materials and methods

2.1 Molecular cloning

The DB domain of human ERα (residues 176-250) was cloned into both the pET101 and pET102 bacterial expression vectors using directional TOPO technology according to manufacturer specifications (Invitrogen). The pET102 vector contains an N-terminal thioredoxin (Trx)-tag as well as a C-terminal polyhistidine (His)-tag. The pET101 vector possesses only a C-terminal His-tag. The Trx-tag was included to maximize protein and help keep the protein soluble as necessary, while the His-tag aids in protein purification using immobilized metal affinity chromatography (IMAC), specifically Ni-NTA affinity chromatography. Thrombin cleavage sequences (LVPRGS) were engineered into both the N-and C-termini (pET102) or the C-terminus (pET101) of the protein insert to permit the removal of the tags subsequent to protein purification as needed.

2.2 Protein expression and purification

All ERα DB domain constructs were expressed in *Escherichia coli* strain BL21* (Invitrogen). The BL21* strain utilizes the DE3 lysogen system to express recombinant protein. In addition, BL21* has a truncated RNase E which enhances mRNA lifetime and thereby maximizes protein expression. Transformed colonies were cultured in LB media and grown at 20°C to an optical density of 0.5 at 600nm. Expression was induced with 0.5 mM isopropyl β-D-1-thioglactopyranoside (IPTG) overnight (18 hours total). Cells were then harvested and resuspended in Lysis Buffer (50mM Tris, 500mM NaCl, 2mM 2-mercaptoethanol, 1% Triton X-100 at pH 8.0). Cells were homogenized utilizing a BeadBeater (Biospec). Cellular debris was separated from lysate via high-speed

centrifugation. The lysate was gravimetrically applied to a Ni-NTA column thrice and subsequently rinsed with a Wash Buffer (50mM Tris, 500mM NaCl, 20mM imidazole, 2mM 2-mercaptoethanol, 1% Triton X-100 at pH 8.0) to remove non-specific binding of bacterial proteins to the column. The recombinant proteins were then eluted with an Elution Buffer (50mM Tris, 500mM NaCl, 200mM imidazole, 2mM 2-mercaptoethanol, 1% Triton X-100 at pH 8.0). The elutant was then dialyzed into an appropriate physiological buffer. Dialyzed protein was further purified using a HiLoad 26/60 Superdex 200 preparatory grade size exclusion chromatography (SEC) column coupled to a GE Akta FPLC system. Protein concentrations were determined by both the fluorescence-based Quant-It assay (Invitrogen) and spectrophotometrically using extinction coefficients of 14,940 $M^{-1}cm^{-1}$ for DB-His (pET101) and 29,045 $M^{-1}cm^{-1}$ for Trx-DB-His (pET102) using the online software ProtParam at Expasy Server [*81*].

2.3 SDS-PAGE analysis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is a ubiquitous technique used to separate proteins according to size [*82*]. SDS is an anionic surfactant that serves as a dissociating and denaturing agent. When protein samples are heated to 100°C, SDS evenly coats the polypeptide backbone thereby rendering any intrinsic charge negligible. This results in proteins with uniform negative charge density. Thus, when an electric field is applied across the gel, SDS-coated proteins are pulled with the same force per unit mass toward the apparatus cathode, and separation occurs according to size as larger proteins migrate slower due to increased resistance provided by the polyacrylamide matrix. The polyacrylamide gel simply serves as a sieving matrix to the protein samples. The observed electrophoretic mobility is a linear function of the

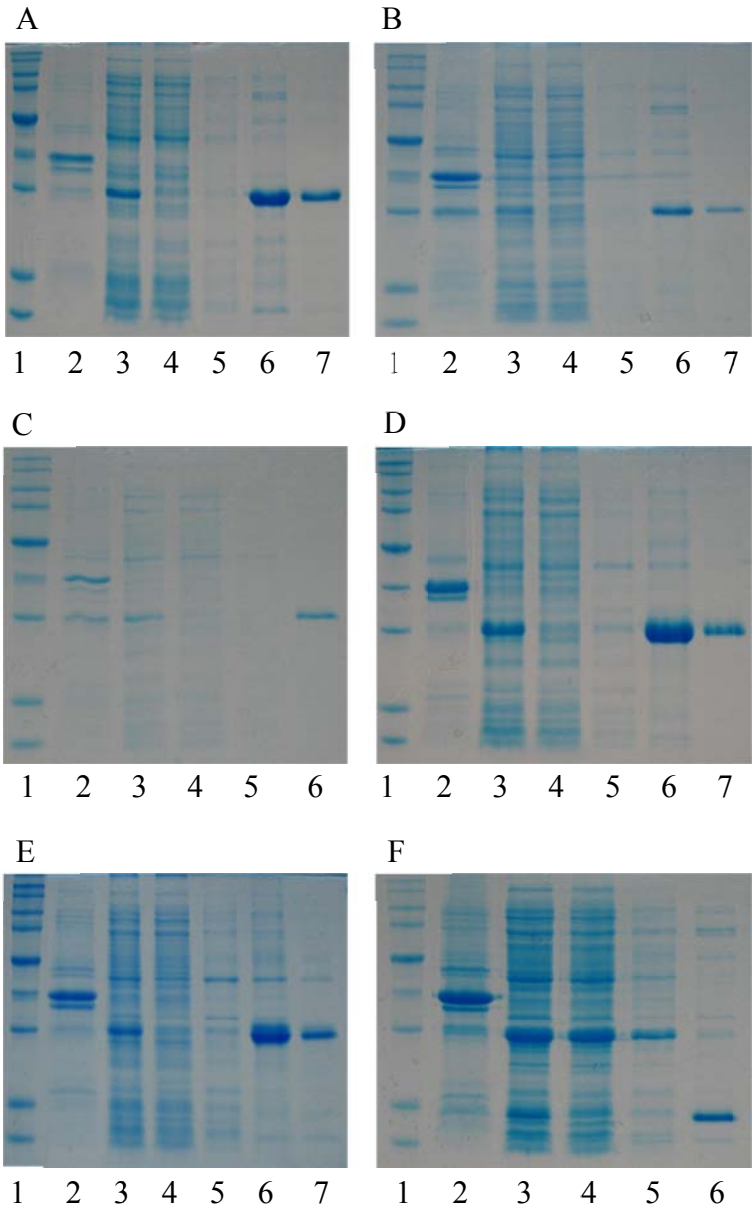

Figure 2-1. SDS-PAGE analysis of the Ni-NTA purification of recombinant proteins. Recombinant human ERα DB (A), ERα DB D190A (B), ERα DB H196A (C), ERα DB E203A (D), ERα DB H196A_E203A (E) utilizing the pET102 vector, while recombinant human ERα DB utilizing pET101 (F) were expressed in E. Coli strain BL21* at 20ºC to an optical density of 0.5 at 600nm and induced overnight with 0.5 mM IPTG. 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 cellular inclusion body was solubilized in 10% SDS and loaded onto a 12% SDS-PAGE gel (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). The column was then washed with 20 mM imidazole to remove non-specific binding of bacterial proteins (lane 5). The recombinant proteins were eluted with 200mM imidazole (lane 6) and dialyzed into appropriate physiological buffers (lane 7). Protein markers (Promega) were added for reference (lane 1).

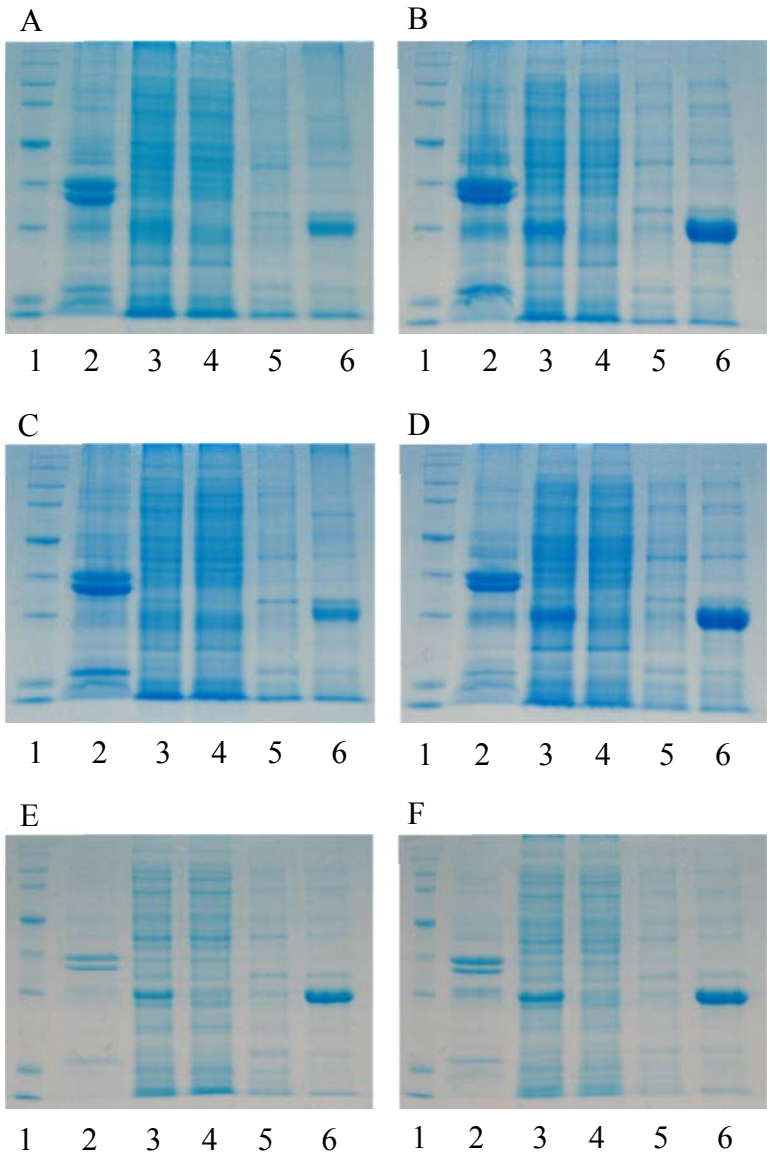

Figure 2-2. SDS-PAGE analysis of the Ni-NTA purification of recombinant proteins. Recombinant human ERα DB S193A (A), ERα DB Y197A (B), ERα DB S201A (C), ERα DB K206A (D), ERα DB K210A (E) ERα DB R211A (F) were expressed in E. Coli strain BL21* at 20ºC to an optical density of 0.5 at 600nm and induced overnight with 0.5 mM IPTG. 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 cellular inclusion body was solubilized in 10% SDS and loaded onto a 12% SDS-PAGE gel (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). The column was then washed with 20 mM imidazole to remove non-specific binding of bacterial proteins (lane 5). The recombinant proteins were eluted with 200mM imidazole (lane 6). Protein markers (Promega) were added for reference (lane 1).

logarithm of molecular weight. Therefore, sample molecular weight can be estimated using a standardized set of protein markers of known molecular weights.

SDS-PAGE analysis of recombinant protein samples (Figure 2-1) were carried out by loading sample onto a 12% (w/v) SDS-PAGE gel run at 150V for 60min using a VWR AccuPower power supply and a Bio-Rad Protean Chamber. Protein bands were visualized by first staining with a 0.1% (w/v) coomassie-blue solution containing 40% (v/v) methanol and 10% (v/v) methanol followed by destaining with a destaining solution comprised of 10% (v/v) methanol and 10% (v/v) acetic acid. Gel images were obtained using a UVP MultiDoc-It Gel Imaging System.

2.4 Site-directed mutagenesis

pET102 bacterial expression vector containing wild-type DB domain of ERα were subjected to Stratagene Quikchange Lightning site-directed mutagenesis to produce mutants of interest. Mutagenesis was conducted according to manufacturer protocol. Recombinant mutant constructs were expressed, purified and characterized as described above.

2.5 DNA synthesis

21-mer oligonucleotides containing the ERE consensus motif (AGGTCAnnnTGACCT) or a single nucleotide variant thereof (complete sequences can be found in Chapter 4) were obtained commercially from Sigma Genosys. The complete nucleotide sequence of both the sense and antisense constituting the consensus ERE sequence is presented below:

5’-cccAGGTCAcagTGACCTgcg-3’

3’-gggTCCAGTgtcACTGGAcgc-5’

G- and C-rich were included (as shown above) in all ERE motifs to prevent end fraying. Additional procedure details can be found in Chapters 3.3.3 and 4.3.2.

2.6 Oligonucleotide annealing

The DNA oligonucleotides of ERE and single nucleotide variants thereof were reconstituted in water. Concentrations were determined spectrophotometrically on the basis of extinction coefficients derived from their nucleotide sequences using the online software OligoAnalyzer 3.0 (Integrated DNA Technologies) based on the nearest-neighbor model [*83*]. In order to obtain double-stranded DNA (dsDNA) duplex, equimolar concentrations of sense and antisense oligonucleotides were mixed at heated to 95°C for 10 min followed by slow cooling to room temperature in a Techne TC-312 thermocycler.

2.7 SEC analysis

Size-exclusion chromatography (SEC) was performed on a Hiload Superdex200 column coupled to a GE Akta FPLC system equipped with UNICORN software capable of automated data collection and operation. SEC is a widely used technique to separate macromolecules based upon effective hydrodynamic volume or size. Hydrodynamic volume or particle size is correlated to molecular weight assuming conditions of a compact spherical and globular fold are met. Briefly, SEC works by applying a sample that is a mixture of proteins onto a column containing a porous gelationous stationary phase running at relatively low pressure (~0.2 MPa). The gelatinous material is comprised of porous "beads" which together establish a pore network. Molecules with a diameter smaller than the pore are permitted to enter the beads thereby effectively increasing the volume through which such molecules must travel to elute (such molecules

must pass both interparticle and intraparticle volumes). On the other hand, molecules larger than the pore are excluded from the pore network and their flow is relatively unimpeded only passing through the interparticle volume and elute earlier. Particles of differing hydrodynamic volumes therefore elute through the stationary phase at different rates. This permits the separation of molecules within a sample based simply on size. On a practical level, because a given protein molecule does not have a single well-defined size, but rather is comprised of an ensemble of sizes, coupled with the fact that the pore sizes are not ideal nor uniform, leads to a normal distribution of elution volume for a given protein. A major benefit of the technique is a sample can be purified and analyzed under native conditions. However, a drawback to SEC is that because filtration rate is simply determined by size, global molecular shape can impact observed elution volumes and hence apparent molecular weights. That is, it is not an absolute method for molecular weight determination.

After purification via Ni-NTA affinity chromatography recombinant proteins were extensively dialyzed in an appropriate buffer. When necessary, depending on protein yield, the recombinant protein was concentrated to an appropriate concentration using Amicon Ultra-15 centrifugal filter units, prior to application on the Superdex200 column pre-equilibrated with identical buffer at 10°C. The elution profile of the recombinant protein was recorded using the UV monitor at 280nm and automatically plotted as a function of elution volume in the UNICORN software. The apparent molecular mass could then be determined by substitution of its observed elution volume maximum into the linear function fit to the logarithm of the molecular mass of known protein standards versus their elution volumes under the same conditions.

2.8 ALS measurements

Analytical light scattering (ALS) is comprised of two types of light scattering modalities – static light scattering (SLS) and dynamic light scattering (DLS). These techniques are used to analyze size and molecular weight (SLS) as well as hydrodynamic properties (DLS) of proteins. When a light source is passed into a medium any non-uniformity (such as a molecule) that is small relative to the wavelength of incident light will scatter light in all directions. This type of scattering is known as Rayleigh scattering. SLS is a technique which measures angular dependence in scattering intensity over a range of concentrations to measure both radius of gyration R_g and molecular weight. SLS takes advantage of anisotropism where scattering by a sample displays an angular dependence in scattering intensity. SLS measures scattering intensity at multiple detection angles and the variation in intensity with respect to angle permits determination of the radius of gyration R_g. Intuitively scattering intensity is directly proportional to molecular weight and concentration, if one knows sample concentration, molecular weight can be calculated. It turns out that double extrapolation to zero angle and zero concentration is what yields molecular weight according to a reduced Zimm equation [*84, 85*].

Dynamic light scattering is a technique used to determine a molecule's translation diffusion coefficient (D_t), which is inversely proportional to hydrodynamic radius (R_h), by measuring time-dependent fluctuation in scattering intensity. These fluctuations arise from dynamic changes in inter-scatterer distances undergoing Brownian motion. Given that the light source is monochromatic and coherent (a laser), measured scattered light will undergo constructive and destructive interference resulting in fluctuation of

measured scattering intensity. Fluctuation rate or periodicity is related to the size of the particles in solution. Larger particles move slower in a solution relative to smaller particles. Thus larger particles will display a slower periodicity of fluctuation intensity corresponding to a smaller D_t and a larger R_h. R_h can be used to indicate the degree to which a protein is folded relative to a plot of the logarithm of R_h versus the logarithm of molecular weight of known protein standards. Values of R_h for a corresponding molecular weight which lay off of the standard curve indicate the protein is likely not folded in an entirely compact manner.

Analytical light scattering (ALS) experiments were conducted on a Wyatt miniDAWN TREOS triple-angle static light scattering detector and Wyatt QELS dynamic light scattering detector coupled in-line with a Wyatt Optilab rEX differential refractive index detector and interfaced to a Hiload Superdex 200 size-exclusion chromatography column under the control of a GE Akta FPLC system within a chromatography refrigerator at 10°C. The DB domain of ERα pre-treated with EDTA to strip divalent zinc ions and upon reconstitution with divalent ions of various metals was loaded onto the column at a flow rate of 1ml/min and the data were automatically acquired using the ASTRA software. All protein samples were prepared in 50mM Sodium phosphate buffer containing 5mM β-mercaptoethanol at pH 7.0 and the starting concentrations injected onto the column were between 20-50μM. The angular- and concentration-dependence of static light scattering (SLS) intensity of each protein species resolved in the flow mode was measured by the Wyatt miniDAWN TREOS detector. The SLS data were analyzed according to the built-in Zimm equation in ASTRA software [*84, 85*]. The time- and concentration-dependence of dynamic light scattering (DLS) intensity

fluctuation of each protein species resolved in the flow mode was measured by the Wyatt QELS detector positioned at 90° with respect to the incident laser beam. The DLS data were iteratively fit using non-linear least squares regression analysis using the built-in equation in ASTRA software [*86-88*]. More detailed and specific procedures can be found in Chapter 5.3.4.

2.9 SSA measurements

Steady-state absorbance (SSA) spectroscopy was employed as a technique to analyze how metal substitution within the zinc fingers of ERα DB domain influence absorbance spectral features in the UV region. It has been previously reported that spectral features of metal-bound protein remain largely similar but the spectral intensity of the 280 nm absorption band undergo reduction upon specific metal-protein interaction [*89, 90*]. This phenomenon is likely due to change in solvation of Trp/Tyr/Phe residues due to protein-metal interactions.

SSA spectra were collected on a Jasco V-630 spectrophotometer using a quartz cuvette with a 10-mm pathlength at 25 °C. Samples were prepared in 50mM Sodium phosphate buffer at pH 7.0. All data were recorded in the 200-350nm wavelength range using a 1.5-nm slit bandwidth. Data were normalized against reference spectra to remove the contribution of buffer. The reference spectra were obtained on a 50mM Sodium phosphate buffer at pH 7.0. More detailed procedures can be found in Chapter 5.3.6

2.10 SSF measurements

Steady-state fluorescence (SSF) spectroscopy was utilized to determine how metal substitution within the zinc fingers of ERα DB domain impact the fluorescence characteristics of metal-reconstituted protein. Changes in tryptophan fluorescence are

reflective of an environmental shift surrounding tryptophan. Thus, observed changes in fluorescence spectral features of metal-reconstituted protein are also are indicative of differing tryptophan environment.

SSF spectra were collected on a Jasco FP-6300 spectrofluorometer using a quartz cuvette with a 10-mm pathlength at 25 °C. Samples were prepared in 50mM Sodium phosphate buffer at pH 7.0. Excitation wavelength was 295nm and emission was acquired from 310nm to 500nm. All data were recorded using a 2.5-nm bandwidth for both excitation and emission. Data were normalized against reference spectra to remove the contribution of buffer. The reference spectra were obtained in a similar manner on a 50mM Sodium phosphate buffer at pH 7.0. More detailed procedures can be found in Chapter 5.3.6

2.11 ITC measurements

Isothermal titration calorimetry (ITC) is the only method for directly studying protein-ligand thermodynamics in solution. ITC provides molecular insight into the ubiquitous yet invisible forces governing protein-ligand complexes [*91, 92*]. ITC directly measures heat evolution associated with a chemical process, whether a reaction or binding, which, under isobaric conditions, is equal to the enthalpy change of the process [*83*]. Therefore ITC, in effect, measures a change in enthalpy (ΔH). Due to the nature of a titration of a protein with a ligand by serial injections, heat response associated with a given injection will be a function of the number of available binding sites. As more binding sites become occupied with ligand, the heat associated with the injection decreases until conditions of saturation are met. Analysis of the integrated isotherm yields a plot which permits calculation of affinity (K_d), Gibbs free energy change (ΔG),

change in entropy (ΔS) and stoichiometry (n). ITC can also determine whether a reaction between two components is under enthalpic or entropic control. Binding events under enthalpic control are exothermic and the associated heat release is a result of favorable contacts formed between protein and ligand due to, for example, hydrogen bonding and electrostatic interactions. In contrast, entropy-driven processes are typically endothermic, or weakly exothermic, owing to a large positive change in entropy of the system reflective of an order-disorder transition arising from conformational, translational and solvent entropic changes upon ligand binding. In addition, ITC can be further utilized to glean mechanistic details surrounding macromolecular assemblies through measurement of heat capacity change (ΔC_P), change in solvent accessible surface area (ΔSASA) [*93-97*], counter-ion uptake [*98*] and protonation/deprotonation events [*99*]. Given its range of capabilities, ITC indeed proves to be an indispensable tool in biophysical characterization macromolecular interactions.

ITC measurements were performed on a Microcal VP-ITC instrument and data were acquired and processed using fully automated features in Microcal ORIGIN. All measurements were repeated at least three times. Briefly, protein and DNA samples were prepared in 50mM Sodium phosphate buffer containing 5mM β-mercaptoethanol at pH 7.0 and de-gassed using the ThermoVac accessory for 5min. The experiments were initiated by injecting 25 x 10μl injections from the syringe into the calorimetric cell, at a fixed temperature. 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 uptake per injection as a function of

concentration. The heats of mixing and dilution were subtracted from the heat of binding per injection. A more detailed and specific procedure can be found in Chapters 3.3.4, 4.3.3 and 5.3.3.

2.12 CD analysis

Circular dichroism (CD) analysis was used to analyze the opticospectroscopic properties of both recombinant protein and oligonucleotide samples. CD is a powerful technique that permits determination of secondary structural features of both protein and DNA as well as a more general higher-ordered tertiary properties as well. It is particularly beneficial in the context of relative comparison of changes in spectral features when introducing perturbations to a system.

CD measurements were conducted on a Bio-Logic MOS450/SFM400 spectropolarimeter thermostatically controlled with a water bath at 25°C. Data were acquired using the BIOKINE software. Samples were prepared in 50mM Sodium phosphate buffer at pH 7.0. For protein samples, far-UV measurements in the 190-250nm wavelength range, experiments were conducted in a quartz cuvette with a 2-mm pathlength, while near-UV measurements in the 250-350nm wavelength range, experiments were conducted in a quartz cuvette with a 10-mm pathlength. For DNA oligonucleotide experiments both far-UV and near-UV measurement were obtained using a 2-mm pathlength quartz cuvette. All spectra were recorded with a slit bandwidth of 2nm at a scan rate of 3nm/min. Data were normalized against reference spectra to remove the contribution of buffer. All data were processed and analyzed using the Microcal ORIGIN software. More detailed and specific procedures can be found in Chapters 4.3.4 and 5.3.5.

2.13 Macromolecular modeling

Molecular modeling (MM) was employed to generate 3D atomic models of the DB domain of ERα in complex with dsDNA oligos containing the consensus ERE motif or a variant motif thereof using the MODELLER software based on homology modeling [*100*]. Briefly, 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 are expected to adopt 3D folds similar to the template structure with the exception sidechain conformation of specific amino acids due to the introduction of defined hydrogen bonding, the rearrangement of domains and DNA spatially to one-another or the modeling of loops not rendered in template structures. Herein, the crystal structure of the DB domain of ERα in complex with a dsDNA oligo containing the ERE motif but with varying flanking sequences was used as a template (PDB entry 1HCQ). Additionally, for atomic models of DB domain in complex with the single nucleotide variant motifs, hydrogen bonding distance restraints were added between appropriate pairs of atoms to allow base pairing between substituted nucleotides within each half-site. For each motif, a total of 100 atomic models were calculated and the structures with the lowest energy, as judged by the MODELLER Objective Function, were selected for further analysis. The atomic models were rendered using RIBBONS [*101*]. Electrostatic surface potentials were generated using MOLMOL [*102*]. Specific modifications made for each model can be found in Chapters 3.3.5 and 4.3.5.

Chapter 3: Binding of the DB Domain of ERα Nuclear Receptor to DNA Is Coupled to Proton Uptake

3.1 Summary

Nuclear receptors act as ligand-modulated transcription factors and orchestrate a plethora of cellular functions central to health and disease. Although studied for more than half a century, many mysteries surrounding the mechanism of action of nuclear receptors remain unresolved. Herein, using isothermal titration calorimetry (ITC) in conjunction with macromolecular modeling (MM), we provide evidence that the binding of ERα nuclear receptor to its DNA response element is coupled to proton uptake by two ionizable residues, H196 and E203, located at the protein-DNA interface. Alanine substitution of these ionizable residues decouples protonation and hampers the binding of ERα to DNA by nearly an order of magnitude. Remarkably, H196 and E203 are predominantly conserved across ~50 members of the nuclear receptor family, implying that proton-coupled equilibrium may serve as a key regulatory switch for modulating protein-DNA interactions central to nuclear receptor function and regulation. Taken together, our findings unearth an unexpected but a critical step in the molecular action of nuclear receptors and suggest that they may act as sensors of intracellular pH.

3.2 Overview

Estrogen receptor α (ERα) is a member of a family of ligand-modulated transcription factors that have come to be known as nuclear receptors (NRs) [*3-6*]. All members of NR family are evolutionarily related and share a core modular architecture comprised of a central DNA-binding (DB) domain flanked between an N-terminal trans-activation (TA) domain and a C-terminal ligand-binding (LB) domain [*58-60*]. A typical scenario for the activation of nuclear receptors involves the secretion of lipophilic

messengers such as hormones and vitamins by appropriate tissues. Upon their diffusion through the cell membrane, these ligands bind to the LB domain of nuclear receptors and allow their translocation into the nucleus to modulate gene expression [*9-11*]. While the DB domain recognizes specific promoter elements, the LB domain additionally serves as a platform for the recruitment of a multitude of cellular proteins, such as transcription factors, co-activators and co-repressors, to the site of DNA transcription and thereby allowing nuclear receptors to exert their action at genomic level in a concerted fashion [*62, 63*]. The TA domain is believed to be responsive to growth factors acting through the MAPK signaling and may further synergize the action of various co-activators and co-repressors recruited by the LB domain at the site of DNA transcription [*28, 64*]. In this manner, nuclear receptors orchestrate a diverse array of cellular functions from embryonic development to metabolic homeostasis and their malfunction has been widely implicated in disease [*58, 103-107*].

Discovered more than half a century ago, ERα mediates the action of estrogens such as estradiol and its hyperactivation leads to the genesis of large fractions of breast cancer [*8, 108-113*]. The DB domain of ERα binds as a homodimer to the AGGTCAnnnTGACCT consensus motif, termed estrogen response element (ERE), located within the promoters of target genes [*65*]. DNA-binding is accomplished through a pair of tandem C4-type Zinc fingers, with each finger containing a Zn^{2+} ion coordinated in a tetrahedral arrangement by four highly conserved cysteine residues [*66, 67*]. The first Zinc finger (ZF-I) within each monomer of DB domain recognizes the hexanucleotide sequence 5'-AGGTCA-3' within the major groove at each end of the ERE duplex, whilst the second Zinc finger (ZF-II) is responsible for the homodimerization of DB domain

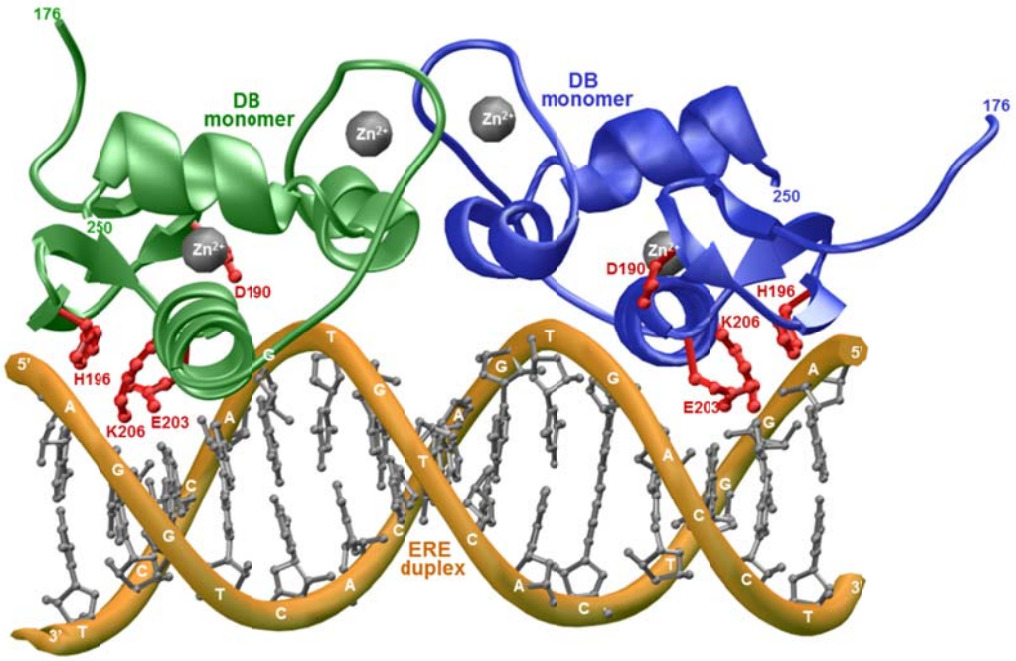

Figure 3-1: 3D atomic model of the DB domain of human ERα in complex with ERE duplex containing the AGGTCAcagTGACCT consensus sequence. Note that the DB domain binds to DNA as a homodimer. One monomer of the DB domain is shown in green and the other in blue. The Zn^{2+} divalent ions are depicted as gray spheres and the sidechain moieties of D190, H196, E203 and K206 within the DB monomers are colored red. The DNA backbone is shown in yellow and the bases are colored gray for clarity. The numerals at the termini of DB monomers indicate the boundaries of DB domain within the amino acid sequence of human ERα.

upon DNA binding. Close scrutiny of 3D structure of the DB domain of ERα in complex with ERE duplex reveals that a triplet of ionizable residues — D190, H196 and E203 — either protrude deep into the comfort of the major groove at the protein-DNA interface or appear to reside within touching distance of DNA backbone [*66*] (Figure 3-1). Given that the placement of these residues in close proximity to the negatively charged phosphate backbone of DNA would be energetically unfavorable due to electrostatic repulsions, we hypothesized that the sidechain moieties of D190, H196 and E203 may become protonated upon the binding of ERα to DNA so as to neutralize the intermolecular repulsions and further enhance the favorable role of electrostatic forces central to driving protein-DNA interactions. This notion is further corroborated by the fact that the

imidazole sidechain of H196 stacks against the highly basic sidechain of K206 — a scenario that could reduce the sidechain pKa of H196 and thereby rendering it more amenable to protonation upon the binding of DB domain to DNA. It is also of worthy note that the acidic sidechains of D190 and E203 are positioned in close proximity to each other within the DB domain. It is thus conceivable that the more acidic sidechain of D190 may be able to increase the pKa value of the sidechain of E203 allowing it to become protonated more easily upon the binding of DB domain to DNA at its own expense. Although both D190 and E203 are located at the protein-DNA interface, protonation of E203 would be more desirable as it directly inserts into the major groove of DNA.

In an effort to test our hypothesis, we have employed here isothermal titration calorimetry (ITC) in conjunction with macromolecular modeling (MM) to analyze the binding of DB domain of ERα to a 21-mer dsDNA oligo containing the ERE motif, hereinafter referred to as ERE duplex. Our data reveal that H196 and E203, but not D190, indeed become protonated upon the binding of ERα to DNA. Furthermore, alanine substitution of these ionizable residues decouples protonation and hampers the binding of ERα to DNA by nearly an order of magnitude. Our study suggests that the proton-coupled equilibrium observed here may be a general feature of the nuclear receptor family.

3.3 Experimental Procedures

3.3.1 Protein Preparation

The DB domain (residues 176-250) of human ERα (Expasy# P03372) 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. 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 DB domain to aid in the removal of tags after purification. The protein was subsequently expressed in Escherichia coli BL21*(DE3) bacterial strain (Invitrogen) and purified on a Ni-NTA affinity column using standard procedures. Briefly, bacterial cells were grown at 20°C in LB media supplemented with 50μM $ZnCl_2$ to an optical density of 0.5 at 600nm prior to induction with 0.5mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The bacterial culture was further grown overnight at 20°C and the cells were subsequently harvested and disrupted using a BeadBeater (Biospec). After separation of cell debris at high-speed centrifugation, the cell lysate was loaded onto a Ni-NTA column and washed extensively with 20mM imidazole to remove non-specific binding of bacterial proteins to the column. The recombinant protein was subsequently eluted with 200mM imidazole and dialyzed against an appropriate buffer to remove excess imidazole. Further treatment on a Hiload Superdex 200 size-exclusion chromatography (SEC) column coupled in-line with GE Akta FPLC system led to purification of recombinant DB domain to apparent homogeneity as judged by SDS-PAGE analysis. The identity of recombinant protein was confirmed by MALDI-TOF mass spectrometry analysis. Final yield was typically between 5-10mg protein of apparent homogeneity per liter of bacterial culture. Treatment with thrombin protease significantly destabilized the recombinant DB domain and it appeared to be proteolytically unstable. For this reason, except for control experiments to ensure that the tags had no effect on the binding of DB domain to DNA, all experiments

reported herein were carried out on recombinant fusion DB domain containing a Trx-tag at the N-terminus and a His-tag at the C-terminus. Protein concentration was determined by the fluorescence-based Quant-It assay (Invitrogen) and spectrophotometrically using an extinction coefficient of 29,045 $M^{-1}cm^{-1}$ calculated for the recombinant fusion DB domain using the online software ProtParam at ExPasy Server [*81*]. Results from both methods were in an excellent agreement.

3.3.2 Site-Directed Mutagenesis

pET102 bacterial expression vector expressing wildtype DB domain of ERα was subjected to QuikChange Lightening kit (Stratagene) to generate single mutants D190A (DB_D190A), S193A (DB_S193A), H196A (DB_H196A), Y197A (DB_Y197A), S201A (DB_S201A), E203A (DB_E203A), K206A (DB_K206A), K210A (DB_K210A), R211A (DB_R211A), and the double mutant H196A/E203A (DB_AA). All mutant DB domains were expressed, purified and characterized as described above. When analyzed by size-exclusion chromatography (SEC) using a Hiload Superdex 200 column, all mutant DB domains exhibited virtually indistinguishable elution volumes to those observed for the wildtype DB domain, implying that the point substitution of specific residues did not lead to protein unfolding and that the mutant DB domains retained the compact globular fold characteristic of wildtype DB domain. These observations were further confirmed by circular dichroism (CD) analysis.

3.3.3 DNA synthesis

21-mer DNA oligos containing the ERE consensus site (AGGTCAnnnTGACCT) were commercially obtained from Sigma Genosys. The complete nucleotide sequences of the sense and antisense oligos constituting the ERE duplex is presented below:

5’-cccAGGTCAcagTGACCTgcg-3’

3’-gggTCCAGTgtcACTGGAcgc-5’

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 [*83*]. To obtain double-stranded DNA (dsDNA) annealed oligos to generate the ERE duplex, equimolar amounts of sense and antisense 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 circular dichroism (CD) analysis.

3.3.4 ITC measurements

Isothermal titration calorimetry (ITC) experiments were performed on a Microcal VP-ITC instrument and data were acquired and processed using fully automized features in Microcal ORIGIN software. Measurements were repeated 3-4 times in Phosphate, Hepes, Tricine or Tris buffer. All buffers were made up to a final concentration of 50mM containing 5mM β-mercaptoethanol at pH 7.0. Additionally, 0-105mM NaCl was added to adjust the ionic strength of all buffers to 110mM. This ionic strength was high enough to prevent non-specific binding of DB domain of ERα to DNA, yet low enough to allow ITC analysis to be conducted with a high signal-to-noise ratio. Various constructs of the DB domain and the ERE duplex were prepared in an appropriate buffer and de-gassed using the ThermoVac accessory for 5min. The experiments were initiated by injecting 25 x 10μl aliquots of 50-100μM of ERE duplex from the syringe into the calorimetric cell

containing 1.8ml of 5-10μM of DB domain solution at 25°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 molar ratio of ERE duplex to dimer-equivalent DB domain. 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 ERE duplex in an identical manner. Control experiments with scrambled dsDNA oligos generated similar thermal power to that obtained for the buffer alone, implying that there was no non-specific binding of DB domain to non-cognate DNA. Experiments on the binding of thrombin-cleaved DB domain to DNA gave similar results to those conducted on recombinant fusion protein, implying that the tags had no effect on DNA-binding. However, due to poor stability and low yield of thrombin-cleaved DB domain and particularly in the case of mutant DB domains, all experiments reported herein were carried out on recombinant fusion DB domain containing a Trx-tag at the N-terminus and a His-tag at the C-terminus. Additionally, titration of a protein construct containing thioredoxin with a C-terminal His-tag (Trx-His) in the calorimetric cell with ERE duplex in the syringe produced no observable signal, implying that the tags do not interact with ERE duplex. In a similar manner, titration of wildtype or mutant DB domains in the calorimetric cell with Trx-His construct in the syringe produced no observable signal, implying that the tags do not interact with any of the wildtype or mutant DB domains. To extract observed affinity (K_{obs}) and observed enthalpy (ΔH_{obs}), 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:

$$q(i) = (nVP\Delta H_{obs}/2)\ \{[1+(L/nP)+(K_{obs}/nP)] - [[1+(L/nP)+(K_{obs}/nP)]^2 - (4L/nP)]^{1/2}\} \quad [3\text{-}1]$$

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 dimer-equivalent concentration of DB domain in the calorimetric cell (μM) and L is the concentration of ERE duplex added (μM). The above equation is derived from the binding of a ligand to a macromolecule using the law of mass action assuming one-site model [*114*]. Observed free energy of binding (ΔG_{obs}) was calculated from the relationship:

$$\Delta G_{obs} = RT\ln K_{obs} \quad [3\text{-}2]$$

where R is the universal molar gas constant (1.99 cal/mol/K) and T is the absolute temperature (298 K). Observed entropic contribution ($T\Delta S_{obs}$) to binding was calculated from the relationship:

$$T\Delta S_{obs} = \Delta H_{obs} - \Delta G_{obs} \quad [3\text{-}3]$$

The net change in the number of protons (Δm) absorbed or released per DB monomer upon binding to DNA and the intrinsic binding enthalpy (ΔH_{int}) due to direct protein-DNA interactions and protonation of ionizable moities were calculated from the slope and y-intercept of ΔH_{obs}-ΔH_{ion} plots by linear fits of data to the equation:

$$\Delta H_{obs} = 2\Delta m\Delta H_{ion} + \Delta H_{int} \quad [3\text{-}4]$$

where ΔH_{obs} is the observed binding enthalpy and ΔH_{ion} is the ionization enthalpy of each buffer. The ΔH_{ion} values of various buffers used were +1.22 kcal/mol (Phosphate), +5.02 kcal/mol (Hepes), +7.64 kcal/mol (Ticine) and +11.35 kcal/mol (Tris) [*115-117*].

3.3.5 Macromolecular modeling

Macromolecular modeling (MM) was employed to generate a 3D atomic model of the DB domain of ERα in complex with the ERE duplex using the MODELLER software based on homology modeling [*100*]. The X-ray structure of DB domain of ERα in complex with a dsDNA oligo containing the ERE motif but with varying flanking sequences was used as a template (with a PDB code of 1HCQ). A total of 100 atomic models were calculated and the structure with the lowest energy, as judged by the MODELLER Objective Function, was selected for further analysis. The atomic model was rendered using RIBBONS [*101*] and the electrostatic surface potentials were generated using MOLMOL [*102*].

3.4 Results and discussion

3.4.1 Binding of the DB domain of ERα to DNA is coupled to proton uptake

To test our hypothesis that the binding of ERα to DNA is coupled to proton uptake, we measured the binding of DB domain of ERα to ERE duplex in buffers of varying ionization enthalpies using ITC. Figure 3-2 shows representative ITC isotherms obtained from such measurements, while detailed thermodynamic parameters are reported in Table 3-1. It should be noted here that a classical test for ligand binding coupled to proton exchange is the dependence of observed enthalpy (ΔH_{obs}) on ionization enthalpy (ΔH_{ion}) of the reaction buffer. Since different buffers are characterized by distinct ionization entahlpies, the observed enthalpy of ligand binding displays sharp

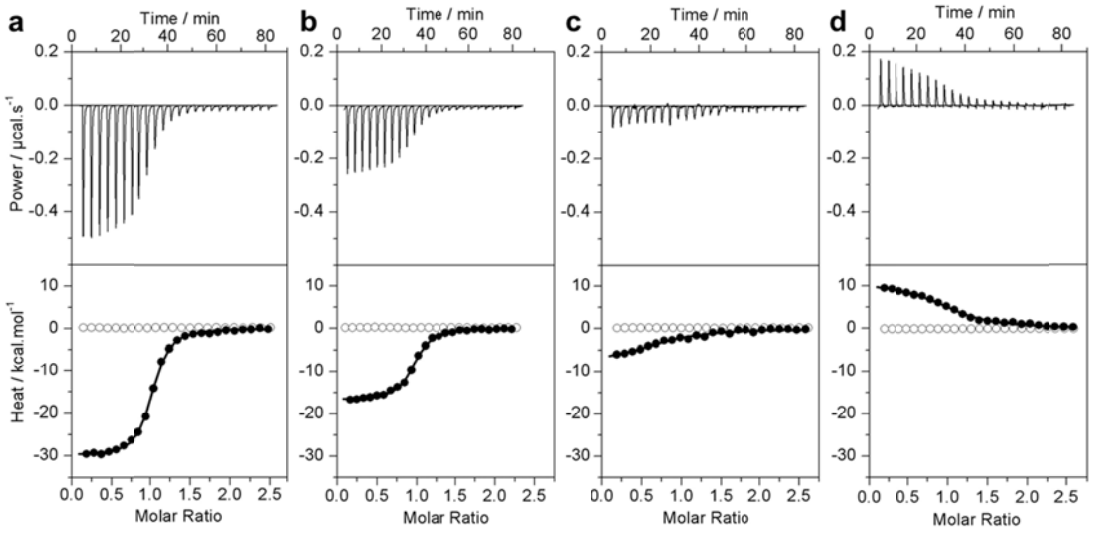

Figure 3-2: Representative ITC isotherms for the binding of ERE duplex to the wildtype DB domain of ERα in Phosphate (a), Hepes (b), Tricine (c) and Tris (d) buffers at pH 7.0 and 25°C. The upper panels show the raw ITC data expressed as change in thermal power with respect to time over the period of titration. In the lower panels, change in molar heat is expressed as a function of molar ratio of ERE duplex to dimer-equivalent DB domain (●). The solid lines in the lower panels show the fit of data to a one-site model, as embodied in Eq [1], using Microcal Origin software. Note also that the DB domain shows no non-specific binding to scrambled dsDNA oligos (○).

dependence on the buffer employed due to varying contributions from coupled protonation/deprotonation. Our data indeed suggest that the ΔH_{obs} for the binding of DB domain of ERα to DNA is highly dependent on the nature of buffer conditions employed (Figure 3-2). Thus, the ΔH_{obs} of binding goes from being highly exothermic (-30.52 kcal/mol) in Phosphate buffer to being endothermic (+9.82 kcal/mol) in Tris buffer and thereby mirrors the ΔH_{ion} of the respective buffers ranging from +1.22 kcal/mol to +11.35 kcal/mol [*115-117*]. This salient observation demonstrates that the binding of DB domain of ERα to DNA is directly coupled to proton uptake. Although such coupled equilibrium could result from the protonation of DNA bases, the sidechain moieties of D190, H196 and E203 within the DB domain must be considered as the major suspects for proton

Table 3-1

Observed thermodynamic parameters for the binding of ERE duplex to the wildtype DB domain of ERα in various buffers at pH 7.0 and 25°C

	K_{obs} / nM	ΔH_{obs} / kcal.mol^{-1}	$T\Delta S_{obs}$ / kcal.mol^{-1}	ΔG_{obs} / kcal.mol^{-1}
Phosphate	43 ± 13	-30.52 ± 0.27	-20.44 ± 0.38	-10.08 ± 0.20
Hepes	59 ± 6	-17.22 ± 0.50	-7.34 ± 0.53	-9.88 ± 0.07
Tricine	238 ± 84	-6.06 ± 1.54	+3.02 ± 1.68	-9.08 ± 0.23
Tris	336 ± 6	+9.82 ± 0.64	+19.10 ± 0.52	-8.94 ± 0.12

The binding stoichiometries to the fits agreed to within ±10%. Errors were calculated from 3-4 independent measurements. All errors are given to one standard deviation.

uptake due to their close proximity to the negatively charged phosphate backbone of DNA — a situation that would be inconceivable on thermodynamic grounds bar their protonation.

It is also of worthy note that while the enthalpy is favorable for the binding of ERα to DNA in Phosphate buffer, it contributes a substantial energetic penalty in Tris buffer (Table 3-1). In fact, close scrutinization of thermodynamic parameters observed for the binding of ERα to DNA in various buffers suggests that while enthalpy solely drives this protein-DNA interaction in Phosphate and Hepes buffers of low ionization enthalpies, entropy plays a major role in Tris and Tricine buffers of high ionization enthalpies. In particular, in the case of Tris buffer, it is the entropy that drives binding against the backdrop of enthalpic penalty. These observations imply that physiological settings with low ionization enthalpies are likely to favor the binding of ERα to DNA while the opposing conditions may be somewhat inhibitory. This is indeed corroborated by the fact that the binding affinity for ERα-DNA complexation drops by nearly an order

of magnitude from 43nM in Phosphate buffer to 336nM in Tris buffer (Table 3-1). It should be borne in mind that the favorable enthalpic contributions to binding largely result from the release of heat upon the formation of tight electrostatic interactions, hydrogen bonding and hydrophobic contacts between protein and DNA. Thus, in buffers of low ionization enthalpy, the favorable enthalpic contribution is only slightly offset to compensate for proton-coupled equilibrium rendering enthalpy as the sole driving force accompanied by entropic penalty. In contrast, in buffers of high ionization enthalpy, the favorable enthalpic contribution is largely offset and even overridden by the enthalpic penalty due to the proton-coupled equilibrium with entropy either contributing favorably or serving as the sole driving force at the expense of enthalpy.

This reciprocal relationship between enthalpy and entropy lies in the enthalpy-entropy compensation phenomenon [*118-122*] — macromolecular interactions are compensated by equal but opposing entropic changes such that there is little or no net gain in the overall free energy. Thus, buffers of high ionization enthalpy gain a substantial increase in entropy upon the release of a proton, presumably due to an increase in the degrees of freedom that become available to water molecules after being freed from their hydration shell surrounding the exchangeable proton prior to its release. However, in the case of buffers of low ionization enthalpy, the exchangeable proton would be expected to be more "economically" hydrated such that the release of water molecules from the rather small hydration shell contributes relatively little to the overall entropy gain but at the same time draws less heat to be removed. Such enthalpy-entropy compensations for the binding of DB domain to ERE duplex in various buffers are illustrated in Figure 3-3a. Consistent with the foregoing arguments, it should also be

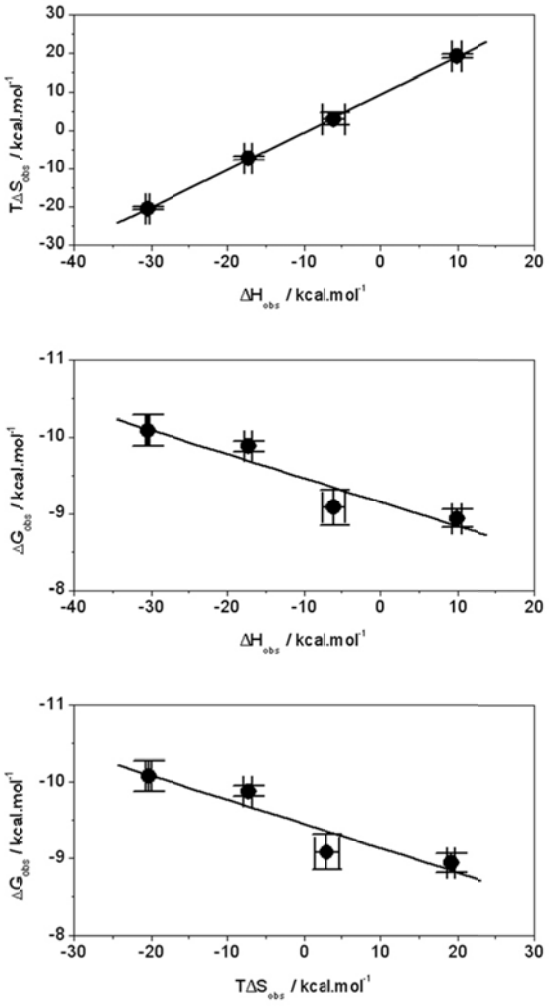

Figure 3-3: Inter-dependence of observed enthalpic change (ΔH_{obs}), entropic change ($T\Delta S_{obs}$) and free energy change (ΔG_{obs}) for the binding of ERE duplex to the wildtype DB domain of ERα in various buffers. (a) ΔH_{obs}-$T\Delta S_{obs}$ plot. (b) ΔH_{obs}-ΔG_{obs} plot. (c) $T\Delta S_{obs}$-ΔG_{obs} plot. Note that the solid lines represent linear fits to the data in all plots. All error bars were calculated from 3-4 independent measurements and are given to one standard deviation.

noted that while the increase in the binding affinity of the DB domain to DNA correlates with overall favorable enthalpy change in various buffers, the increase in favorable entropy change seems to oppose such protein-DNA interactions (Figures 3-3b and 3-3c).

3.4.2 Residues H196 and E203 serve as sole proton acceptors upon the binding of ERα to DNA

For processes in which ligand binding is coupled to proton exchange, the observed enthalpy (ΔH_{obs}) is related to the ionization enthalpy (ΔH_{ion}) by the relationship

$\Delta H_{obs}=2\Delta m\Delta H_{ion}+\Delta H_{int}$, where Δm is the net change in the number of protons absorbed or released per DB monomer upon binding to DNA and ΔH_{int} is the intrinsic binding enthalpy due to direct protein-DNA interactions and protonation of ionizable moities. A plot of ΔH_{obs} versus ΔH_{ion} should thus yield a linear curve with the slope $2\Delta m$ and y-intercept equal to ΔH_{int}. As shown in Figure 3-4, such analysis reveals that the binding of wildtype DB domain (DB_WT) of ERα to DNA results in the uptake of two protons per DB monomer. It should be noted here that a positive slope equates to proton uptake and a negative slope to proton release in this analysis. The fact that the binding of each monomer of DB domain to DNA is coupled to a net uptake of two protons implies that at least two of the three possible residues in D190, H196 and E203 may serve as proton acceptors. Could it be possible that only two of these residues are involved in proton uptake, or do all three residues fractionally contribute to a net uptake of two protons?

To address this question, we introduced single alanine substitutions at positions D190, H196 and E203 within the DB domain and then conducted the binding of these mutant domains to ERE duplex using ITC. The ΔH_{obs}-ΔH_{ion} plot for the binding of D190A mutant of DB domain (DB_D190A) to DNA reveals that there is no net change in the number of protons exchanged relative to DB_WT domain (Figure 3-4), implying that the residue D190 is most likely not responsible for the proton-coupled equilibrium observed here. In striking contrast, the ΔH_{obs}-ΔH_{ion} plots for the binding of H196A (DB_H196A) and E203A (DB_E203A) mutants of the DB domain to DNA reveal that only one proton is exchanged in each case (Figure 3-4), arguing strongly that the residues H196 and E203 are the sole sites of protonation.

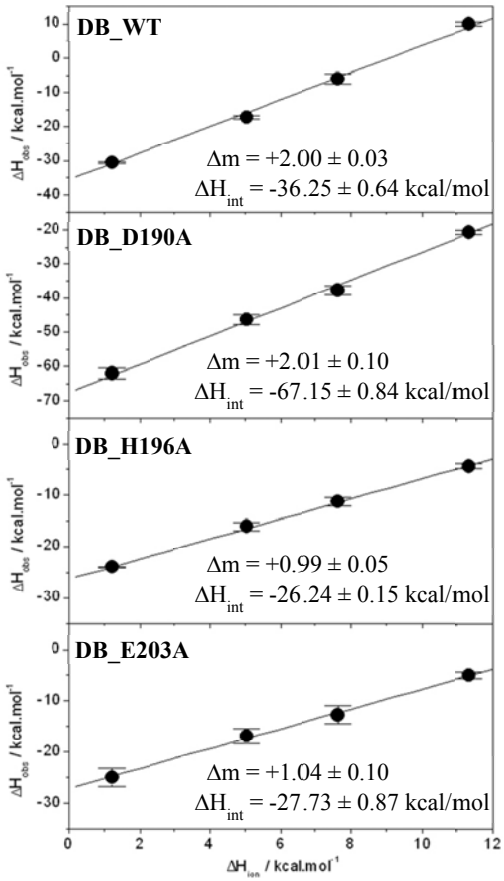

Figure 3-4: Dependence of observed enthalpy (ΔH_{obs}) as a function of ionization enthalpy (ΔH_{ion}) of various buffers upon the binding of ERE duplex to the wildtype DB domain (DB_WT), the D190A single mutant of DB domain (DB_D190A), the H196A single mutant of DB domain (DB_H196A) and the E203A single mutant of DB domain (DB_E203A) of ERα at pH 7.0 and 25°C. The ΔH_{ion} of various buffers used were +1.22 kcal/mol (Phosphate), +5.02 kcal/mol (Hepes), +7.64 kcal/mol (Ticine) and +11.35 kcal/mol (Tris) [*115-117*]. The solid lines within each panel represent fit of data points to Eq [4]. Note that the net change in the number of protons (Δm) absorbed or released per DB monomer upon binding to DNA and the intrinsic binding enthalpy (ΔH_{int}) due to direct protein-DNA interactions and protonation of ionizable moieties for each DB construct are provided within the corresponding panels. Error bars were calculated from 3-4 independent measurements. All errors are given to one standard deviation.

Table 3-2 provides complete thermodynamic parameters for the binding of wildtype and various mutants of the DB domain to ERE duplex in Phosphate buffer. It is clearly evident from these data that while the D190A mutation has little effect on the binding affinity of DB domain to DNA, H196A and E203A mutations both reduce the binding affinity by several folds. Remarkably, the binding of double mutant H196A/E203A of DB domain (DB_AA) to DNA is about an order of magnitude weaker

Table 3-2

Observed thermodynamic parameters for the binding of ERE duplex to wildtype and various mutant constructs of the DB domain of ERα in Phosphate buffer at pH 7.0 and 25°C

	K_{obs} / nM	ΔH_{obs} / kcal.mol^{-1}	$T\Delta S_{obs}$ / kcal.mol^{-1}	ΔG_{obs} / kcal.mol^{-1}
DB_WT	43 ± 13	-30.52 ± 0.27	-20.44 ± 0.38	-10.08 ± 0.20
DB_D190A	67 ± 6	-62.05 ± 1.64	-52.33 ± 1.48	-9.80 ± 0.06
DB_H196A	190 ± 24	-23.95 ± 0.08	-14.76 ± 0.14	-9.18 ± 0.07
DB_E203A	172 ± 48	-24.97 ± 1.81	-15.66 ± 1.97	-9.26 ± 0.17
DB_AA	387 ± 90	-23.77 ± 2.29	-15.00 ± 2.36	-8.77 ± 0.14
DB_S193A	102 ± 10	-45.10 ± 1.29	-35.54 ± 1.25	-9.55 ± 0.04
DB_Y197A	316 ± 7	-31.12 ± 0.41	-22.24 ± 0.37	-8.88 ± 0.02
DB_S201A	119 ± 9	-28.70 ± 0.69	-19.23 ± 0.64	-9.46 ± 0.05
DB_K206A	313 ± 6	-30.86 ± 0.86	-21.97 ± 0.85	-8.89 ± 0.07
DB_K210A	326 ± 11	-25.96 ± 0.17	-17.09 ± 0.18	-8.86 ± 0.09
DB_R211A	745 ± 76	-8.21 ± 0.09	+0.17 ± 0.10	-8.38 ± 0.11

The various constructs of the DB domain are the wildtype consruct (DB_WT), the single mutant constructs (DB_D190A, DB_H196A, DB_E203A, DB_S193A, DB_Y197A, DB_S201A, DB_K206A, DB_K210A, DB_R211A) and the H196A/E203A double mutant construct (DB_AA). The binding stoichiometries to the fits agreed to within ±10%. Errors were calculated from 3-4 independent measurements. All errors are given to one standard deviation.

relative to the wildtype DB domain (DB_WT), arguing further that both H196 and E203 residues are likely involved in proton uptake upon the binding of ERα to DNA. It should however be noted that the poor stability of the DB_AA construct made measurements feasible only in Phosphate buffer and no reliable analysis could be carried out in other buffers for direct comparison with the DB_WT construct. It is also of interest that while the H196A and E203A mutations result in the reduction of enthalpy for the binding of DB domain to DNA by about 5-7 kcal/mol due to removal of enthalpic contribution of

protonation at H196 or E203 and corresponding protein-DNA interactions at these positions, the enthalpy change for the D190A mutation is nearly two-fold more favorable relative to the wildtype DB domain (Table 3-2). It is thus conceivable that the D190A mutation results in local secondary and tertiary structural changes within the DB domain and that the partial folding of DB_D190A mutant domain upon binding to DNA also favorably contributes to the binding enthalpy. However, such enhancement in favorable enthalpy does not translate into higher binding affinity of D190A mutant domain to DNA due to an equally compensating entropic contribution as discussed in the previous section.

3.4.3 pH tightly regulates the binding of DB domain of ERα to DNA

In an effort to further support the notion that the residues H196 and E203 serve as the sole sites of protonation upon the binding of ERα to DNA, we analyzed the binding of wildtype (DB_WT) and double mutant H196A/E203A (DB_AA) constructs of DB domain to ERE duplex as a function of solution pH (Figure 3-5). Our data reveal that while the binding affinity of DB_WT construct to DNA is sharply dependent on solution pH in a sigmoidal fashion, binding affinity of DB_AA construct to DNA is independent of solution pH (Figure 3-5a). Taken collectively, our data suggest strongly that the binding of DB domain of ERα is coupled to proton uptake and that the sidechain moieties of H196 and E203 serve as sole proton acceptors in this capacity.

It is however of worthy note that the protonation of H196 and E203 within ERα could precede or follow the subsequent binding of DNA. In this manner, ERα could bind to DNA in both the protonated and unprotonated forms. Figure 3-6 provides a thermodynamic cycle for the various equilibria linked to the binding of ERα to DNA. It

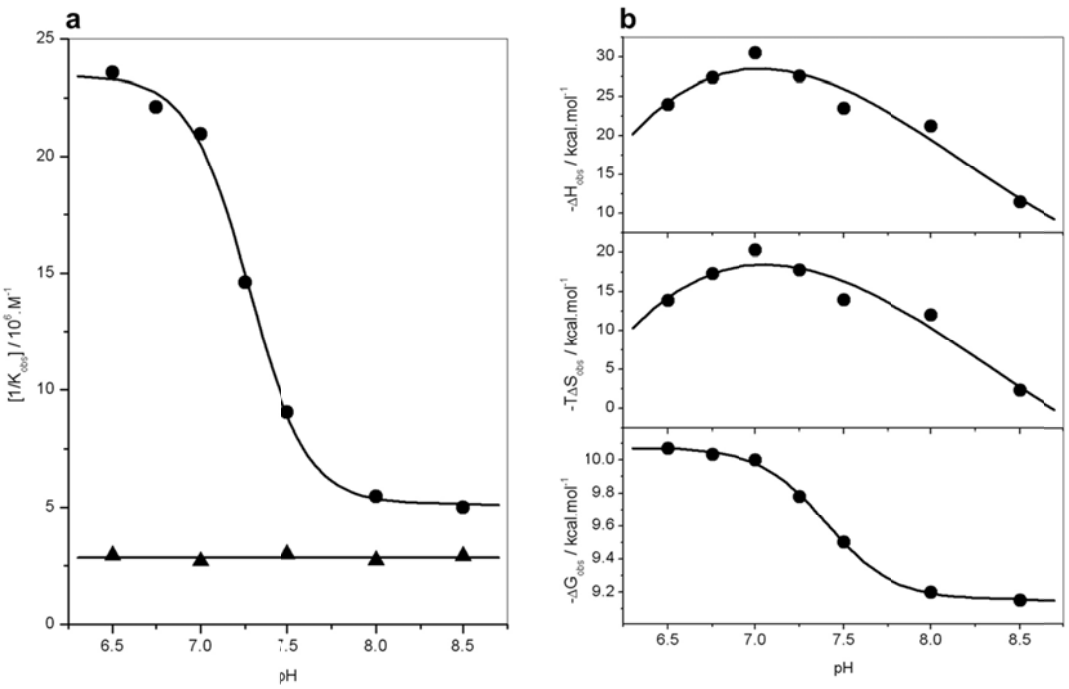

Figure 3-5: Dependence of thermodynamics on pH for the binding of ERE duplex to the wildtype DB domain (DB_WT) and the H196A/E203A double mutant of DB domain (DB_AA) of ERα in Phosphate buffer at 25°C. (a) Representative $[1/K_{obs}]$-pH plots for the DB_WT (●) and DB_AA (▲) domains. The solid lines respectively indicate sigmoidal and linear fits to data points for clarity. (b) Representative ΔH_{obs}-pH (top panel), $T\Delta S_{obs}$-pH (middle panel) and ΔG_{obs}-pH (bottom panel) plots for the DB_WT domain. In the top and middle panels, the solid lines indicate polynomial fits to data points for clarity. In the bottom panel, the solid line indicates sigmoidal fit to data points for clarity.

is clearly evident from such a cycle that the binding of ERα to DNA may or may not be coupled to proton uptake depending on solution pH. Thus, under low pH values, ERα may become fully protonated prior to binding DNA. On the other hand, under high pH values, ERα may become fully unprotonated and the proton uptake may be decoupled for its binding to DNA. That ERα can bind to DNA both in the protonated and unprotonated forms is further supported by the sigmoidal response of binding affinity (K_{obs}) of DB_WT to DNA as a function of pH (Figure 3-5a). Thus, the plateau values of $[1/K_{obs}]$-pH plot at low and high pH values correspond to the intrinsic binding affinities of the protonated and unprotonated forms of DB_WT domain to DNA, respectively.

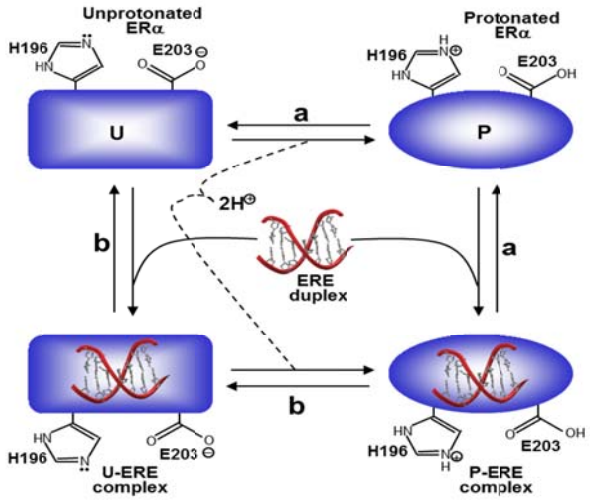

Figure 3-6: A thermodynamic cycle for the various equilibria linked to the binding of ERα to DNA. (a) ERα becomes protonated in the free form and the resulting protonated form (P) binds to DNA. (b) ERα binds to DNA in the unprotonated form (U) and the resulting liganded form becomes protonated.

It should also be noted that the ΔH_{obs}-pH and $T\Delta S_{obs}$-pH plots for the binding of DB_WT domain to DNA display bell-shaped curves characteristic of a proton-coupled ligand-binding event (Figure 3-5b, top and middle panels). Such behavior arises due to the fact that the enthalpy of protonation of free form is different from that of liganded form. Since the ratio of free and liganded forms of the protein varies as a function of pH in going from a low pH value to high, the enthalpic contribution due to their protonation to the observed enthalpy varies accordingly, reaching zero at the extreme values and a maximum in between these extremes where the ratio of the protonation of free form to liganded form equals unity. Expectedly, the ΔG_{obs}-pH plot for the binding of DB_WT to DNA follows sigmoidal behavior in agreement with the ability of both the protonated and unprotonated forms to bind to DNA with distinct affinities (Figure 3-5b, bottom panel).

3.4.4 Electrostatic surface potentials reveal that the protonation of H196 and E203 optimizes thermodynamic constraints

In an attempt to rationalize the effect of protonation of H196 and E203 on electrostatics at the protein-DNA interface, we generated molecular surfaces of the DB

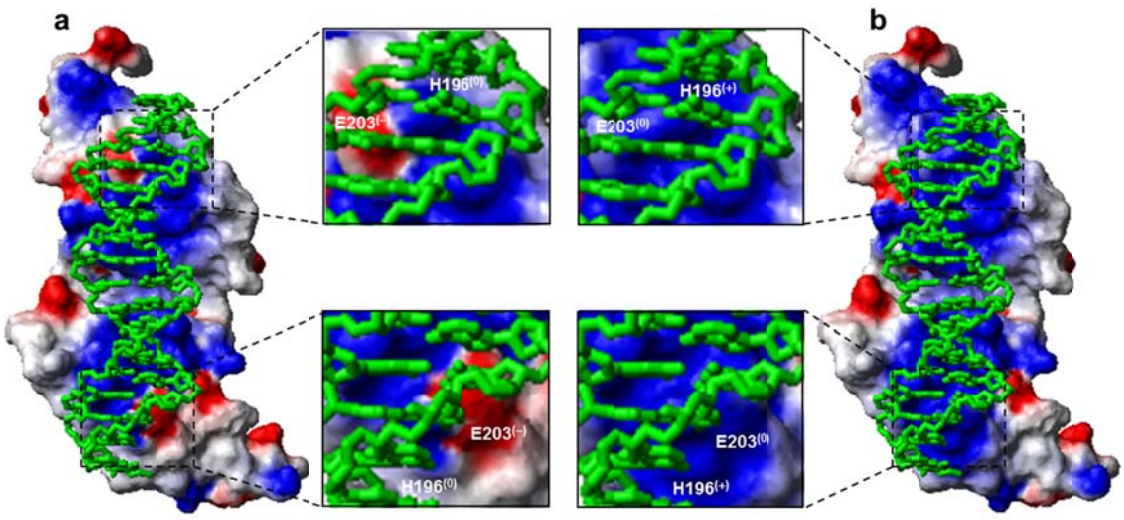

Figure 3-7: Molecular surfaces depicting electrostatic potentials of the DB domain of ERα containing H196 and E203 in unprotonated forms (a) and protonated forms (b) in complex with the ERE duplex. The blue and red colors respectively denote the density of positive and negative charges, while the apolar and polar surfaces are indicated by white/gray color on the molecular surfaces. In the expanded views, the locations of H196 and E203 are clearly marked on the molecular surfaces with the parenthesis indicating the overall charge on each residue under unprotonated and protonated forms. The ERE duplex is displayed as a stick model and colored green for clarity.

domain of ERα in complex with the ERE duplex depicting protein electrostatic potentials (Figure 3-7). Our data reveal how such protonation switches polarization of protein surface at residues H196 and E203 so as to render it thermodynamically more favorable for coming into contact with DNA. In the free conformation, H196 and E203 occupy what appear to be respectively neutral and negatively charged spots on the protein surface that is destined to come in close contact with DNA. It is further evident that while the presence of neutral charge at H196 may not in any way compromise the subsequent binding of DNA, protonation at this position could bring about favorable energetic contributions as a direct result of favorable electrostatic interactions with the negatively charged phosphate backbone. In contrast, the buildup of negative charge at E203 would hamper the subsequent binding of DNA due to electrostatic repulsions with the negatively charged phosphate backbone suggesting that protonation at this position would

relieve such energetic barriers. Taken together, the electrostatic surface potentials of the DB domain of ERα alone and in complex with DNA argue strongly that the protonation of H196 and E203 would optimize thermodynamic constraints so to allow the two molecular surfaces to come in close proximity to attain a tight molecular fit worthy of the rather high affinity that this DNA-protein complex displays.

3.4.5 Proton-coupled binding to DNA appears to be a hallmark of nuclear receptor family

In an attempt to analyze the extent to which the ability of ERα to become protonated upon binding to DNA is shared by other members, we generated amino acid sequence alignment of the DB domains of the entire human NR family (Figure 3-8). It should be noted that the DB domains of NR family are poorly conserved and display less than 15% sequence identity outside the quartet of cysteine residues involved in coordinating the Zn^{2+} divalent ion within each of the two Zinc fingers of the DB domain. Thus, the residues conserved among the various DB domains bear a significant importance and must have co-evolved for a common physiological function.

Remarkably, H196 and E203 rank among these conserved residues within the DB domains of human NR family. Thus, while H196 is absolutely conserved within all members of human NR family, E203 is predominantly conserved in most members, with notable exceptions being the androgen receptor (AR), glucocorticoid receptor (GR), mineralocorticoid receptor (MR) and progesterone receptor (PR), which all have a glycine substitution for E203. Interestingly, E203 is substituted by an asparagine in photoreceptor-specific nuclear receptor (PNR), implying that hydrogen bonding at this position may play a critical role in protein-DNA interaction pertinent to the physiological function of this nuclear receptor. Finally, E203 is substituted by a related acidic and

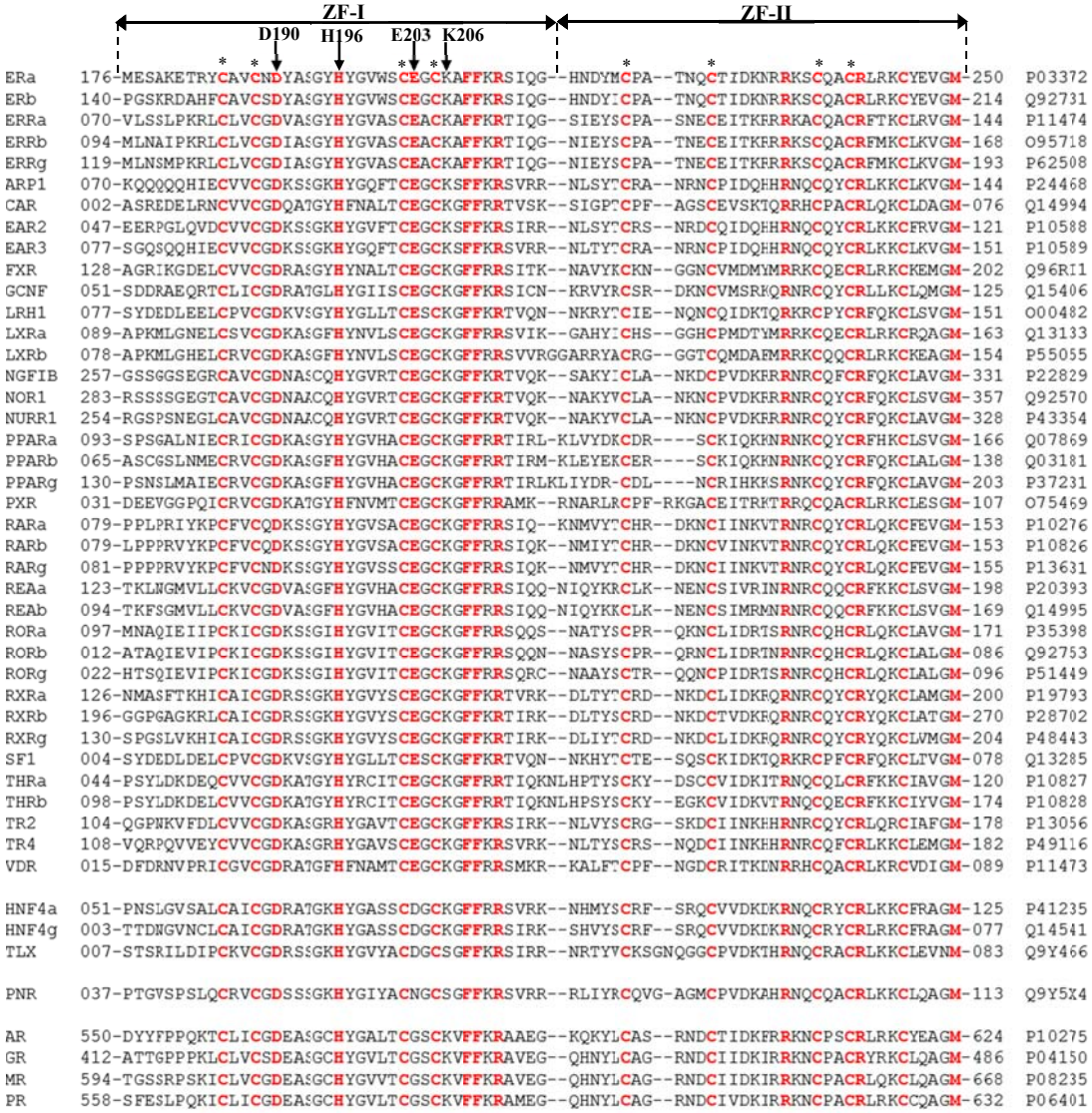

Figure 3-8: Amino acid sequence alignment of DB domains of all known members of human NR family. Absolutely conserved residues are shown in red, while all other residues are depicted in black. Each member is denoted by its acronym in the left column with the corresponding Expasy code provided in the right column for access to complete proteomic details on each member. The numerals hyphenated to amino acid sequence at each end denote the boundaries of DB domains for each member. The cysteine residues within each of the two Zinc fingers of DB domains, denoted ZF-I and ZF-II, that coordinate the Zn^{2+} ion in a tetrahedral arrangement are marked by asterisks. Residues D190, H196, E203 and K206, located within the DB domain of ERα, are indicated by vertical arrows.

ionizable aspartate residue in hepatocyte nuclear factor 4a (HNF4a), hepatocyte nuclear factor 4g (HNF4g) and the tail-less orphan receptor (TLX), indicating that protonation at this position upon binding to DNA may also be critical for these nuclear receptors. Thus,

the highly conserved nature of H196 and E203 among the functionally-diverse members of the human NR family argues strongly that the proton-coupled binding to DNA may have evolved as a general mechanism for nuclear receptor function and regulation.

It is also of worthy note that although the residue D190 is absolutely conserved among all members of the human NR family, implying that although it does not serve as a proton acceptor, it must also play a pivotal role in protein-DNA interactions pertinent to nuclear receptors. Additionally, our thermodynamic data indicate that the D190A substitution has no bearable effect on the binding affinity of DB domain of ERα to DNA (Table 3-2) — an observation that is in conflict with evolutionary constraints being placed upon this residue in the human NR family. Further scrutiny of the binding of wildtype DB domain (DB_WT) versus the D190A mutant (DB_D190A) to DNA suggests that although they bind with virtually indistinguishable affinities, the underlying thermodynamic forces display distinct features. Thus, while the binding of both domains is driven by favorable enthalpic factors accompanied by entropic penalties, DB_D190A generates twice as much heat relative to DB_WT implying that this residue may be critical for the folding and stability of ERα and that the more favorable heat likely results from the partial folding of DB_D190A mutant domain upon binding to DNA as noted earlier.

In an effort to further decipher the molecular basis of how nuclear receptors bind to their cognate DNA promoter elements with specificity, we also analyzed and compared the thermodynamics of binding of ERE duplex to DB domain of ERα containing alanine substitutions for a number of additional amino acid residues located at the protein-DNA interface (Table 3-2). These point mutations include S193A

(DB_S193A), Y197A (DB_Y197A), S201A (DB_S201A), K206A (DB_K206A), K210A (DB_K210A) and R211A (DB_R211A). As shown in Table 3-2, alanine substitution of these residues reduces the binding of DB domain to DNA by as little as 2-fold in the case of S193A mutation to as large as 17-fold in the case of R211A mutation relative to the wildtype construct, implying that these residues contribute differentially to the free energy of binding. Of these six residues at the protein-DNA interface, only R211 is absolutely conserved within the DB domains of all nuclear receptors (Figure 3-8). This salient observation suggests strongly that in addition to H196 and E203, R211 is also likely to be a critical residue involved in the binding of all nuclear receptors to their cognate DNA sequences. However, the fact that the residues S193, Y197, S201, K206 and K210 show variability within the DB domains of nuclear receptors argues strongly in favor of their role in determining the specificity of binding of nuclear receptors to DNA. Nonetheless, it should be borne in mind that complete understanding of molecular basis of DNA-specificity of nuclear receptors awaits detailed thermodynamic analysis coupled with site-directed mutagenesis of specific amino acid residues within DB domains of other nuclear receptors that we hope to accomplish in our future studies.

3.5 Concluding remarks

Nuclear receptor function is tightly regulated by a multitude of post-translational modifications such as phosphorylation, acetylation, sumoylation, ubiquitination and glycosylation [*123-126*]. However, such modifications usually occur in regions outside the DB domain. The fact that the DB domain of ERα is directly regulated via proton-coupled equilibrium of two critical residues, H196 and E203, located at the protein-DNA interface not only adds to the repertoire of tricks and treats employed by nuclear receptors

but also bears significant implications for furthering our understanding of this important family of transcription factors.

Although the binding of DB domain of ERα to DNA appears to be coupled to proton uptake, it is not clear from our data as to how such coupled equilibrium might dictate the physiological role of this important nuclear receptor. Changes in intracellular pH regulate a multitude of cellular processes such as metabolic homeostasis and apoptosis [*127*]. Furthermore, it is believed that ionizable residues within proteins sense such changes and activate a variety of proton pumps and ion transporters that in turn mediate extracellular transport of protons and anions to regulate intracellular pH [*128-130*]. It is thus conceivable that changes in intracellular pH may also tightly regulate the transcriptional activity of ERα through direct modulation of two ionizable residues, H196 and E203, located at the protein-DNA interface. Protonation of such residues would clearly enhance intermolecular hydrogen bonding and electrostatic interactions critical to driving this key protein-DNA interaction and vice versa. Although pKa values of sidechains of histidine and glutamate within proteins are respectively around 6 and 4 [*131*], these are likely to be influenced by the neighboring ionizable amino acid residues in the DB domain as noted earlier (Figure 3-1). Thus, protonation/deprotonation of H196 and E203 may not necessarily require large changes but may be mediated by small changes in intracellular pH. Whatever the exact physiological role of proton-coupled equilibrium observed here, our current study clearly warrants further investigating the role of pH in physiological processes governed by ERα and other nuclear receptors.

In the crystal structure of the DB domain of ERα in complex with ERE duplex solved nearly two decades ago [*66*], it was proposed that the negative charge on E203

was largely neutralized through the formation of a salt bridge with the neighboring K206. On the contrary, our study here shows that the negative charge on E203 is rather neutralized through its protonation allowing it to participate in the formation of hydrogen bonding with DNA in a more harmonious manner. Additionally, the crystal structural analysis also suggested the involvement of H196 in dictating protein-DNA interactions through hydrogen bonding with the phosphate backbone. The fact that H196 acquires a net positive charge through protonation upon the binding of DNA suggests that H196 is more likely to engage in the formation of a salt bridge with the phosphate backbone. Taken together, our study exquisitely reveals how a combined approach involving site-directed mutagenesis in conjunction with thermodynamics can complement structural data and further define key residues involved in protein-DNA interactions.

In short, our present study demonstrates that the protonation of H196 and E203 in ERα is coupled to the binding of DNA and that such protonation is required for high-affinity protein-DNA interaction through thermodynamic optimization of intermolecular contacts. Given that H196 and E203 are conserved in a vast majority of ~50 members of the nuclear receptor family, our findings suggest that the nuclear receptors may act as sensors of intracellular pH and bear important consequences for a paradigm shift of their molecular action. Finally, the proton-coupled equilibrium characterized here may serve as a novel target for therapeutic intervention of nuclear receptors.

Chapter 4: Genetic Variations within the ERE Motif Modulate Plasticity and Energetics of Binding of DNA to the ERα Nuclear Receptor

4.1 Summary

Upon binding to estrogens, the ERα nuclear receptor acts as a transcription factor and mediates a multitude of cellular functions central to health and disease. Herein, using isothermal titration calorimetry (ITC) and circular dichroism (CD) in conjunction with molecular modeling (MM), we analyze the effect of symmetric introduction of single nucleotide variations within each half-site of the estrogen response element (ERE) on the binding of ERα nuclear receptor. Our data reveal that ERα exudes remarkable tolerance and binds to all genetic variants in the physiologically relevant nanomolar-micromolar range with the consensus ERE motif affording the highest affinity. We provide rationale for how genetic variations within the ERE motif may reduce its affinity for ERα by orders of magnitude at the atomic level. The data also suggest that the introduction of genetic variations within the ERE motif allows it to sample a much greater conformational space. Surprisingly, ERα displays no preference for binding to ERE variants with higher AT content, implying that any advantage due to DNA plasticity may be largely compensated by unfavorable entropic factors. Collectively, our study bears important consequences for how genetic variations within DNA promoter elements may fine-tune the physiological action of ERα and other nuclear receptors.

4.2 Overview

Nuclear receptors (NRs) act as ligand-modulated transcription factors and orchestrate a plethora of cellular functions central to health and disease [*3-6*]. Some notable examples of ~50 members of the NR family are the androgen receptor (AR), estrogen receptor α (ERα), glucocorticoid receptor (GR) and progesterone receptor (PR).

All members of the NR family are evolutionarily related and share a core modular architecture comprised of a central DNA-binding (DB) domain flanked between an N-terminal trans-activation (TA) domain and a C-terminal ligand-binding (LB) domain [*58-60*]. Nuclear receptors become activated by lipophilic messengers such as hormones and vitamins secreted by appropriate tissues. Upon their diffusion through the cell membrane, the binding of these ligands to the LB domain culminates in a series of events involving the translocation of nuclear receptors into the nucleus and subsequent modulation of expression of target genes [*9-11*].

ERα mediates the action of estrogens such as estradiol and its hyperactivation leads to the genesis of large fractions of breast cancer [*8, 108-113*]. The DB domain of ERα binds as a homodimer with a two-fold axis of symmetry to the ERE motif, containing the AGGTCAnnnTGACCT consensus sequence, located within the promoters of target genes [*65*]. DNA-binding is accomplished through a pair of tandem C4-type Zinc fingers, with each finger containing a Zn^{2+} ion coordinated in a tetrahedral arrangement by four highly conserved cysteine residues [*66, 67*]. It is important to note that the DB domain contains two Zinc fingers. The first Zinc finger (ZF-I) within each monomer of DB domain recognizes the hexanucleotide sequence 5'-AGGTCA-3' within the major groove at each end of the ERE duplex, whilst the second Zinc finger (ZF-II) is responsible for the homodimerization of DB domain upon DNA binding.

Although nuclear receptors recognize the target genes in a DNA-sequence-dependent manner, genetic variations within specific promoter response elements are extremely common within the eukaryotic genomes [*65*]. Given that the nucleotide sequence is a key determinant of the ability of DNA to behave as a flexible polymer and

undergo physical phenomena such as bending, stretching, deformation and distortion coupled with its ability to exist in various structural conformations (such as the B-DNA, A-DNA and Z-DNA) [*132-134*], our knowledge of how genetic variations within the promoter elements influence the ability of nuclear receptors to bind and subsequently affect gene transcription remains largely elusive. Several lines of evidence indeed suggest that genetic variations within the cognate response elements play a key role in modulating the affinity and specificity of binding of AR, GR and PR nuclear receptors [*72-75*]. In an effort to build on these earlier studies, we set out here to investigate how single nucleotide variations within the estrogen response element (ERE) affect the plasticity and energetics of binding of DNA to the ERα nuclear receptor.

4.3 Experimental Procedures

4.3.1 Protein Preparation

The DB domain (residues 176-250) of human ERα (Expasy# P03372) was cloned into pET101 bacterial expression vector with a C-terminal polyhistidine (His)-tag, to aid in protein purification through Ni-NTA affinity chromatography, using Invitrogen TOPO technology. The protein was subsequently expressed in Escherichia coli BL21*(DE3) bacterial strain (Invitrogen) and purified on a Ni-NTA affinity column using standard procedures. Briefly, bacterial cells were grown at 20°C in TB media supplemented with 50μM $ZnCl_2$ to an optical density of 0.5 at 600nm prior to induction with 0.5mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The bacterial culture was further grown overnight at 20°C and the cells were subsequently harvested and disrupted using a BeadBeater (Biospec). After separation of cell debris at high-speed centrifugation, the cell lysate was loaded onto a Ni-NTA column and washed extensively with 20mM

imidazole to remove non-specific binding of bacterial proteins to the column. The recombinant protein was subsequently eluted with 200mM imidazole and dialyzed against an appropriate buffer to remove excess imidazole. Further treatment on a Hiload Superdex 200 size-exclusion chromatography (SEC) column coupled in-line with GE Akta FPLC system led to purification of recombinant DB domain to apparent homogeneity as judged by SDS-PAGE analysis. The identity of recombinant protein was confirmed by MALDI-TOF mass spectrometry. Final yield was typically between 5-10mg protein of apparent homogeneity per liter of bacterial culture. Protein concentration was determined by the fluorescence-based Quant-It assay (Invitrogen) and spectrophotometrically using an extinction coefficient of 14,940 $M^{-1}cm^{-1}$ calculated for the recombinant DB domain using the online software ProtParam at ExPasy Server [*81*]. Results from both methods were in an excellent agreement.

4.3.2 DNA synthesis

21-mer DNA oligos containing the consensus ERE motif (AGGTCAnnnTGACCT) and all possible symmetric single nucleotide variants were commercially obtained from Sigma Genosys. The design of such oligos and their numbering relative to the central 3-bp spacer is illustrated in Figure 4-1. 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 [*83*]. Double-stranded DNA (dsDNA) oligos were generated as described earlier [*135*].

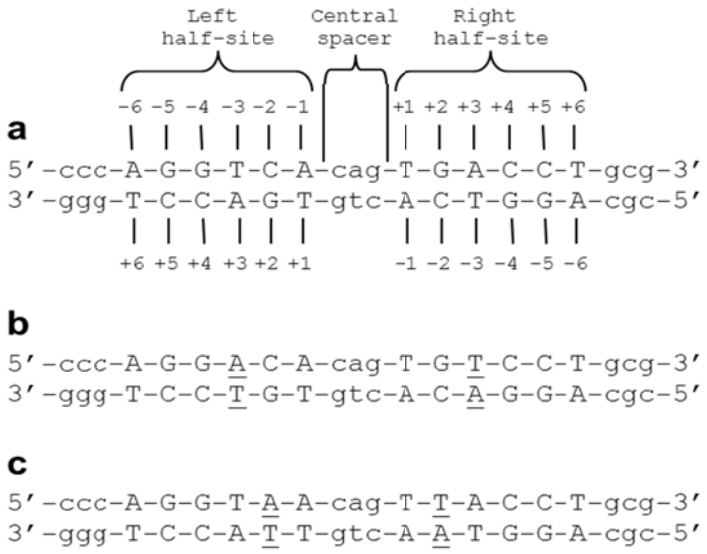

Figure 4-1: Nucleotide sequence of dsDNA oligos. (a) Consensus ERE motif. The AGGTCA and TGACCT half-sites are clearly marked. The consensus nucleotides within each half-site are capitalized whilst the flanking nucleotides and the intervening nucleotides within the central spacer are shown in small letters. The numbering of various nucleotides within each half-site relative to the CAG central spacer in the sense (upper) and antisense (lower) strands are indicated. (b) Am3Tp3 motif, wherein adenine and thymine are respectively substituted at –3 and +3 positions within the sense strand in a symmetric manner within each half-site. The variant nucleotides relative to the consensus ERE motif in both strands are underlined. (c) Am2Tp2 motif, wherein adenine and thymine are respectively substituted at –2 and +2 positions within the sense strand in a symmetric manner within each half-site. The variant nucleotides relative to the consensus ERE motif in both strands are underlined.

4.3.3 ITC Measurements

Isothermal titration calorimetry (ITC) experiments were performed on a Microcal VP-ITC instrument and data were acquired and processed using fully automized features in Microcal ORIGIN software. All measurements were repeated at least three times. Briefly, protein and DNA samples were prepared in 50mM Sodium phosphate containing 5mM β-mercaptoethanol at pH 7.0 and de-gassed using the ThermoVac accessory for 5min. The experiments were initiated by injecting 25 x 10μl aliquots of 50-200μM of a dsDNA oligo containing the ERE motif, or a variant thereof, from the syringe into the calorimetric cell containing 1.8ml of 5-10μM of DB domain of ERα at 25°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 molar ratio of dsDNA oligo to dimer-equivalent DB domain. 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 oligo in an identical manner. Control experiments with scrambled dsDNA oligos generated similar thermal power to that obtained for the buffer alone, implying that there was no non-specific binding of DB domain to non-cognate DNA. To extract various thermodynamic parameters, the binding isotherms were iteratively fit to a built-in one-site model by non-linear least squares regression analysis using the ORIGIN software as described previously [*114, 135*].

4.3.4 CD analysis

Circular dichroism (CD) measurements were conducted on a Bio-Logic MOS450/SFM400 spectropolarimeter thermostatically controlled with a water bath at 25°C. All data were acquired and processed using the Biokine software. Briefly, experiments were conducted on 20μM of dsDNA oligos containing the ERE motif, or a variant thereof, in 50mM Sodium phosphate at pH 7.0. All experiments were conducted in a quartz cuvette with a 2-mm pathlength in the wavelength range 190-310nm. Data were recorded with a slit bandwidth of 2nm at a scan rate of 3nm/min. Data were normalized against reference spectra to remove the contribution of buffer. The reference spectra were obtained in a similar manner on a 50mM Sodium phosphate at pH 7.0. Each data set represents an average of at least four scans acquired at 1nm intervals. Data were

converted to molar ellipticity, [θ], as a function of wavelength (λ) of electromagnetic radiation using the equation:

$$[\theta] = [(10^5 \Delta\varepsilon)/cl] \text{ deg.cm}^2\text{.dmol}^{-1}$$

where Δε is the observed ellipticity in mdeg, c is the dsDNA concentration in μM and l is the cuvette pathlength in cm.

4.3.5 Macromolecular modeling

Molecular modeling (MM) was employed to generate 3D atomic models of the DB domain of ERα in complex with dsDNA oligos containing the consensus ERE motif and the variant Am2Tp2 motif using the MODELLER software based on homology modeling [*100*]. In each case, the crystal structure of the DB domain of ERα in complex with a dsDNA oligo containing the ERE motif but with varying flanking sequences was used as a template (PDB entry 1HCQ). Additionally, for the atomic model of DB domain in complex with the Am2Tp2 motif, hydrogen bonding distance restraints were added between appropriate pairs of atoms to allow base pairing between A-2 and T+2 within each half-site. For each motif, a total of 100 atomic models were calculated and the structures with the lowest energy, as judged by the MODELLER Objective Function, were selected for further analysis. The atomic models were rendered using RIBBONS [*101*].

4.4 Results and discussion

4.4.1 ERα tolerates genetic variations within the ERE motif at the expense of reduced affinities

In order to assess the effect of genetic variations within the ERE motif on the binding of ERα, we analyzed the binding of DB domain of ERα to the consensus ERE motif and its genetic variants containing single nucleotide substitutions within each half-

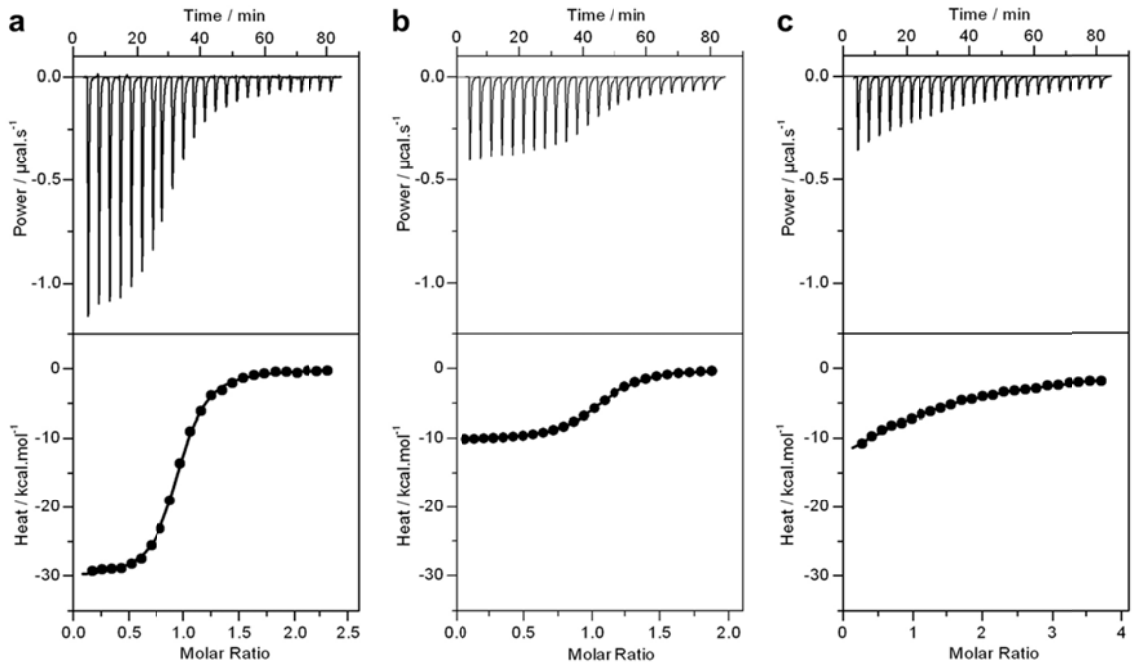

Figure 4-2: Representative ITC isotherms for the binding of DB domain of ERα to dsDNA oligos containing the consensus ERE motif (a), the variant Am3Tp3 motif (b) and the variant Am2Tp2 motif (c). The upper panels show the raw ITC data expressed as change in thermal power with respect to time over the period of titration. In the lower panels, change in molar heat is expressed as a function of molar ratio of corresponding dsDNA oligos to dimer-equivalent DB domain. The solid lines in the lower panels represent the fit of data to a one-site model, based on the binding of a ligand to a macromolecule assuming the law of mass action, using the ORIGIN software [*114, 135*].

site in a symmetrical manner using ITC (Figure 4-2 and Table 4-1). Our analysis suggests that the DB domain not only tolerates such genetic variations but also binds in the physiologically relevant nanomolar-micromolar range, with the consensus ERE motif affording highest affinity. These findings are thus consistent with the knowledge that genetic variations within the estrogen-responsive genes can dramatically affect the affinity of ERα-DNA interactions [*65*]. Although these affinities vary over nearly two orders of magnitude, binding is universally driven by favorable enthalpic changes accompanied by unfavorable entropy. The favorable enthalpic changes observed here are consistent with the formation of an extensive network of hydrogen bonding and ion pairing of amino acid residues with the bases and backbone phosphates at the protein-

Table 4-1

Thermodynamic parameters for the binding of DB domain of ERα to dsDNA oligos containing the consensus ERE (cERE) motif and single nucleotide symmetrical variants thereof obtained from ITC measurements

Motif	Sequence	K_d / μM	ΔH / kcal.mol^{-1}	TΔS / kcal.mol^{-1}	ΔG / kcal.mol^{-1}	Gene promoter
cERE	AGGTCAcagTGACCT	0.08 ± 0.01	-30.15 ± 0.28	-20.42 ± 0.23	-9.73 ± 0.05	Cell cycle kinase CDK5
Cm1Gp1	AGGTCCcagGGACCT	0.79 ± 0.01	-18.25 ± 0.25	-9.91 ± 0.25	-8.34 ± 0.01	Signaling protein WISP2
Gm1Cp1	AGGTCGcagCGACCT	1.58 ± 0.05	-22.76 ± 0.28	-14.83 ± 0.26	-7.93 ± 0.02	Hemoprotein CYP1B1
Tm1Ap1	AGGTCTcagAGACCT	0.85 ± 0.06	-22.59 ± 1.08	-14.29 ± 1.08	-8.30 ± 0.04	Transctiption factor HNF3A
Am2Tp2	AGGTAAcagTTACCT	7.02 ± 0.04	-25.57 ± 0.03	-18.53 ± 0.03	-7.04 ± 0.01	Nuclear receptor ERα
Gm2Cp2	AGGTGAcagTCACCT	2.31 ± 0.11	-20.06 ± 0.22	-12.35 ± 0.19	-7.70 ± 0.03	Mitotic protein PRCC
Tm2Ap2	AGGTTAcagTAACCT	0.11 ± 0.01	-24.36 ± 0.42	-14.83 ± 0.45	-9.53 ± 0.02	Nuclear receptor SHP
Am3Tp3	AGGACAcagTGTCCT	0.26 ± 0.08	-11.48 ± 0.53	-2.47 ± 0.35	-9.01 ± 0.18	Protease inhibitor CST5
Cm3Gp3	AGGCCAcagTGGCCT	0.86 ± 0.04	-19.58 ± 0.28	-11.29 ± 0.30	-8.28 ± 0.03	Cytoskeletal protein TNS1
Gm3Cp3	AGGGCAcagTGCCCT	0.50 ± 0.02	-20.06 ± 0.44	-11.45 ± 0.41	-8.61 ± 0.02	Retinoblastoma protein RBL2
Am4Tp4	AGATCAcagTGATCT	0.31 ± 0.01	-19.20 ± 0.17	-10.30 ± 0.14	-8.89 ± 0.01	Skeletal muscle protein MYOT
Cm4Gp4	AGCTCAcagTGAGCT	2.56 ± 1.01	-11.77 ± 0.41	-4.85 ± 0.20	-7.66 ± 0.24	Cytoskeletal protein EPS8
Tm4Ap4	AGTTCAcagTGAACT	1.17 ± 0.04	-23.87 ± 0.68	-15.76 ± 0.66	-8.10 ± 0.02	GTPase Rab7L1
Am5Tp5	AAGTCAcagTGACTT	0.50 ± 0.15	-17.70 ± 0.49	-9.08 ± 0.31	-8.62 ± 0.18	Cytokine IL20
Cm5Gp5	ACGTCAcagTGACGT	4.09 ± 0.19	-21.51 ± 0.20	-14.15 ± 0.17	-7.36 ± 0.03	DNA helicase RecQ4
Tm5Ap5	ATGTCAcagTGACAT	0.40 ± 0.01	-21.86 ± 0.87	-13.12 ± 0.88	-8.75 ± 0.01	MDR protein MRP5
Cm6Gp6	CGGTCAcagTGACCG	0.14 ± 0.01	-22.69 ± 0.52	-13.34 ± 0.54	-9.35 ± 0.02	Transcription factor PROP1
Gm6Cp6	GGGTCAcagTGACCC	0.15 ± 0.01	-26.10 ± 0.20	-16.76 ± 0.17	-9.34 ± 0.03	Caspase CASP7
Tm6Ap6	TGGTCAcagTGACCA	0.75 ± 0.07	-17.38 ± 0.36	-9.00 ± 0.31	-8.37 ± 0.05	Secretory protein TFF1

Note that the DNA sequence shown for all motifs corresponds to the sense strand only and the flanking nucleotides have been omitted for clarity (see Figure 1). The symmetrically-substituted nucleotides within each half-site relative to the cERE motif are underlined. One example of an estrogen-responsive gene promoter that contains at least one of the substitutions within the corresponding ERE motif is provided for physiological relevance [*136-138*]. The values for the affinity (K_d) and enthalpy change (ΔH) accompanying the binding of DB domain of ERα to dsDNA oligos were obtained from the fit of a one-site model, based on the binding of a ligand to a macromolecule using the law of mass action, to the corresponding ITC isotherms as described earlier [*114, 135*]. 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 $T\Delta S=\Delta H-\Delta G$. Binding stoichiometries generally agreed to within ±10%. Errors were calculated from at least three independent measurements. All errors are given to one standard deviation.

DNA interface [*66*]. The unfavorable entropic changes most likely result from the loss of conformational degrees of freedom that both the protein and DNA experience upon complexation. Given that the DNA undergoes bending upon binding to the DB domain

[*139*], the entropic penalty observed here may also in part be attributed to such physical distortion of DNA necessary for it to wrap around the protein so as to attain a close molecular fit. A comparative analysis of thermodynamic properties of variant motifs relative to the consensus ERE motif for binding to the DB domain reveals that the variant motifs exhibit remarkable thermodynamic versatility (Figure 4-3). Thus, while the binding of Am3Tp3 variant motif to the DB domain is only weaker by about three-fold relative to the consensus ERE motif, its underlying enthalpic and entropic contributions to the free energy are remarkably different. The much smaller entropic penalty incurred upon the binding of Am3Tp3 motif to the DB domain underscores how genetic variations within the ERE motif may dictate the underlying thermodynamics of protein-DNA interactions. On the other hand, the Am2Tp2 motif binds to the DB domain with an affinity that is nearly two orders of magnitude weaker relative to the ERE motif, yet both motifs share very similar underlying thermodynamic signatures. Taken together, our data presented above suggest strongly that genetic variations within the ERE motif modulate the energetics of binding of DB domain of ERα to DNA. Accordingly, genetic variations within the ERE motif at target gene promoters may play a key role in gauging the transcriptional output of ERα in response to estrogens. The genetic variations within the ERE motif may have thus evolved to provide a differential response to the expression of estrogen-responsive genes. Additionally, the genetic variations within the ERE motif may exert their effect by modulating the affinity of protein-DNA interactions through fine-tuning the contributions of the underlying thermodynamic forces to the free energy. In a related study from our laboratory, we demonstrated that genetic variations within the DNA response element of Jun-Fos heterodimeric transcription factor modulate its

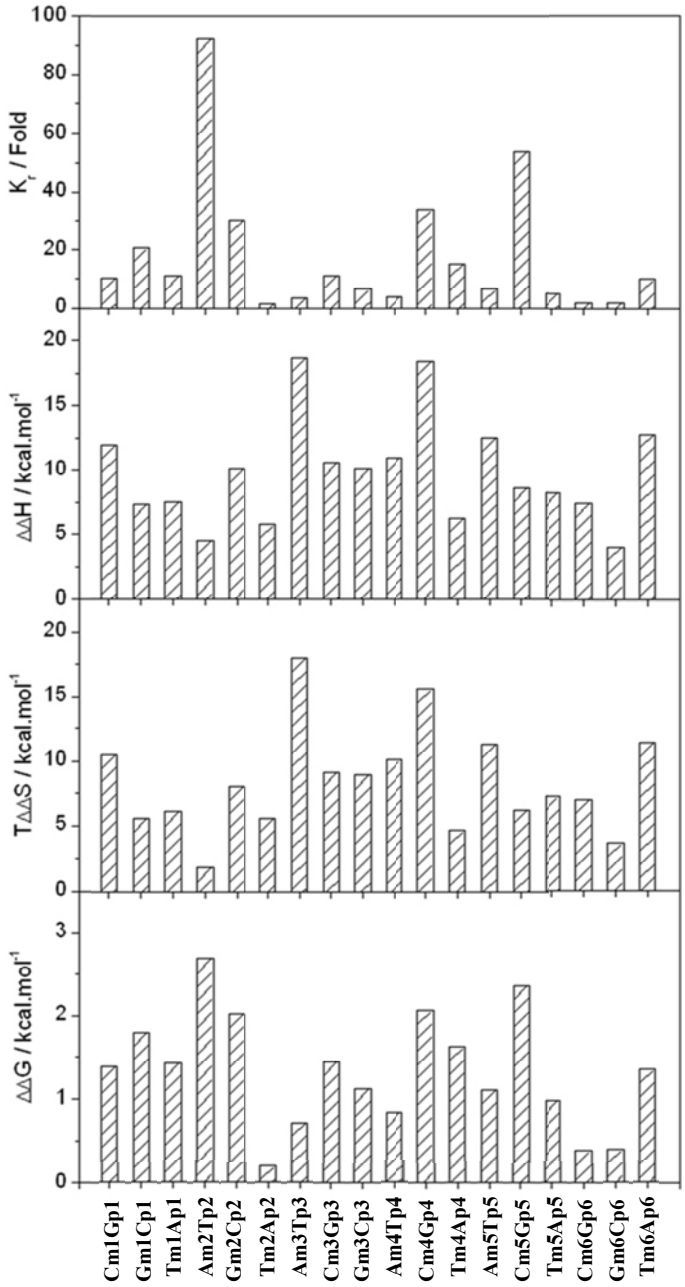

Figure 4-3: Analysis of the binding of DB domain of ERα to variant motifs relative to the consensus ERE motif in terms of relative binding affinity (K_r), relative enthalpic contribution ($\Delta\Delta H$) and relative entropic contribution ($T\Delta\Delta S$) to the relative free energy ($\Delta\Delta G$). K_r is defined as $K_r=K_v/K_c$, where K_v and K_c are respectively the binding affinities of the variant and consensus ERE motifs to the DB domain (Table 4-1). $\Delta\Delta H$ is defined as $\Delta\Delta H=\Delta H_v-\Delta H_c$, where ΔH_v and ΔH_c are respectively the enthalpy changes observed for the variant and consensus ERE motifs upon binding to the DB domain (Table 4-1). $T\Delta\Delta S$ is defined as $T\Delta\Delta S=T\Delta S_v-T\Delta S_c$, where $T\Delta S_v$ and $T\Delta S_c$ are respectively the entropic contributions observed for the variant and consensus ERE motifs upon binding to the DB domain (Table 4-1). $\Delta\Delta G$ is defined as $\Delta\Delta G=\Delta G_v-\Delta G_c$, where ΔG_v and ΔG_c are respectively the free energy changes observed for the variant and consensus ERE motifs upon binding to the DB domain (Table 4-1).

orientation [*140*]. Given that ERα also cooperates with ERβ in binding as a heterodimer to the promoters of target genes [*141*], it is conceivable that genetic variations within the ERE motif may also dictate the orientation of ERα-ERβ heterodimeric transcription factor. Such an orientation may in turn be an important determinant of the nature of other interacting cellular partners being recruited to the site of DNA transcription and thereby may play a key role in further modulating gene expression.

4.4.2 Binding of ERE motif and its genetic variants thereof to ERα is enthalpy-entropy compensated

Macromolecular interactions are often governed by enthalpy-entropy compensation phenomenon, whereby favorable enthalpic changes are largely compensated by unfavorable entropic factors, and vice versa, such that there is little or no gain in the overall free energy of binding. In an effort to test whether the binding of ERE motif and its variants thereof to the DB domain of ERα is also subject to enthalpy-entropy compensation, we generated the enthalpy-entropy plot (Figure 4-4a). Evidently, the binding of ERE motif and its variants thereof to the DB domain indeed appears to be enthalpy-entropy compensated. Consistent with these observations, it is also important to note that the enthalpic (ΔH) and entropic (TΔS) contributions for the binding of ERE motif and its variants thereof to the DB domain show poor correlation with the overall free energy (ΔG) (Figures 4-4b and 4-4c). For example, an increase in favorable ΔH or a decrease in unfavorable TΔS does not necessarily lead to an increase in ΔG and vice versa. In light of the knowledge that the binding of proteins to major grooves within the DNA is under enthalpic control [*142-148*], the negative contribution of entropic penalty to the free energy appears to be an equally important regulator of such protein-DNA interactions. Taken together, these salient observations bear important consequences for

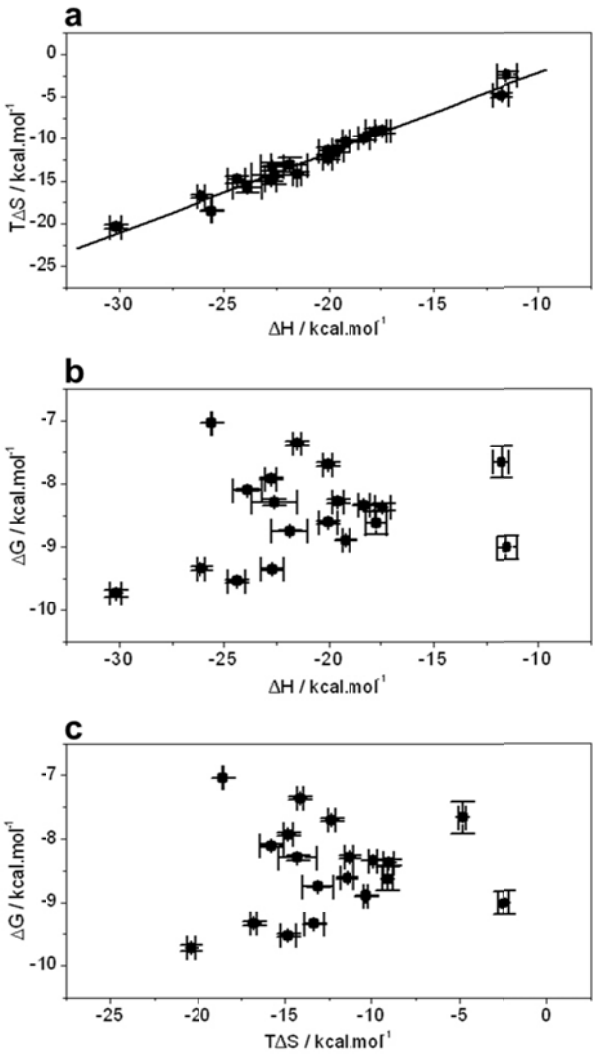

Figure 4-4: Interdependence of enthalpic (ΔH) and entropic (TΔS) contributions to the free energy (ΔG) for the binding of DB domain of ERα to dsDNA oligos containing the consensus ERE and variant motifs. (a) TΔS vs ΔH plot. (b) ΔG vs ΔH plot. (c) ΔG vs TΔS plot. Note that the solid line in (a) represents linear fit to the data. Error bars were calculated from at least three independent measurements. All errors are given to one standard deviation.

the rationale design of novel drugs in that attempts to improve the efficacy of drugs through optimizing drug-target interactions may prove futile due to enthalpy-entropy compensation. Nevertheless, a better understanding of protein-ligand interactions in thermodynamic terms is a pre-requisite for the development of thermodynamic rules that may ultimately guide the design of novel drugs harboring greater efficacy coupled with low toxicity. It should also be noted here that although optimization of favorable ΔH in

drug design appears to be more intuitive than minimizing unfavorable TΔS, enthalpically optimized drugs may not necessarily be effective universally and that entropically-optimized drugs may offer a viable alternative.

4.4.3 Effect of genetic variations within the ERE motif on the binding of ERα is governed by both the chemical nature of the substituted nucleotide and position of substitution

To test how the chemistry of the substituted nucleotide is coupled to the position at which it is substituted within the ERE motif in dictating the binding of DB domain of ERα, we generated the plots shown in Figure 4-5. It is important to note that there are four possible symmetric pairs of nucleotide substitutions within the ERE motif. These include the –A/+T, –C/+G, –G/+C and –T/+A symmetric pairs, where the – and + signs respectively indicate symmetric substitution of corresponding nucleotides within the left and right half-sites of the sense strand at a given position. Our data suggest that the introduction of these symmetric pairs of nucleotide substitutions within the ERE motif is highly position-dependent and that substitutions at certain positions are less tolerable than others. Thus, for example, the substitution of –A/+T and –G/+C symmetric pairs are least tolerable at ±2 positions, while –C/+G and –T/+A symmetric pairs are least tolerable at ±5 and ±4 positions, respectively. In contrast, the ±1, ±3 and ±6 positions within the ERE motif seem to be most tolerable for nucleotide substitutions. These salient observations are consistent with the crystal structure of the DB domain of ERα in complex with ERE duplex [*66*], wherein nucleotides at the ±2, ±4 and ±5 positions engage in closer intermolecular contacts with the protein in comparison with those at the ±1, ±3 and ±6 positions. It is noteworthy that the optimal contribution of underlying enthalpic and

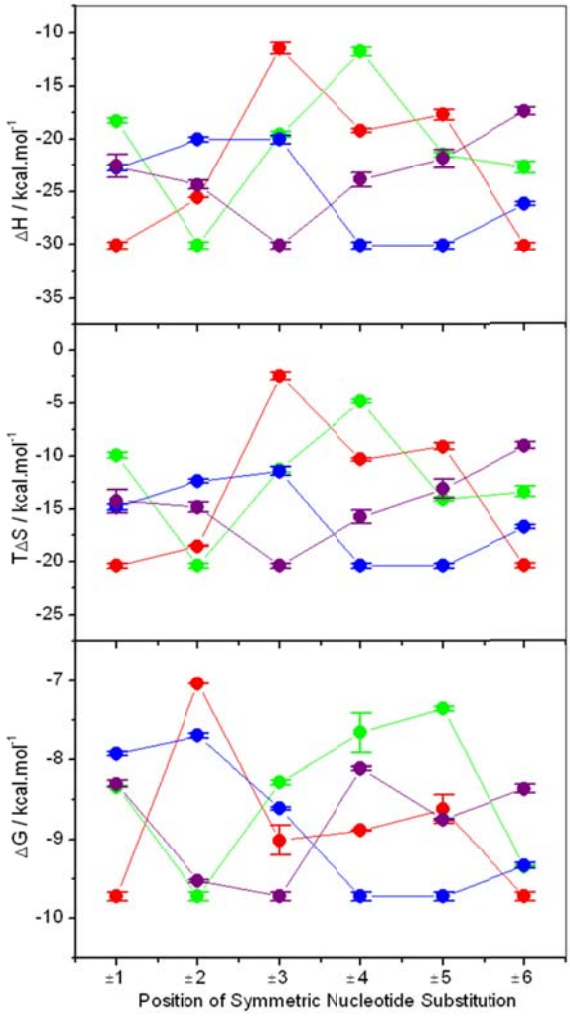

Figure 4-5: Dependence of free energy (ΔG) and the underlying enthalpic (ΔH) and entropic (TΔS) contributions on the position of symmetric nucleotide substitution within both half-sites of the consensus ERE and variant motifs for the binding of DB domain of ERα. The changes in various thermodynamic parameters upon the introduction of –A/+T (red), –C/+G (green), –G/+C (blue) and –T/+A (purple) are color-coded and connected by solid lines for clarity, where the – and + signs respectively indicate symmetric substitution of corresponding nucleotides within the left and right half-sites of the sense strand at the specified position. Error bars were calculated from at least three independent measurements. All errors are given to one standard deviation.

entropic contributions to the overall free energy does not correlate with the least and most tolerable positions within the ERE motif. Thus, for example, the substitution of –A/+T symmetric pair is least tolerable at ±2 position, but the least favorable enthalpic contribution for this pair occurs at ±3 position. However, the ±3 position is also the most preferred position for the –A/+T symmetric pair in terms of encountering the least

entropic penalty, implying that this position is reasonably tolerant for the substitution of –A/+T symmetric pair despite a significant loss in favorable enthalpic contribution. Of particular interest is also the observation that the most preferred positions for the –A/+T, –C/+G, –G/+C and –T/+A symmetric pairs are consistent with their positions within the consensus ERE motif, which displays the highest affinity for the DB domain compared to all other variants (Table 4-1). In other words, the nucleotides within the consensus AGGTCAnnnTGACCT motif are energetically the most preferred in their corresponding positions. Furthermore, the nucleotides within the consensus AGGTCAnnnTGACCT motif are also the most preferred in their corresponding positions in terms of enthalpic contributions to the free energy, implying that the consensus ERE motif is enthalpically-optimized. But is it also entropically-optimized? Close scrutiny of data presented in Figure 4-5 reveals that the nucleotides with least unfavorable entropic contributions to the overall free energy at each position within the ERE motif conform to the TACAGCnnnGCTGTA motif, implying that none of the nucleotides within the consensus ERE motif are entropically-optimized. In short, our data suggest strongly that the effect of genetic variations within the ERE motif on the energetics of binding of DB domain of ERα are strongly dependent upon both the chemistry of the substituted nucleotide and the position at which it is substituted.

4.4.4 ERα shows no preference for binding to ERE variants rich in AT content

It is widely believed that AT sequences within DNA account for its intrinsic conformational flexibility such as bending and curvature [*132, 133, 149-154*]. Such intrinsic propensity of AT sequences to undergo bending is believed to be largely due to increased propeller twist of these sequences by virtue of the fact that A-T base pairs are

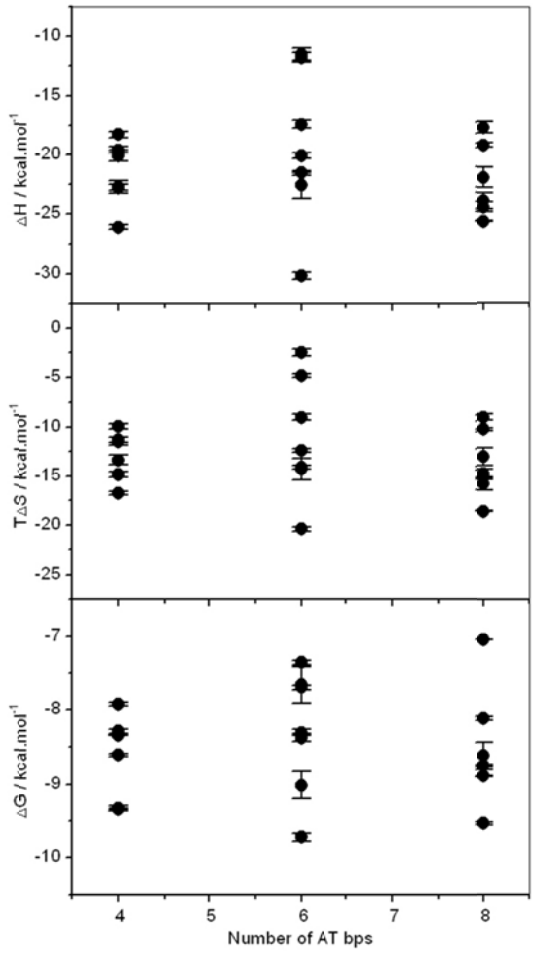

Figure 4-6: Dependence of free energy (ΔG) and the underlying enthalpic (ΔH) and entropic (TΔS) contributions on the total number of AT bps within both half-sites of the consensus ERE and variant motifs for the binding of DB domain of ERα.

held together by only two hydrogen bonds in lieu of three formed between G-C base pairs. Given that the binding of the DB domain of ERα to the ERE motif results in the bending of DNA [*139*], the AT content of ERE motif may therefore play an important role in modulating this protein-DNA interaction. In an attempt to analyze how the AT content of the ERE motif and its variants thereof correlates with the energetics of binding of the DB domain of ERα, we generated thermodynamic plots shown in Figure 4-6. To our surprise, the increase in AT content of ERE variants neither correlates with an increase in favorable enthalpic contribution nor unfavorable entropic contribution with

the net result that the increased DNA flexibility does not result in enhanced binding to ERα. Thus, although the conformational plasticity such as the ability of ERE variants to bend and wrap around the DB domain so as to attain a close molecular fit may be critical for high-affinity binding, the unfavorable entropy arising from DNA becoming more constrained upon the binding of more flexible AT-rich variants relative to more rigid GC-rich variants may override such conformational advantage.

4.4.5 Genetic variations within the ERE motif allow it to sample much greater conformational space

As discussed above, the nucleotide sequence is a key determinant of the ability of DNA to behave as a flexible polymer and undergo physical phenomena such as bending and curvature [*132-134*]. In an effort to further test how the introduction of genetic variations within the ERE motif modulates its conformational flexibility, we next conducted CD analysis (Figure 4-7). As expected, the CD spectra of ERE motifs and its variants thereof exhibit features characteristic of a right-handed double-stranded B-DNA with bands centered around 195nm, 220nm and 280nm. It is important to note that while the 195-nm and 220-nm bands arise from secondary structural DNA features, the 280-nm band probes the 3D conformation of DNA and therefore it is highly sensitive to physical changes in DNA such as bending and curvature. Accordingly, it is evident that the CD spectra of the variant motifs do not superimpose upon the CD spectrum of the consensus ERE motif but rather fluctuate around it and fan out forming a cluster of closely related optical spectra. Additionally, the wavelength maxima of the spectral bands centered around 220nm and 280nm show remarkable heterogeneity and lie as much as more than 10nm apart for some variant motifs relative to the consensus ERE motif. These salient observations imply that the introduction of single nucleotide substitutions within the ERE

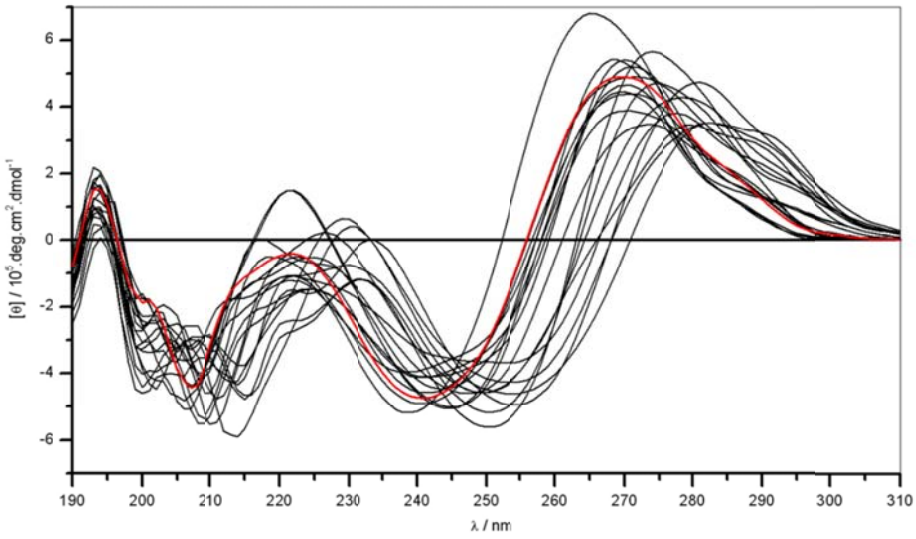

Figure 4-7: Representative CD spectra of dsDNA oligos containing the consensus ERE and variant motifs. The spectrum of consensus ERE motif (red) is superimposed onto spectra of variant motifs (black).

motif tightly governs its conformational flexibility and that the varying flexibility is likely to be an integral feature of their ability to bind to the DB domain of ERα with distinct underlying energetics (Table 4-1). We also note that while the wavelength maximum and the intensity of the 280-nm band are related to the overall 3D conformation of DNA, such optical properties are not easily interpretable in structural terms such as bending and curvature. Nonetheless, our CD data indicate that the introduction of genetic variations within the ERE motif allows it to sample a much greater conformational space that might be a key feature of the ability of its variants to bind to ERα at distinct promoters in a selective manner.

4.4.6 Atomic models provide the physical basis of how genetic variations within the ERE may gauge its binding affinity toward ERα

Our thermodynamic analysis presented here suggests strongly that genetic variations within the ERE motif can modulate its affinity to the DB domain of ERα by orders of magnitude. In an attempt to rationalize such a broad spectrum of binding affinities, we modeled and compared 3D structures of the DB domain in complex with

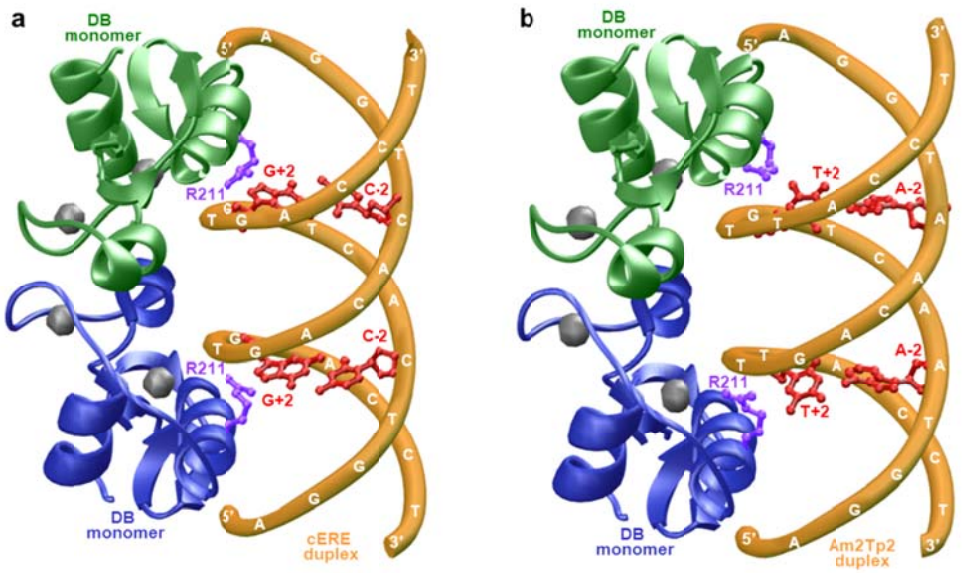

Figure 4-8: 3D atomic models of the DB domain of ERα in complex with dsDNA oligos containing the consensus ERE motif (a) and the variant Am2Tp2 motif (b). Note that the DB domain binds to DNA as a homodimer with a two-fold axis of symmetry. One monomer of the DB domain is shown in green and the other in blue. The Zn^{2+} divalent ions within each DB monomer are depicted as gray spheres and the sidechain moieties of R211 within each DB monomer are colored purple. The DNA backbone is shown in yellow and the bases at –2 and +2 positions within each motif are colored red.

the consensus ERE motif and the Am2Tp2 variant motif (Figure 4-8). It should be noted that the Am2Tp2 motif binds to the DB domain by nearly two orders of magnitude weaker than the consensus ERE motif (Table 4-1). Our modeled structures provide an exquisite explanation for such dramatic differences in the binding affinities of these two motifs. As shown in the crystal structure [*66*], the DB domain binds as a homodimer with a two-fold axis of symmetry within the major grooves of DNA duplex. One monomer contacts the antisense strand within the first half-site of the DNA duplex, while the other monomer contacts the sense strand within the second half-site. The key amino acid residue within each monomer involved in making differential contacts with the consensus ERE motif relative to the Am2Tp2 motif is R211. In the case of consensus ERE duplex, the guanidino moiety of R211 hydrogen bonds to a nitrogen atom with a lone pair of electrons in the G+2 base within both the antisense and sense strands. In the Am2Tp2

motif, the G+2 base is substituted to T+2 base. Of the all four DNA bases, thymine stands out for its unique hydrophobic character due to the presence of a methyl group. Thus, the introduction of a thymine at +2 positions within both strands of Am2Tp2 not only imparts hydrophobicity but, unlike guanine at this position, thymine can no longer serve as a hydrogen bonding partner for the guanidino moiety of R211 due to the lack of an available nitrogen atom with a lone pair of electrons. Additionally, our 3D atomic models reveal that the hydrophobic methyl group of T+2 bases would be highly destabilizing for subsequent protein-DNA contacts due to its close proximity to the charged guanidino moiety of G+2 bases. Accordingly, the sidechain of R211 must undergo a rotation to minimize contact with the methyl group of T+2 bases upon the binding of DB domain to Am2Tp2 motif. Additionally, the small size of thymine base compared to a much bulkier guanine may also result in the formation of cavities and subsequent entrapment of water molecules at the protein-DNA interface. In sum, our 3D atomic models suggest that the loss of key hydrogen bonding contacts at two critical points coupled with a number of other destabilizing factors could significantly weaken the binding of DB domain of ERα to Am2Tp2 motif relative to consensus ERE motif in agreement with our thermodynamic data reported here.

4.5 Concluding remarks

The ability of ERα to serve as a transcription factor is largely dependent upon its ability to recognize the promoters of target genes. Although it is generally believed that ERα recognizes the ERE motif containing the consensus AGGTCAnnnTGACCT sequence, the promoters of a vast majority of estrogen-responsive genes are comprised of unusual elements that are related to the consensus ERE motif but differ in one or more

nucleotides [*65*]. Despite the knowledge that such genetic variations within the promoters of target genes modulate the transcriptional activity of nuclear receptors [*72-75*], deciphering the underlying protein-DNA interactions in quantitative terms has remained a daring challenge for the past two decades or so. Herein, we have provided a detailed ITC analysis of how genetic variations within the ERE motif may affect the binding of ERα nuclear receptor and hence its transcriptional output in response to estrogens. Our data suggest strongly that genetic variations can modulate the binding of ERα by orders of magnitude and that such modulation may or may not involve drastic changes in the contribution of underlying thermodynamic forces driving subtle protein-DNA interactions. Furthermore, the binding of ERE motif and its variants thereof to ERα faithfully obeys the enthalpy-entropy compensation phenomenon, arguing strongly that thermodynamic considerations should form an integral part of rational drug design. Although it is widely believed that AT sequences within DNA account for its intrinsic conformational flexibility such as bending and curvature [*132, 133, 149-154*], our data presented here suggest a poor correlation between DNA flexibility and its binding to ERα. To account for such a discrepancy, we argue that although the conformational plasticity such as the ability of ERE variants to bend and wrap around ERα so as to attain a close molecular fit may be critical for high-affinity binding, the unfavorable entropy arising from DNA becoming more constrained upon the binding of more flexible AT-rich variants relative to more rigid GC-rich variants may override such conformational advantage. Nevertheless, our CD analysis shows that the introduction of genetic variations within the ERE motif allows it to sample a much greater conformational space that might be a key feature of the ability of its variants to bind to ERα at distinct

promoters in a selective manner. Our atomic models also provide structural basis of how the symmetrical introduction of A-2/T+2 nucleotide pair within both half-sites of ERE can result in the reduction of binding of DB domain by nearly two orders of magnitude. Likewise, we postulate that the introduction of nucleotides at other positions within the ERE motif is likely to result in the loss of hydrogen bonding and other stabilizing interactions due to the rearrangement of amino acid sidechains in the DB domain. Finally, it should be noted that the genetic variations within the ERE motif may not necessarily act alone but rather in concert with other factors to regulate the transcriptional activity of ERα within the milieu of the cell. It has been previously reported that the nucleotides flanking the ERE motif affect the transcription activity of ERα [*155-158*]. Additionally, many estrogen-responsive genes contain the ERE motif in tandem or as composite elements containing an ERE motif and the binding site for another transcription factor. All these scenarios thus could dictate how genetic variations within the ERE motif at a given promoter might influence the overall transcriptional activity of ERα. Taken together, our study bears important consequences for how genetic variations within DNA promoter elements may fine-tune the physiological action of ERα and other nuclear receptors.

Chapter 5: Structural and Thermodynamic Consequences of the Replacement of Zinc with Environmental Metals on ERα-DNA Interactions

5.1 Summary

Estrogen receptor α (ERα) acts as a transcription factor by virtue of the ability of its DNA-binding (DB) domain, comprised of a tandem pair of zinc fingers, to recognize the estrogen response element (ERE) within the promoters of target genes. Herein, using an array of biophysical methods, we probe structural consequences of the replacement of zinc within the DB domain of ERα with various environmental metals and their effects on the thermodynamics of binding to DNA. Our data reveal that while the DB domain reconstituted with divalent ions of zinc, cadmium, mercury and cobalt binds to DNA with affinities in the nanomolar range, divalent ions of barium, copper, iron, lead, manganese, nickel and tin are unable to regenerate DB domain with DNA-binding potential though they can compete with zinc for coordinating the cysteine ligands within the zinc fingers. We also show that the metal-free DB domain is a homodimer in solution and that the binding of various metals only results in subtle secondary and tertiary structural changes, implying that metal-coordination may only be essential for DNA-binding. Collectively, our findings provide mechanistic insights into how environmental metals may modulate the physiological function of a key nuclear receptor involved in mediating a plethora of cellular functions central to human health and disease.

5.2 Overview

Estrogen receptor α (ERα) is a member of a family of ligand-modulated transcription factors that have come to be known as nuclear receptors (NRs) [*3-6*]. ERα mediates the action of estrogens such as estradiol in a diverse array of cellular processes and its hyperactivation leads to the genesis of large fractions of breast cancer [*8*]. ERα is

constructed on a modular architecture, also shared by other members of the NR family, comprised of a central DNA-binding (DB) domain flanked between an N-terminal trans-activation (TA) domain and a C-terminal ligand-binding (LB) domain. Upon the binding of estrogens to the LB domain, ERα translocates to the nucleus and binds as a homodimer with a two-fold axis of symmetry to the estrogen response element (ERE), containing the AGGTCAnnnTGACCT consensus sequence, located within the promoters of target genes [*65*]. DNA-binding is accomplished through a pair of tandem C4-type Zinc fingers located within the DB domain, with each finger containing a Zn^{2+} ion coordinated in a tetrahedral arrangement by four highly conserved cysteine residues to generate the $Zn^{2+}[Cys]_4$ metal-protein complex [*66, 67*] (Figure 5-1). Importantly, while the first Zinc finger (ZF-I) within each monomer of DB domain recognizes the hexanucleotide sequence 5'-AGGTCA-3' within the major groove at each end of the ERE duplex, the second Zinc finger (ZF-II) is responsible for the homodimerization of DB domain.

Discovered more than a quarter of century ago [*159*], the zinc finger is one of the most common motifs found in transcription factors [*160, 161*]. Several lines of evidence suggest that metals other than zinc can serve as coordination sites for cysteine ligands within zinc fingers with important consequences on cellular processes involved in gene expression, DNA repair and genomic stability [*77, 89, 90, 162-165*]. In an effort to further our understanding of the interaction of metals with zinc fingers, we analyze here structural consequences of the replacement of zinc within the DB domain of ERα with various environmental metals and their effects on the thermodynamics of binding to DNA using an array of biophysical methods. Our data reveal that while the DB domain

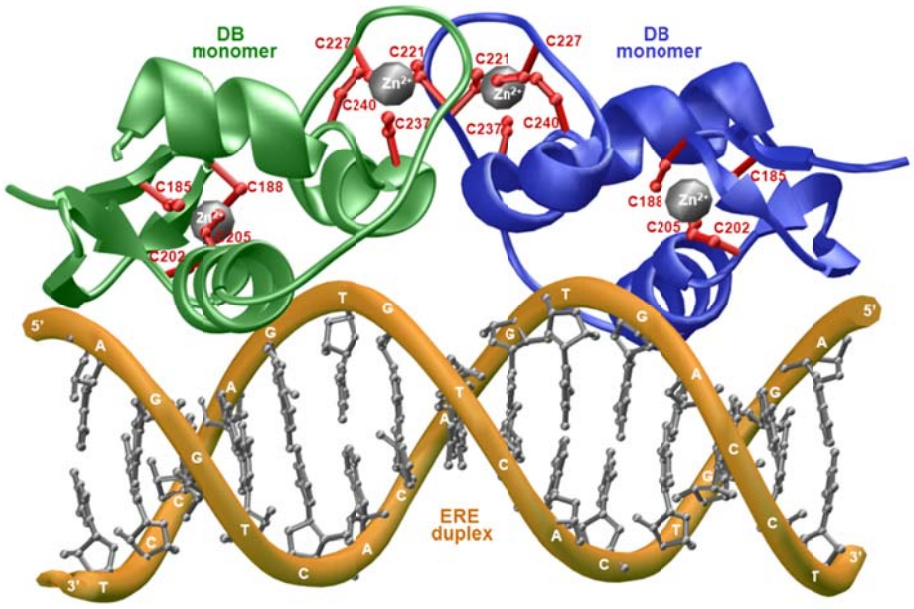

Figure 5-1: 3D structural model of the DB domain of human ERα in complex with ERE duplex containing the AGGTCAcagTGACCT consensus sequence based on the crystal structure (PDB# 1HCQ) determined by Rhodes and co-workers [*66*]. The structural model was built as described earlier [*135*] and rendered using RIBBONS [*101*]. Note that the DB domain binds to DNA as a homodimer with a two-fold axis of symmetry. One monomer of DB domain is shown in green and the other in blue. The Zn^{2+} divalent ions are depicted as gray spheres and the sidechain moieties of cysteine residues being coordinated in red. The DNA backbone is shown in yellow and the bases are colored gray for clarity.

reconstituted with divalent ions of zinc, cadmium, mercury and cobalt binds to DNA with affinities in the nanomolar range, divalent ions of barium, copper, iron, lead, manganese, nickel and tin are unable to regenerate DB domain with DNA-binding potential though they can compete with zinc for coordinating the cysteine ligands within the zinc fingers. We also show that the metal-free DB domain is a homodimer in solution and that the binding of various metals only results in subtle secondary and tertiary structural changes, implying that metal-coordination may only be essential for DNA-binding. Collectively, our findings provide mechanistic insights into how environmental metals may modulate the physiological function of a key nuclear receptor involved in mediating a plethora of cellular functions central to human health and disease.

5.3 Experimental Procedures

5.3.1 Protein Preparation

The DB domain (residues 176-250) of human ERα was cloned into pET101 bacterial expression vector with a C-terminal polyhistidine (His)-tag using Invitrogen TOPO technology. The recombinant protein was expressed in bacteria supplemented with 50μM $ZnCl_2$ and purified to apparent homogeneity on a Ni-NTA affinity column followed by treatment on a Hiload Superdex 200 size-exclusion chromatography (SEC) column coupled in-line with GE Akta FPLC system as described previously [*135, 166*]. Zinc divalent ions were stripped by the treatment of purified protein in Tris buffer (50mM Tris, 200mM NaCl and 10mM β-mercaptoethanol at pH 8.0) containing 8M urea and 10mM EDTA. After denaturation of protein overnight, EDTA was removed under denatured conditions by dialysis in acetate buffer (50mM Sodium acetate, 200mM NaCl and 10mM β-mercaptoethanol at pH 6.0) containing 8M urea. Further dialysis of protein in acetate buffer containing one of the metal chlorides at a final protein-to-metal molar ratio of 1:10 led to simultaneous removal of urea and reconstitution of the DB domain with the corresponding metal divalent ions. It is important to note that the reconstitution of DB domain at pH 6.0 was necessary to prevent the formation of insoluble salts such as lead chloride. Additionally, acetate buffer was preferred over phosphate buffer due to insolubility of various metal phosphates. Metal-reconstituted DB domain was extensively dialyzed in an appropriate buffer to remove excess metal ions. Protein concentration was determined by the fluorescence-based Quant-It assay (Invitrogen) and spectrophotometrically using an extinction coefficient of 14,940 $M^{-1}cm^{-1}$. Results from both assays were in an excellent agreement.

5.3.2 DNA synthesis

21-mer DNA oligos containing the ERE consensus site AGGTCAnnnTGACCT were commercially obtained from Sigma Genosys. The complete nucleotide sequences of the sense and antisense oligos constituting the ERE duplex is presented below:

5'-cccAGGTCAcagTGACCTgcg-3'

3'-gggTCCAGTgtcACTGGAcgc-5'

Oligo concentrations were determined spectrophotometrically on the basis of their extinction co-efficients derived from their nucleotide sequences. Sense and antisense oligos were annealed together to generate the ERE duplex as described earlier [*135, 166*].

5.3.3 ITC Measurements

Isothermal titration calorimetry (ITC) measurements were performed on a Microcal VP-ITC instrument and data were acquired and processed using fully automized features in Microcal ORIGIN. All measurements were repeated at least three times. Briefly, protein and DNA samples were prepared in 50mM Sodium phosphate buffer containing 5mM β-mercaptoethanol at pH 7.0 and de-gassed using the ThermoVac accessory for 5min. The experiments were initiated by injecting 25 x 10μl aliquots of 50-100μM of ERE duplex from the syringe into the calorimetric cell containing 1.8ml of 2-5μM of DB domain of ERα reconstituted with various metals at 25°C. The change in thermal power as a function of each injection was automatically recorded using the ORIGIN software and the raw data were further processed to yield binding isotherms of heat release per injection as a function of molar ratio of ERE duplex to dimer-equivalent DB domain. 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 ERE duplex in an identical manner. Control experiments with scrambled dsDNA oligos generated similar thermal power to that obtained for the buffer alone, implying that there was no non-specific binding of DB domains to non-cognate DNA. To extract various thermodynamic parameters, the binding isotherms were iteratively fit to a built-in one-site model by non-linear least squares regression analysis using the ORIGIN software as described previously [*114, 135*].

5.3.4 ALS measurements

Analytical light scattering (ALS) experiments were conducted on a Wyatt miniDAWN TREOS triple-angle static light scattering detector and Wyatt QELS dynamic light scattering detector coupled in-line with a Wyatt Optilab rEX differential refractive index detector and interfaced to a Hiload Superdex 200 size-exclusion chromatography column under the control of a GE Akta FPLC system within a chromatography refrigerator at 10°C. The DB domain of ERα pre-treated with EDTA to strip divalent zinc ions and upon reconstitution with divalent ions of various metals was loaded onto the column at a flow rate of 1ml/min and the data were automatically acquired using the ASTRA software. All protein samples were prepared in 50mM Sodium phosphate buffer containing 5mM β-mercaptoethanol at pH 7.0 and the starting concentrations injected onto the column were between 20-50μM. The angular- and concentration-dependence of static light scattering (SLS) intensity of each protein species resolved in the flow mode was measured by the Wyatt miniDAWN TREOS detector. The SLS data were analyzed according to the following built-in Zimm equation in ASTRA software [*84, 85*]:

$$(Kc/R_\theta) = ((1/M_{obs})+2A_2c)[1+((16\pi^2(R_g)^2/3\lambda^2)\sin^2(\theta/2))] \qquad [5\text{-}1]$$

where R_θ is the excess Raleigh ratio due to protein in the solution as a function of protein concentration c (mg/ml) and the scattering angle θ (42°, 90° and 138°), M_{obs} is the observed molecular mass of each protein species, A_2 is the second virial coefficient, λ is the wavelength of laser light in solution (658nm), R_g is the radius of gyration of protein, and K is given by the following relationship:

$$K = [4\pi^2 n^2 (dn/dc)^2]/N_A\lambda^4 \quad [5\text{-}2]$$

where n is the refractive index of the solvent, dn/dc is the refractive index increment of the protein in solution and N_A is the Avogadro's number ($6.02x10^{23}mol^{-1}$). If we assume that $c \rightarrow 0$ and $\theta \rightarrow 0$, then Eq [1] reduces to:

$$(Kc/R_\theta) = 1/M_{obs} \quad [5\text{-}3]$$

Thus, under dilute protein concentrations ($c \rightarrow 0$) and at low scattering angles ($\theta \rightarrow 0$), the y-intercept of Eq [1] equates to $1/M_{obs}$. Accordingly, the weighted average value for M_{obs} was obtained from the y-intercept of linear fits of a range of (Kc/R_θ)-$\sin^2(\theta/2)$ plots as a function of protein concentration along the elution profile of each protein species using SLS measurements at three scattering angles. The time- and concentration-dependence of dynamic light scattering (DLS) intensity fluctuation of each protein species resolved in the flow mode was measured by the Wyatt QELS detector positioned at 90° with respect to the incident laser beam. The DLS data were iteratively fit using non-linear least squares regression analysis to the following built-in equation in ASTRA software: software [*86-88*]:

$$G(\tau) = \alpha Exp(-2\Gamma\tau) + \beta \quad [5\text{-}4]$$

where $G(\tau)$ is the autocorrelation function of dynamic light scattering intensity fluctuation I, τ is the delay time of autocorrelation function, Γ is the decay rate constant

of autocorrelation function, α is the initial amplitude of autocorrelation function at zero delay time, and β is the baseline offset (the value of autocorrelation function at infinite delay time). Thus, fitting the above equation to a range of $G(\tau)$-τ plots as a function of protein concentration along the elution profile of each protein species computes the average value of Γ using DLS measurements at a scattering angle of 90°. Accordingly, the translational diffusion coefficient (D_t) of each protein species was calculated from the following relationship:

$$D_t = [(\Gamma\lambda^2)/(16\pi^2n^2\sin^2(\theta/2))] \quad [5\text{-}5]$$

where λ is the wavelength of laser light in solution (658nm), n is the refractive index of the solvent and θ is the scattering angle (90°). Additionally, the hydrodynamic radius (R_h) of each protein species was calculated from the Stokes-Einstein relationship:

$$R_h = [(k_BT)/(6\pi\eta D_t)] \quad [5\text{-}6]$$

where k_B is Boltzman's constant ($1.38\text{x}10^{-23}JK^{-1}$), T is the absolute temperature and η is the solvent viscosity. It should be noted that, in both the SLS and DLS measurements, protein concentration (c) along the elution profile of each protein species was automatically quantified in the ASTRA software from the change in refractive index (Δn) with respect to the solvent as measured by the Wyatt Optilab rEX detector using the following relationship:

$$c = (\Delta n)/(dn/dc) \quad [5\text{-}7]$$

where dn/dc is the refractive index increment of the protein in solution.

5.3.5 CD analysis

Circular dichroism (CD) analysis was conducted on a Bio-Logic MOS450/SFM400 spectropolarimeter thermostatically controlled with a water bath at

25°C and data were acquired using the BIOKINE software. Briefly, experiments were conducted on a 20-50μM of DB domain of ERα pre-treated with EDTA to strip divalent zinc ions and upon reconstitution with divalent ions of various metals in 50mM Sodium phosphate buffer at pH 7.0. For far-UV measurements in the 190-250nm wavelength range, experiments were conducted in a quartz cuvette with a 2-mm pathlength. For near-UV measurements in the 250-350nm wavelength range, experiments were conducted in a quartz cuvette with a 10-mm pathlength. All data were recorded with a slit bandwidth of 2nm at a scan rate of 3nm/min. Data were normalized against reference spectra to remove the contribution of buffer. The reference spectra were obtained in a similar manner on a 50mM Sodium phosphate buffer at pH 7.0. Each data set represents an average of at least four scans acquired at 1nm intervals. Data were converted to molar ellipticity, [θ], as a function of wavelength (λ) of electromagnetic radiation using the following equation:

$$[\theta] = [(10^5 \Delta\varepsilon)/cl] \text{ deg.cm}^2\text{.dmol}^{-1} \qquad [5\text{-}8]$$

where Δε is the observed ellipticity in mdeg, c is the protein concentration in μM and l is the cuvette pathlength in cm. All data were processed and analyzed using the Microcal ORIGIN software.

5.3.6 SSA measurements

Steady-state absorbance (SSA) spectra were collected on a Jasco V-630 spectrophotometer using a quartz cuvette with a 10-mm pathlength at 25 °C. Briefly, experiments were conducted on 5-10μM of DB domain of ERα pre-treated with EDTA to strip divalent zinc ions and upon reconstitution with divalent ions of various metals in 50mM Sodium phosphate buffer at pH 7.0. All data were recorded in the 200-350nm wavelength range using a 1.5-nm slit bandwidth. Data were normalized against reference

spectra to remove the contribution of buffer. The reference spectra were obtained in a similar manner on a 50mM Sodium phosphate buffer at pH 7.0.

5.3.7 SSF measurements

Steady-state fluorescence (SSF) spectra were collected on a Jasco FP-6300 spectrofluorometer using a quartz cuvette with a 10-mm pathlength at 25 °C. Briefly, experiments were conducted on 5-10μM of DB domain of ERα pre-treated with EDTA to strip divalent zinc ions and upon reconstitution with divalent ions of various metals in 50mM Sodium phosphate buffer at pH 7.0. Excitation wavelength was 295nm and emission was acquired from 310nm to 500nm. All data were recorded using a 2.5-nm bandwidth for both excitation and emission. Data were normalized against reference spectra to remove the contribution of buffer. The reference spectra were obtained in a similar manner on a 50mM Sodium phosphate buffer at pH 7.0.

5.4 Results and discussion

5.4.1 Binding of the DB domain of ERα to DNA is restored upon substitution of zinc with only specific divalent metal ions

To determine the extent to which environmental metals may be able to replace zinc within the zinc fingers of ERα, we measured the binding of ERE duplex to the DB domain pre-treated with EDTA to remove divalent zinc ions (as a control) and upon reconstitution with divalent ions of various metals using ITC (Figure 5-2). Our data reveal that while the DB domain reconstituted with divalent ions of zinc, cadmium, mercury and cobalt binds to DNA with affinities in the nanomolar range (Table 5-1), divalent ions of barium, copper, iron, lead, manganese, nickel and tin are unable to regenerate DB domain with DNA-binding potential. These data are consistent with

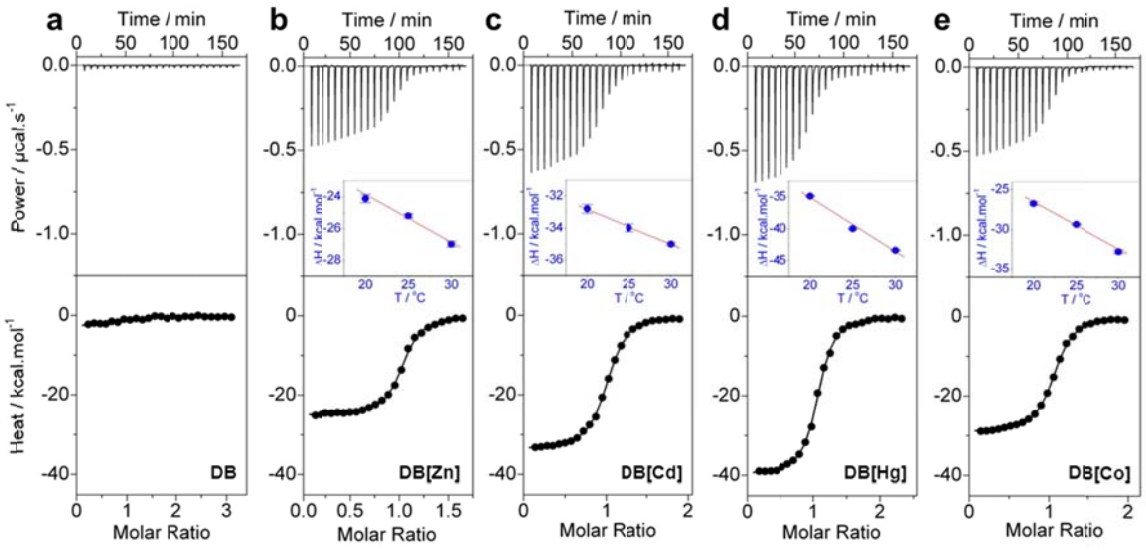

Figure 5-2: Representative ITC isotherms for the binding of ERE duplex to the DB domain of ERα pre-treated with EDTA to strip divalent zinc ions (a) and upon reconstitution with divalent ions of zinc (b), cadmium (c), mercury (d) and cobalt (e). The upper panels show the raw ITC data expressed as change in thermal power with respect to time over the period of titration. In the lower panels, change in molar heat is expressed as a function of molar ratio of ERE duplex to dimer-equivalent DB domain. The solid lines in the lower panels represent the fit of data to a one-site model, based on the binding of a ligand to a macromolecule assuming the law of mass action, using the ORIGIN software [*114, 135*]. The insets in top panels represent ΔH-T plots for the corresponding protein-DNA complexes. The solid line in red represents linear fit to the corresponding data.

previous studies demonstrating that DB domain of ERα regenerated with divalent ions of zinc, cadmium and cobalt binds DNA but not that regenerated with divalent ions of copper and nickel [*78, 79*].

Although the replacement of zinc with cadmium, mercury and cobalt restores DNA-binding with very similar affinities, the underlying thermodynamic forces display remarkable contrast (Table 5-1). Thus, while the binding of DB domain reconstituted with various metals is universally driven by favorable enthalpic changes accompanied by entropic penalty, the binding of DB domain reconstituted with cadmium and mercury results in the release of 5-15 kcal/mol of additional enthalpic contribution to the overall free energy relative to reconstitution with zinc and cobalt. We believe that such enthalpic advantage is due to the fact that the DB domain reconstituted with cadmium and mercury

Table 5-1

Thermodynamic parameters obtained from ITC measurements at 25°C for the binding of ERE duplex to the DB domain of ERα pre-treated with EDTA to strip divalent zinc ions (DB) and upon reconstitution with divalent ions of zinc (DB[Zn]), cadmium (DB[Cd]), mercury (DB[Hg]) and cobalt (DB[Co])

	K_d / nM	ΔH / kcal.mol^{-1}	$T\Delta S$ / kcal.mol^{-1}	ΔG / kcal.mol^{-1}	ΔC_p / kcal.mol^{-1}.K^{-1}
DB	NB	NB	NB	NB	NB
DB[Zn]	68 ± 8	-25.20 ± 0.15	-15.40 ± 0.21	-9.80 ± 0.07	-0.28 ± 0.04
DB[Cd]	69 ± 5	-33.99 ± 0.25	-24.20 ± 0.20	-9.78 ± 0.04	-0.22 ± 0.04
DB[Hg]	59 ± 4	-39.95 ± 0.13	-30.07 ± 0.16	-9.88 ± 0.04	-0.85 ± 0.01
DB[Co]	81 ± 2	-29.35 ± 0.10	-19.66 ± 0.10	-9.69 ± 0.01	-0.60 ± 0.01

The values for the affinity (K_d) and enthalpy change (ΔH) accompanying the binding of ERE duplex to the DB domain reconstituted with various metals were obtained from the fit of a one-site model, based on the binding of a ligand to a macromolecule using the law of mass action, to the corresponding ITC isotherms as described earlier [*114, 135*]. 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 ($T\Delta S$) to binding was calculated from the relationship $T\Delta S=\Delta H-\Delta G$. Heat capacity change ($\Delta Cp$) was calculated from the slope of ΔH-T plot for the corresponding protein-DNA complex (Figure 2). Binding stoichiometries generally agreed to within ±10%. Errors were calculated from at least three independent measurements. All errors are given to one standard deviation. Note that the DB domain pre-treated with EDTA to strip divalent zinc ions and upon reconstitution with divalent metal ions of barium, copper, iron, lead, manganese, nickel and tin showed no binding (NB) to the ERE duplex.

is only partially structured and that it only becomes fully structured upon binding to DNA in a binding-coupled-folding manner. To shed further light on this phenomenon, we also measured heat capacity changes associated with the binding of various metal-coordinated DB domains to DNA (Figure 5-2 and Table 5-1). It has been previously reported that a large negative heat capacity change (ΔC_p) associated with protein-DNA interactions is indicative of protein folding coupled to DNA-binding [*167*]. Remarkably, our analysis reveals that while a large negative ΔC_p indeed accompanies the binding of mercury-coordinated DB domain to DNA, ΔC_p associated with the binding of cadmium-coordinated DB domain is smaller than that observed for DB domain reconstituted with zinc and cobalt. In fact, ΔC_p accompanying the binding of cobalt-coordinated DB domain to DNA is similar to that observed for mercury-coordinated DB domain. However, it is

important to note that such discrepancies in the values of ΔC_p do not necessarily contradict the corresponding enthalpic contributions to the overall free energy of binding of various metal-coordinated DB domains to DNA. On the contrary, it is believed that factors other than protein folding can also contribute to ΔC_p. In particular, factors such as entrapment of water molecules and counterions within interfacial cavities as well as proton-linked equilibria during the formation of macromolecular complexes may also contribute to heat capacity changes [*120, 168*]. Thus, the large negative values of ΔC_p reported here may not solely reflect the folding of DB domain upon binding to DNA but they could also arise from other factors.

We have previously shown that the binding of DB domain of ERα to DNA is coupled to proton uptake [*135*]. Interestingly, this phenomenon is not affected by the replacement of zinc with divalent ions of cadmium, mercury and cobalt. As shown in Figure 5-3, the observed enthalpies for the binding of various metal-reconstituted DB domains to DNA are strongly dependent on buffer, implying that DNA-binding is coupled to proton uptake irrespective of the nature of metal coordination. Our analysis also suggests that the binding of metal-reconstituted DB domains to DNA involves a net uptake of two protons in agreement with our previous study [*135*].

Importantly, the metal-coordination of cysteine ligands in a tetrahedral arrangement to generate the $M^{2+}[Cys]_4$ metal-protein may be important for the proper folding of the DB domain such that it can recognize the ERE duplex in a specific manner. This is further corroborated by the knowledge that the divalent ions of zinc, cadmium, mercury and cobalt are all capable of coordinating their ligands with tetrahedral geometries [*169*]. However, the fact that divalent ions of metals such as nickel and

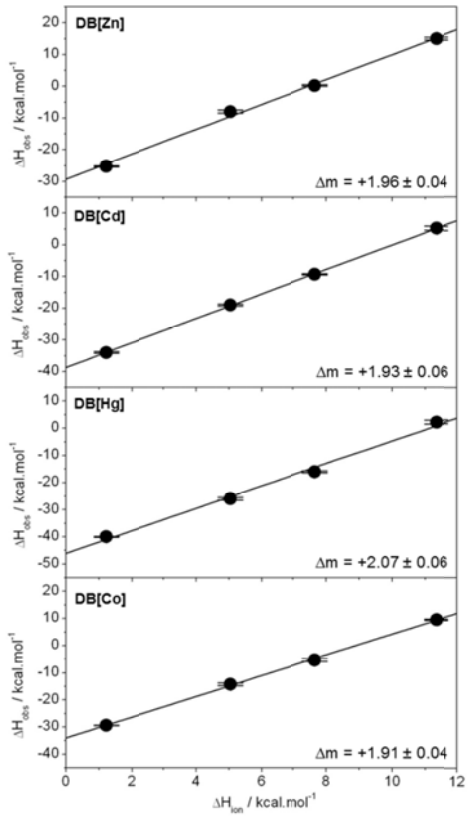

Figure 5-3: Dependence of observed enthalpy (ΔH_{obs}) as a function of ionization enthalpy (ΔH_{ion}) of various buffers upon the binding of ERE duplex to the DB domain of ERα reconstituted with divalent ions of zinc (DB[Zn]), cadmium (DB[Cd]), mercury (DB[Hg]) and cobalt (DB[Co]). The ΔH_{ion} of various buffers used were +1.22 kcal/mol (Phosphate), +5.02 kcal/mol (Hepes), +7.64 kcal/mol (Ticine) and +11.35 kcal/mol (Tris) [*115-117*]. The solid lines within each panel represent linear fits to data points and the net change in the number of protons (μm) uptaken per DB monomer upon binding to DNA was calculated from the corresponding slopes as described earlier [*135*]. Error bars were calculated from at least three independent measurements. All errors are given to one standard deviation.

manganese can also adopt tetrahedral geometry implies that the factors other than coordination geometry may also hold key to determining whether a particular divalent metal ion can replace zinc within the zinc fingers of the DB domain. Of particular importance are the ionic radii and internuclear coordinaton distances of various divalent metal ions in complex with their ligands. Interestingly, the ionic radii of hydrated divalent ions of all divalent metal ions analyzed here fall in the 100-150 pm range [*170*], and there seems to be no correlation between their ionic radii and their ability to replace zinc within the zinc fingers of the DB domain. To what extent the hydration shell, or the

extent to which a divalent metal ion becomes hydrated in solution, may be an important determinant of its ability to coordinate a given ligand also remains debatable. In short, it is not clear from our studies as to why some divalent ions can replace zinc within the zinc fingers of the DB domain while others cannot on basis of their physicochemical properties such as coordination geometry, ionic radius and internuclear coordination distance. Our study thus clearly warrants further investigation of the precise mechanisms driving protein-metal interactions.

5.4.2 Substitution of zinc with other divalent metal ions does not affect hydrodynamic properties of the DB domain of ERα

To further investigate how the substitution of zinc with cadmium, mercury and cobalt affects the hydrodynamic properties of the DB domain of ERα, we conducted ALS analysis based on first principles of hydrodynamics with no assumptions (Figure 5-4). It is generally believed that the DB domain is monomeric in solution and that it only homodimerizes upon binding to DNA [*66, 67*]. Contrary to this school of thought, our analysis reveals that the DB domain predominantly exists as a homodimer in solution even in the absence of DNA (Table 5-2). Analysis on a non-reducing SDS-PAGE further confirmed that the ability of DB domain to homodimerize in solution was not an artifact of intermolecular disulfide bridges.

Strikingly, the DB domain not only homodimerizes in the absence of DNA but even the metal-coordination does not appear to be obligatory. Thus, the metal-free DB domain not only appears to behave as a homodimer in solution in a manner akin to when reconstituted with divalent metal ions but its hydrodynamic radius also does not seem to change, implying that the protein possesses a globular fold even in the absence of metal-

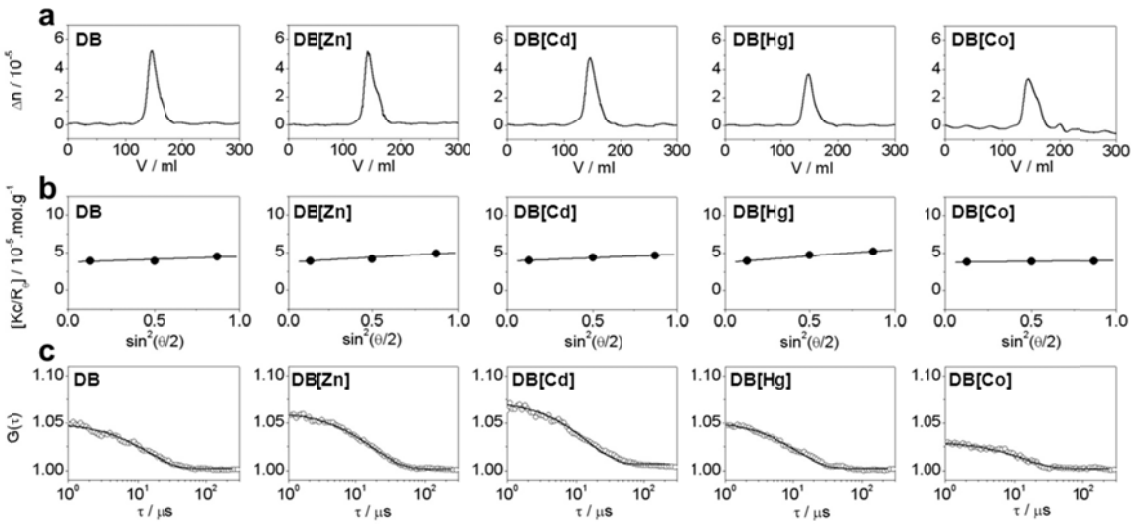

Figure 5-4: Representative ALS chromatograms for the DB domain of ERα pre-treated with EDTA to strip divalent zinc ions (DB) and upon reconstitution with divalent ions of zinc (DB[Zn]), cadmium (DB[Cd]), mercury (DB[Hg]) and cobalt (DB[Co]). (a) Elution profiles as monitored by the differential refractive index (Δn) plotted as a function of elution volume (V) for the indicated species. (b) Partial Zimm plots obtained from analytical SLS measurements at a specific protein concentration for indicated species. The solid lines through the data points represent linear fits. (c) Autocorrelation function plots obtained from analytical DLS measurements at a specific protein concentration for the indicated species. The solid lines through the data points represent non-linear least squares fits to Eq [5-4].

Table 5-2

Hydrodynamic parameters obtained from ALS measurements for the DB domain of ERα pre-treated with EDTA to strip divalent zinc ions (DB) and upon reconstitution with divalent ions of zinc (DB[Zn]), cadmium (DB[Cd]), mercury (DB[Hg]) and cobalt (DB[Co])

	M_{obs} / kD	D_t / $\mu m^2.s^{-1}$	R_h / Å	Associativity
DB	26 ± 1	102 ± 3	24 ± 1	Dimer
DB[Zn]	27 ± 1	104 ± 9	23 ± 2	Dimer
DB[Cd]	26 ± 1	102 ± 5	25 ± 1	Dimer
DB[Hg]	27 ± 1	103 ± 4	23 ± 1	Dimer
DB[Co]	26 ± 1	108 ± 5	23 ± 1	Dimer

M_{obs}, D_t and R_h are respectively the observed molecular mass, translational diffusion coefficient and hydrodynamic radius for each indicated species. Note that the calculated molecular mass of the recombinant DB domain from amino acid sequence alone is 13 kD. Errors were calculated from at least three independent measurements. All errors are given to one standard deviation.

coordination. It is however important to note that our ALS measurements were conducted on DB domains in the 20-50μM range. Although all DB domains behaved as homodimers within this concentration range, it is conceivable that the DB domain in the

absence or presence of metal-coordination may behave as a monomer at protein concentrations in the submicromolar range. It should however be noted that ALS measurements for the DB domain outside the 20-50μM concentration range were not feasible. Thus, while the ALS signal-to-noise ratio starts to get poorer at protein concentrations below 20μM making hydrodynamic analysis less reliable, the DB domain appears to precipitate on the SEC column at concentrations above 50μM. Although it has not been possible at this stage, our future efforts will further explore the ability of DB domain to homodimerize in solution using alternative methodologies such as analytical ultracentrifugation and native mass spectrometry. Nonetheless, these observations suggest strongly that the DB domain may be able to attain a globular fold alone and that metal-coordination may only be essential for DNA-binding. Importantly, these findings are consistent with the knowledge that the LB domain of ERα also behaves as a homodimer in solution [*171*], suggesting that both the LB and DB domains contribute to the dimerization of ERα in solution [*172, 173*].

5.4.3 Substitution of zinc with other divalent metal ions results in subtle secondary and tertiary structural changes within the DB domain of ERα

To further understand how the binding of various metals changes secondary and tertiary structural features of the DB domain of ERα, we conducted CD analysis of DB domain in complex with various divalent metal ions (Figure 5-5). Consistent with our ALS analysis above, the metal-free DB domain indeed displays spectral features in the far-UV region characteristic of a mixed αβ-fold, with bands centered around 210nm and 220nm (Figure 5-5a). The metal-free DB domain also displays near-UV spectral features characteristic of a globular fold with bands centered around 260nm and 280nm (Figure 5-5b), due respectively to Phe and Trp/Tyr residues. Upon the addition of divalent ions of

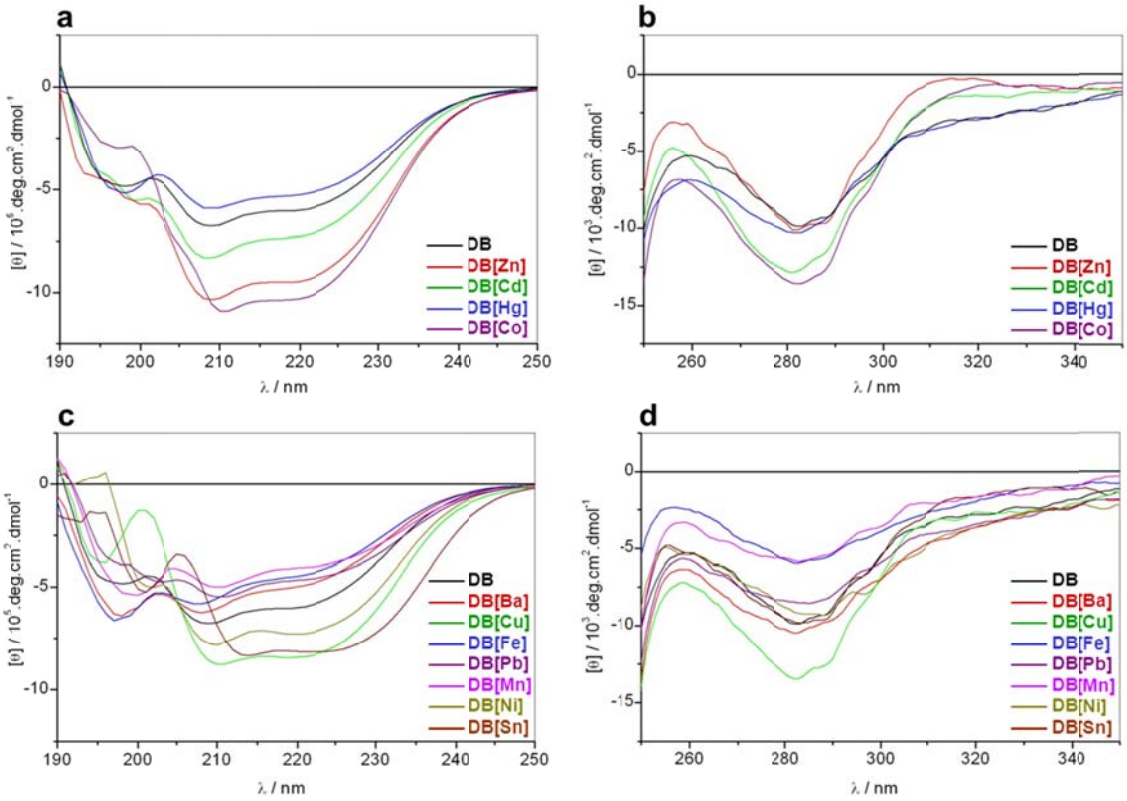

Figure 5-5: Representative far-UV (a and c) and near-UV (b and d) CD spectra for the DB domain of ERα pre-treated with EDTA to strip divalent zinc ions (DB) and upon reconstitution with divalent ions of zinc (DB[Zn]), cadmium (DB[Cd]), mercury (DB[Hg]), cobalt (DB[Co]), barium (DB[Ba]), copper (DB[Cu]), iron (DB[Fe]), lead (DB[*25*]), manganese (DB[Mn]), nickel (DB[Ni]) and tin (DB[Sn]). Note that the spectra with divalent metal ions that can regenerate the DB domain with DNA-binding potential are presented in (a-b), while spectra for divalent metal ions that fail to do so in (c-d).

zinc and cobalt, the far-UV and near-UV spectra of the DB domain undergo sharp enhancement in intensity, implying that the binding of zinc and cobalt induces substantial folding of the protein. In contrast, addition of cadmium results in only slight changes in the spectral intensities of the DB domain in the far-UV and near UV-regions, suggesting that cadmium-coordination also results in some degree of folding but is not sufficient to lead to fully structured protein. Strikingly, in the presence of mercury, the far-UV and near-UV spectral intensities of the DB domain undergo slight reduction. This salient observation implies that the coordination of mercury has little or no effect on the secondary or tertiary structure of the DB domain in agreement with our thermodynamic

data (Table 5-1). We believe that this is most likely due to the preference of mercury to coordinate its ligands with linear geometry [*169*]. Accordingly, the cysteine ligands within the DB domain may be coordinated by mercury with a linear geometry as opposed to tetrahedral arrangement necessary for its proper folding as observed in the case of coordination with zinc, cobalt and cadmium divalent ions. In light of our thermodynamic data (Table 5-1), it is conceivable that the binding of DNA causes mercury to switch its coordination from linear to tetrahedral geometry so as to allow the DB domain to undergo proper folding necessary for DNA recognition.

Taken together, our CD data are consistent with the notion that the DB domain may be able to attain a globular fold alone and that metal-coordination may only be essential for DNA-binding. It should also be noted here that the failure of DB domain regenerated with divalent ions of barium, copper, iron, lead, manganese, nickel and tin to bind DNA does not necessarily imply failure to replace zinc within the zinc fingers of the DB domain. Indeed, it has been previously reported that the divalent ions of copper can replace zinc within the zinc fingers of DB and that such substitution is met with profound effects on the physiological action of ERα [*78, 174-176*]. It is thus conceivable that divalent ions of barium, copper, iron, lead, manganese, nickel and tin can replace zinc within the zinc fingers of DB domain but their inability to coordinate cysteine ligands in a tetrahedral arrangement results in improper folding such that the DB domain can no longer recognize the target DNA. This is indeed further supported by our CD measurements of DB domain in complex with divalent metal ions of barium, copper, iron, lead, manganese, nickel and tin (Figures 5-5c and 5-5d). Evidently, the addition of these metal divalent ions results in marked changes in both the secondary and tertiary

structural features of the DB domain, implying that these metals can also coordinate the protein but in a non-productive manner. In light of these observations, we believe that metals such as barium, copper, iron, lead, manganese, nickel and tin may also be able to compete with zinc for coordinating to the DB domain within living cells with important consequences on the physiological action of ERα.

5.4.4 DB domain of ERα reconstituted with various metals displays distinct spectroscopic properties

In an attempt to further analyze how reconstitution of DB domain of ERα with various divalent metal ions results in structural changes, we also measured SSA and SSF spectra (Figures 5-6a and 5-6b). As expected, the metal-free DB domain displays characteristic absorbance spectral features in the UV region with maxima centered around 225nm and 280nm (Figure 5-6a), due respectively to peptide bonds and Trp/Tyr/Phe residues. Additionally, the DB domain reconstituted with various metals also displays similar spectroscopic features but the spectral intensity of the 280-nm band appears to undergo reduction, implying that metal-binding most likely induces structural changes within the DB domain, due for example to burial of Trp/Tyr/Phe residues. Such metal-mediated modulation of spectral features is further indicative of specific metal-protein interactions as noted previously [*89, 90*]. Consistent with our absorbance measurements, the DB domain reconstituted with various metals also displays fluorescence properties distinct from those observed for the metal-free DB domain (Figure 5-6b). Thus, the increase in fluorescence intensity and the emission wavelength maximum (λ_{max}) undergoing a slight blue-shift in the DB domain reconstituted with various metals relative to metal-free DB domain is indicative of the transfer of Trp to a

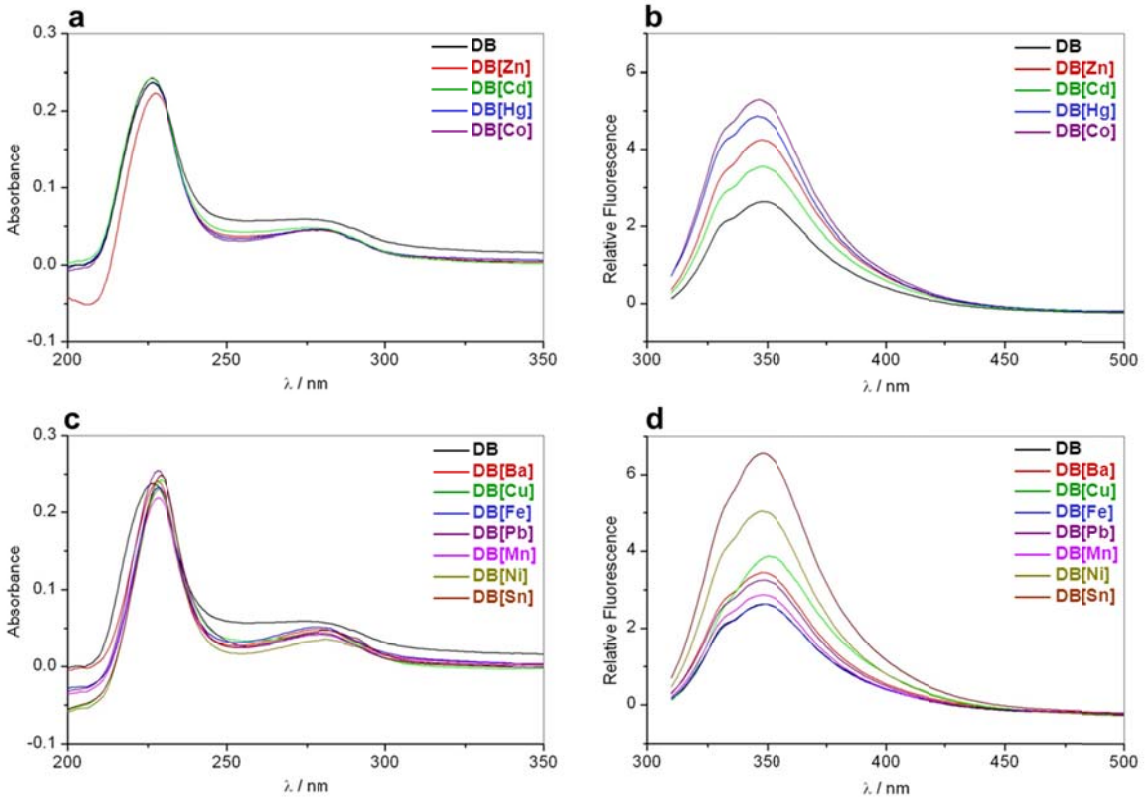

Figure 5-6: Representative SSA (a and c) and SSF (b and d) spectra for the DB domain of ERα pre-treated with EDTA to strip divalent zinc ions (DB) and upon reconstitution with divalent ions of zinc (DB[Zn]), cadmium (DB[Cd]), mercury (DB[Hg]), cobalt (DB[Co]), barium (DB[Ba]), copper (DB[Cu]), iron (DB[Fe]), lead (DB[*25*]), manganese (DB[Mn]), nickel (DB[Ni]) and tin (DB[Sn]). Note that the spectra with divalent metal ions that can regenerate the DB domain with DNA-binding potential are presented in (a-b), while spectra for divalent metal ions that fail to do so in (c-d).

more hydrophobic environment and thereby suggesting that the protein undergoes conformational changes upon metal-coordination.

Interestingly, mercury-coordinated DB domain also undergoes substantial enhancement in fluorescence, implying that coordination of mercury also results in structural changes within the DB domain. However, this observation is at odds with our CD data above, wherein the CD spectral features of mercury-coordinated DB domain did not differ much from those observed for the metal-free DB domain. It is however conceivable that mercury-coordination of cysteine ligands within the DB domain with Cys-Hg^{2+}-Cys linear geometry primarily results in perturbation of environment around

Trp, which could also account for the enhancement of fluorescence observed here. Notably, our SSA and SSF measurements also suggest that the addition of divalent ions of barium, copper, iron, lead, manganese, nickel and tin also results in changes in spectroscopic properties of DB domain (Figures 5-6c and 5-6d), implying that although these metals cannot regenerate DB domain with DNA-binding potential, they are nevertheless capable of competing with zinc for the zinc fingers in agreement with our CD data.

5.5 Concluding remarks

In the industrialized world, the environment has come to play an increasing role on human health due to the presence of obnoxious chemicals in a wide variety of sources. In particular, over the past decade or so, it has become clear that metal ions absorbed from various environmental sources can potentiate the transcriptional activity of ERα within the body leading to the development and progression of breast cancer [*177-179*]. Although it is generally believed that such metals up-regulate ERα by virtue of their ability to mimic the action of endogenous estrogens, our data presented here suggest that some of these metals may also be able to replace zinc within the zinc fingers of the DB domain and thereby modulate its binding to the promoters of target genes [*65*]. In particular, our data suggest that metals such as barium, copper, iron, lead, manganese, nickel and tin may coordinate to cysteine ligands within the DB domain in a manner that it can longer bind to DNA. It is thus quite possible that these metals may influence the physiological action of ERα by virtue of their ability to compete with zinc for coordinating to DB domain though their concentration is likely to be much lower than zinc within living cells. More importantly, given that the hyperactivation of ERα is

linked to the genesis of large fractions of breast cancer [*8, 30*], deciphering the molecular basis of how environmental metals modulate the transcriptional activity of ERα bears the potential to not only expand our biomedical knowledge but may also lead to the development of novel anti-cancer therapies harboring greater efficacy coupled with low toxicity. Toward this goal, our present study provides mechanistic insights into how environmental metals may replace zinc within the zinc fingers of ERα and thus bears important consequences on understanding its physiological function in human health and disease.

Chapter 6: Conclusion

Nuclear receptors act as ligand-modulated transcription factors and orchestrate a multitude of cellular functions central to health and disease. Although studied for more than half a century, many mysteries surround the mechanism of action of nuclear receptors. In particular, despite the determination of the ERE consensus motif [*65*] and the availability of the crystal structure of the DB domain of ERα in complex with ERE for decades [*66, 67*] , it is surprising such a lack of thermodynamic data addressing this key protein-DNA assembly existed. Knowledge of the thermodynamic forces governing biomolecular interactions is inherent to establishing the fundamental molecular mechanisms behind them. To this end, this thesis reports a thorough biophysical characterization of the binding of the ERα nuclear receptor to DNA.

Our initial studies were conducted with the aim of describing the basic thermodynamic phenomena behind the DB domain of ERα binding to its DNA response element, ERE. Using ITC in conjunction with site-directed mutagenesis this thesis demonstrates that binding of the DB domain to ERE is coupled to proton uptake at two specific residues, H196 and E203. Interestingly, nuclear receptor function is tightly regulated by a multitude of post-translational modifications such as phosphorylation, acetylation, sumoylation, ubiquitination and glycosylation [*123-126*], yet, such modifications typically occur in regions outside the DB domain. The fact that the DB domain of ERα is directly regulated via proton-coupled equilibrium at two critical residues, H196 and E203, located at the protein-DNA interface not only adds to the

repertoire of regulation employed by nuclear receptors but also bears significant implications for furthering our understanding of this important family of transcription factors.

In the crystal structure of the DB domain of ERα in complex with ERE duplex solved nearly two decades ago [*66*], it was proposed that the negative charge on E203 was largely neutralized through the formation of a salt bridge with the neighboring K206. On the contrary, the thesis demonstrates that the negative charge on E203 is rather neutralized through its protonation allowing it to participate in the formation of hydrogen bonding with DNA in a more harmonious manner. Additionally, the crystal structural analysis also suggested the involvement of H196 in dictating protein-DNA interactions through hydrogen bonding with the phosphate backbone. The fact that H196 acquires a net positive charge through protonation upon the binding of DNA suggests that H196 is more likely to engage in the formation of a salt bridge with the phosphate backbone. This thesis illustrates how a combined approach involving site-directed mutagenesis in conjunction with thermodynamics can complement structural data and further define key residues involved in protein-DNA interactions.

Though the binding of the DB domain of ERα to DNA appears to be coupled to proton uptake, it is not clear from our data as to how such coupled equilibrium might dictate the physiological role of this important nuclear receptor. Changes in intracellular pH regulate a multitude of cellular processes such as metabolic homeostasis and apoptosis [*127*]. It is thus conceivable that changes in intracellular pH may also regulate the transcriptional activity of ERα through direct modulation of H196 and E203. Protonation of such residues would clearly enhance intermolecular hydrogen bonding and

electrostatic interactions critical to driving this key protein-DNA interaction. Although pKa values of sidechains of histidine and glutamate within proteins are respectively around 6 and 4 [*131*], these are likely to be influenced by the neighboring ionizable amino acid residues in the DB domain as noted earlier (Figure 3-1). Thus, protonation/deprotonation of H196 and E203 may not necessarily require large changes but may be mediated by small changes in intracellular pH. Whatever the exact physiological role of proton-coupled equilibrium observed here, our current study clearly warrants further investigating the role of pH in physiological processes governed by ERα and other nuclear receptors.

Furthermore, it has been reported that most estrogen-dependent genes contain imperfect ERE sequences [*65*], yet it is poorly understood how sequence variations impact binding energetics beyond the notion they generally bind weaker [*180*]. Herein, this thesis has provided a detailed and systematic analysis of how genetic variations within the ERE motif may affect the binding of ERα nuclear receptor and hence its transcriptional output in response to estrogens. Our data suggest strongly that genetic variations can modulate the binding of ERα by orders of magnitude and that such modulation may or may not involve drastic changes in the contribution of underlying thermodynamic forces driving subtle protein-DNA interactions. Although it is widely believed that AT sequences within DNA account for its intrinsic conformational flexibility such as bending and curvature [*132, 133, 149-154*], our data presented here suggest a poor correlation between DNA flexibility and its binding to ERα. Nevertheless, our CD analysis shows that the introduction of genetic variations within the ERE motif allows it to sample a much greater conformational space which may be a key

determinant of the ability of its variants to bind to ERα at distinct promoters in a selective manner. Our atomic models also provide structural basis of how the symmetrical introduction of A-2/T+2 nucleotide pair within both half-sites of ERE can result in the reduction of binding of DB domain by nearly two orders of magnitude. Likewise, we postulate that the introduction of nucleotides at other positions within the ERE motif is likely to result in the loss of hydrogen bonding and other stabilizing interactions due to the rearrangement of amino acid sidechains in the DB domain. Thus, the energetic and conformational heterogeneity associated with sequence variants of the ERE may have emerged as an evolutionary tool to fine-tune the physiological action of ERα as well as other nuclear receptors.

Lastly, after extensive characterization of the DB-ERE protein-DNA interface, this thesis investigated the effects of metal-substitution within the zinc-fingers of the DB domain of ERα on structure and function. Over the past decade, it has become clear that metal ions absorbed from various environmental sources can potentiate the transcriptional activity of ERα within the body leading to the development and progression of breast cancer [*177-179*]. Although it is generally believed that such metals activate ERα by virtue of their ability to mimic the action of endogenous estrogens, our data suggest that some of these metals may also be able to replace zinc within the zinc fingers of the DB domain and thereby modulate its binding to the promoters of target genes [*65*]. Divalent cations of cobalt, cadmium and mercury are, in fact, able to reconstitute a functional DB domain retaining the capacity to bind ERE. These particular metal-reconstituted DB domains all bind ERE with similar affinities despite differences in the underlying thermodynamics. Our data suggest that metals such as barium, copper, iron, lead,

manganese, nickel and tin may coordinate to cysteine ligands within the DB domain in a manner that it can longer bind to DNA. Interestingly, DNA binding capacity cannot be predicted from CD data, indicating that, in the case of reconstituted zinc-finger proteins, isostructure does not imply isofunction. In addition, the identity of the reconstituting metal does not influence hydrodynamic properties. It is thus quite possible that these metals may influence the physiological action of ERα by virtue of their ability to compete with zinc for coordination to the DB domain. Though their concentration is likely to be much lower than zinc within living cells, chronic lifetime accumulation could pose a threat. More importantly, given that the hyperactivation of ERα is linked to the genesis of large fractions of breast cancer [*8, 30*], deciphering the molecular basis of how environmental metals modulate the transcriptional activity of ERα bears the potential to not only expand our biomedical knowledge but may also lead to the development of novel anti-cancer therapies harboring greater efficacy coupled with low toxicity. Toward this goal, this thesis provides mechanistic insights into how environmental metals may replace zinc within the zinc fingers of ERα and thus bears important consequences on understanding its physiological function in human health and disease.

In conclusion, our data reported here reveal key insights into the biophysical mechanisms governing the ERα-DNA interaction. Through structural analysis and extensive calorimetric studies, we demonstrated that the protonation of H196 and E203 in ERα is coupled to the binding of DNA and that such protonation is required for high-affinity protein-DNA interaction through thermodynamic optimization of intermolecular contacts. Amino acid sequence alignment of the DB domains of the nuclear receptor family members revealed that protonation-coupled DNA binding may be a hallmark of

the entire family. Similar studies of other nuclear receptors are warranted to test whether protonation is in fact conserved. Clearly confirmation of these observations at the cellular level would enhance our understanding of the potential impact of proton-linkage phenomena. Our systematic analysis on the effects of single nucleotide variations within the ERE has at last defined the ability of ERα to bind these imperfect motifs. Our data reveal that ERα is able to bind all symmetric single nucleotide variants in the physiologically relevant nanomolar - micromolar range. It is conceivable such data could be applied toward predicting new estrogen-dependent genes or used to make predictions at the functional genomic level concerning global estrogen-dependent gene response to estrogen exposure. Finally, our data indicate that while several metals can replace zinc within the zinc-fingers of the DB domain of ERα, only a few are able to reconstitute and retain DNA-binding. It would be interesting to determine the precise coordination geometry of each metal within the DB domain both in the absence and presence of DNA as well as the stoichiometry. Also, because these metals are considered to have estrogen-like effects, and are sometimes referred to as metalloestrogens, it would be intriguing to determine if estrogenic potency is related to the ability of a metal to yield a functioning DB domain. As is the case with most science, the work of this thesis and its goal of understanding the molecular mechanisms driving the ERα-DNA interaction likely beg more questions than it answered. Nonetheless, these studies further enhance to our understanding of the molecular action of nuclear receptors, and in particular the ERα-ERE interaction. Furthermore, novel therapeutic avenues for targeting the nuclear receptors may one day open up as a result of these studies.

Works Cited

1. **Savage, D. C. (1977) Microbial ecology of the gastrointestinal tract, *Annu Rev Microbiol 31*, 107-133.**

2. **Bjornstrom, L., and Sjoberg, M. (2005) Mechanisms of estrogen receptor signaling: convergence of genomic and nongenomic actions on target genes, *Mol Endocrinol 19*, 833-842.**

3. **Evans, R. M. (1988) The steroid and thyroid hormone receptor superfamily, *Science 240*, 889-895.**

4. **Escriva, H., Bertrand, S., and Laudet, V. (2004) The evolution of the nuclear receptor superfamily, *Essays Biochem 40*, 11-26.**

5. **Thornton, J. W. (2001) Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions, *Proc Natl Acad Sci U S A 98*, 5671-5676.**

6. **McKenna, N. J., Cooney, A. J., DeMayo, F. J., Downes, M., Glass, C. K., Lanz, R. B., Lazar, M. A., Mangelsdorf, D. J., Moore, D. D., Qin, J., Steffen, D. L., Tsai, M. J., Tsai, S. Y., Yu, R., Margolis, R. N., Evans, R. M., and O'Malley, B. W. (2009) Minireview: Evolution of NURSA, the Nuclear Receptor Signaling Atlas, *Mol Endocrinol 23*, 740-746.**

7. **Hall, J. M., Couse, J. F., and Korach, K. S. (2001) The multifaceted mechanisms of estradiol and estrogen receptor signaling, *J Biol Chem 276*, 36869-36872.**

8. **Heldring, N., Pike, A., Andersson, S., Matthews, J., Cheng, G., Hartman, J., Tujague, M., Strom, A., Treuter, E., Warner, M., and Gustafsson, J. A. (2007) Estrogen receptors: how do they signal and what are their targets, *Physiol Rev 87*, 905-931.**

9. **Egea, P. F., Klaholz, B. P., and Moras, D. (2000) Ligand-protein interactions in nuclear receptors of hormones, *FEBS Lett 476*, 62-67.**

10. **Claessens, F., and Gewirth, D. T. (2004) DNA recognition by nuclear receptors, *Essays Biochem 40*, 59-72.**

11. **Green, S., and Chambon, P. (1988) Nuclear receptors enhance our understanding of transcription regulation, *Trends Genet 4*, 309-314.**

12. **Rosenfeld, M. G., and Glass, C. K. (2001) Coregulator codes of transcriptional regulation by nuclear receptors, *J Biol Chem 276*, 36865-36868.**

13. O'Lone, R., Frith, M. C., Karlsson, E. K., and Hansen, U. (2004) Genomic targets of nuclear estrogen receptors, *Mol Endocrinol 18*, 1859-1875.

14. Kushner, P. J., Agard, D. A., Greene, G. L., Scanlan, T. S., Shiau, A. K., Uht, R. M., and Webb, P. (2000) Estrogen receptor pathways to AP-1, *J Steroid Biochem Mol Biol 74*, 311-317.

15. Saville, B., Wormke, M., Wang, F., Nguyen, T., Enmark, E., Kuiper, G., Gustafsson, J. A., and Safe, S. (2000) Ligand-, cell-, and estrogen receptor subtype (alpha/beta)-dependent activation at GC-rich (Sp1) promoter elements, *J Biol Chem 275*, 5379-5387.

16. Liu, M. M., Albanese, C., Anderson, C. M., Hilty, K., Webb, P., Uht, R. M., Price, R. H., Jr., Pestell, R. G., and Kushner, P. J. (2002) Opposing action of estrogen receptors alpha and beta on cyclin D1 gene expression, *J Biol Chem 277*, 24353-24360.

17. Umayahara, Y., Kawamori, R., Watada, H., Imano, E., Iwama, N., Morishima, T., Yamasaki, Y., Kajimoto, Y., and Kamada, T. (1994) Estrogen regulation of the insulin-like growth factor I gene transcription involves an AP-1 enhancer, *J Biol Chem 269*, 16433-16442.

18. Webb, P., Lopez, G. N., Uht, R. M., and Kushner, P. J. (1995) Tamoxifen activation of the estrogen receptor/AP-1 pathway: potential origin for the cell-specific estrogen-like effects of antiestrogens, *Mol Endocrinol 9*, 443-456.

19. Li, C., Briggs, M. R., Ahlborn, T. E., Kraemer, F. B., and Liu, J. (2001) Requirement of Sp1 and estrogen receptor alpha interaction in 17beta-estradiol-mediated transcriptional activation of the low density lipoprotein receptor gene expression, *Endocrinology 142*, 1546-1553.

20. Losel, R., and Wehling, M. (2003) Nongenomic actions of steroid hormones, *Nat Rev Mol Cell Biol 4*, 46-56.

21. Losel, R. M., Falkenstein, E., Feuring, M., Schultz, A., Tillmann, H. C., Rossol-Haseroth, K., and Wehling, M. (2003) Nongenomic steroid action: controversies, questions, and answers, *Physiol Rev 83*, 965-1016.

22. Migliaccio, A., Di Domenico, M., Castoria, G., de Falco, A., Bontempo, P., Nola, E., and Auricchio, F. (1996) Tyrosine kinase/p21ras/MAP-kinase pathway activation by estradiol-receptor complex in MCF-7 cells, *EMBO J 15*, 1292-1300.

23. Endoh, H., Sasaki, H., Maruyama, K., Takeyama, K., Waga, I., Shimizu, T., Kato, S., and Kawashima, H. (1997) Rapid activation of MAP kinase by estrogen in the bone cell line, *Biochem Biophys Res Commun 235*, 99-102.

24. Jessop, H. L., Sjoberg, M., Cheng, M. Z., Zaman, G., Wheeler-Jones, C. P., and Lanyon, L. E. (2001) Mechanical strain and estrogen activate estrogen receptor alpha in bone cells, *J Bone Miner Res 16*, 1045-1055.

25. Watters, J. J., Campbell, J. S., Cunningham, M. J., Krebs, E. G., and Dorsa, D. M. (1997) Rapid membrane effects of steroids in neuroblastoma cells: effects of estrogen on mitogen activated protein kinase signalling cascade and c-fos immediate early gene transcription, *Endocrinology 138*, 4030-4033.

26. Aronica, S. M., Kraus, W. L., and Katzenellenbogen, B. S. (1994) Estrogen action via the cAMP signaling pathway: stimulation of adenylate cyclase and cAMP-regulated gene transcription, *Proc Natl Acad Sci U S A 91*, 8517-8521.

27. Improta-Brears, T., Whorton, A. R., Codazzi, F., York, J. D., Meyer, T., and McDonnell, D. P. (1999) Estrogen-induced activation of mitogen-activated protein kinase requires mobilization of intracellular calcium, *Proc Natl Acad Sci U S A 96*, 4686-4691.

28. Kato, S., Endoh, H., Masuhiro, Y., Kitamoto, T., Uchiyama, S., Sasaki, H., Masushige, S., Gotoh, Y., Nishida, E., Kawashima, H., Metzger, D., and Chambon, P. (1995) Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase, *Science 270*, 1491-1494.

29. Coutts, A. S., and Murphy, L. C. (1998) Elevated mitogen-activated protein kinase activity in estrogen-nonresponsive human breast cancer cells, *Cancer Res 58*, 4071-4074.

30. Deroo, B. J., and Korach, K. S. (2006) Estrogen receptors and human disease, *J Clin Invest 116*, 561-570.

31. Mueller, S. O., and Korach, K. S. (2001) Estrogen receptors and endocrine diseases: lessons from estrogen receptor knockout mice, *Curr Opin Pharmacol 1*, 613-619.

32. Couse, J. F., and Korach, K. S. (1999) Estrogen receptor null mice: what have we learned and where will they lead us?, *Endocr Rev 20*, 358-417.

33. Lubahn, D. B., Moyer, J. S., Golding, T. S., Couse, J. F., Korach, K. S., and Smithies, O. (1993) Alteration of reproductive function but not prenatal sexual development after insertional disruption of the mouse estrogen receptor gene, *Proc Natl Acad Sci U S A 90*, 11162-11166.

34. Hewitt, S. C., Harrell, J. C., and Korach, K. S. (2005) Lessons in estrogen biology from knockout and transgenic animals, *Annu Rev Physiol 67*, 285-308.

35. Visvader, J. E., and Lindeman, G. J. (2003) Transcriptional regulators in mammary gland development and cancer, *Int J Biochem Cell Biol 35*, 1034-1051.

36. Couse, J. E., Mahato, D., Eddy, E. M., and Korach, K. S. (2001) Molecular mechanism of estrogen action in the male: insights from the estrogen receptor null mice, *Reprod Fertil Dev 13*, 211-219.

37. Mahato, D., Goulding, E. H., Korach, K. S., and Eddy, E. M. (2000) Spermatogenic cells do not require estrogen receptor-alpha for development or function, *Endocrinology 141*, 1273-1276.

38. Mahato, D., Goulding, E. H., Korach, K. S., and Eddy, E. M. (2001) Estrogen receptor-alpha is required by the supporting somatic cells for spermatogenesis, *Mol Cell Endocrinol 178*, 57-63.

39. Losordo, D. W., Kearney, M., Kim, E. A., Jekanowski, J., and Isner, J. M. (1994) Variable expression of the estrogen receptor in normal and atherosclerotic coronary arteries of premenopausal women, *Circulation 89*, 1501-1510.

40. Post, W. S., Goldschmidt-Clermont, P. J., Wilhide, C. C., Heldman, A. W., Sussman, M. S., Ouyang, P., Milliken, E. E., and Issa, J. P. (1999) Methylation of the estrogen receptor gene is associated with aging and atherosclerosis in the cardiovascular system, *Cardiovasc Res 43*, 985-991.

41. Herrington, D. M., Howard, T. D., Hawkins, G. A., Reboussin, D. M., Xu, J., Zheng, S. L., Brosnihan, K. B., Meyers, D. A., and Bleecker, E. R. (2002) Estrogen-receptor polymorphisms and effects of estrogen replacement on high-density lipoprotein cholesterol in women with coronary disease, *N Engl J Med 346*, 967-974.

42. Almeida, S., Franken, N., Zandona, M. R., Osorio-Wender, M. C., and Hutz, M. H. (2005) Estrogen receptor 2 and progesterone receptor gene polymorphisms and lipid levels in women with different hormonal status, *Pharmacogenomics J 5*, 30-34.

43. Rossouw, J. E., Anderson, G. L., Prentice, R. L., LaCroix, A. Z., Kooperberg, C., Stefanick, M. L., Jackson, R. D., Beresford, S. A., Howard, B. V., Johnson, K. C., Kotchen, J. M., and Ockene, J. (2002) Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial, *JAMA 288*, 321-333.

44. Parfitt, A. M. (1994) Osteonal and hemi-osteonal remodeling: the spatial and temporal framework for signal traffic in adult human bone, *J Cell Biochem 55*, 273-286.

45. Jilka, R. L., Takahashi, K., Munshi, M., Williams, D. C., Roberson, P. K., and Manolagas, S. C. (1998) Loss of estrogen upregulates osteoblastogenesis in the murine bone marrow. Evidence for autonomy from factors released during bone resorption, *J Clin Invest 101*, 1942-1950.

46. Di Gregorio, G. B., Yamamoto, M., Ali, A. A., Abe, E., Roberson, P., Manolagas, S. C., and Jilka, R. L. (2001) Attenuation of the self-renewal of transit-amplifying osteoblast progenitors in the murine bone marrow by 17 beta-estradiol, *J Clin Invest 107*, 803-812.

47. Syed, F., and Khosla, S. (2005) Mechanisms of sex steroid effects on bone, *Biochem Biophys Res Commun 328*, 688-696.

48. Korach, K. S., Couse, J. F., Curtis, S. W., Washburn, T. F., Lindzey, J., Kimbro, K. S., Eddy, E. M., Migliaccio, S., Snedeker, S. M., Lubahn, D. B., Schomberg, D. W., and Smith, E. P. (1996) Estrogen receptor gene disruption: molecular characterization and experimental and clinical phenotypes, *Recent Prog Horm Res 51*, 159-186; discussion 186-158.

49. Jemal, A., Tiwari, R. C., Murray, T., Ghafoor, A., Samuels, A., Ward, E., Feuer, E. J., and Thun, M. J. (2004) Cancer statistics, 2004, *CA Cancer J Clin 54*, 8-29.

50. Yager, J. D., and Davidson, N. E. (2006) Estrogen carcinogenesis in breast cancer, *N Engl J Med 354*, 270-282.

51. **Key, T. J., Appleby, P. N., Reeves, G. K., Roddam, A., Dorgan, J. F., Longcope, C., Stanczyk, F. Z., Stephenson, H. E., Jr., Falk, R. T., Miller, R., Schatzkin, A., Allen, D. S., Fentiman, I. S., Wang, D. Y., Dowsett, M., Thomas, H. V., Hankinson, S. E., Toniolo, P., Akhmedkhanov, A., Koenig, K., Shore, R. E., Zeleniuch-Jacquotte, A., Berrino, F., Muti, P., Micheli, A., Krogh, V., Sieri, S., Pala, V., Venturelli, E., Secreto, G., Barrett-Connor, E., Laughlin, G. A., Kabuto, M., Akiba, S., Stevens, R. G., Neriishi, K., Land, C. E., Cauley, J. A., Kuller, L. H., Cummings, S. R., Helzlsouer, K. J., Alberg, A. J., Bush, T. L., Comstock, G. W., Gordon, G. B., and Miller, S. R. (2003) Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women, *J Natl Cancer Inst 95*, 1218-1226.**

52. **Chlebowski, R. T., Hendrix, S. L., Langer, R. D., Stefanick, M. L., Gass, M., Lane, D., Rodabough, R. J., Gilligan, M. A., Cyr, M. G., Thomson, C. A., Khandekar, J., Petrovitch, H., and McTiernan, A. (2003) Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women: the Women's Health Initiative Randomized Trial, *JAMA 289*, 3243-3253.**

53. **Li, K. M., Todorovic, R., Devanesan, P., Higginbotham, S., Kofeler, H., Ramanathan, R., Gross, M. L., Rogan, E. G., and Cavalieri, E. L. (2004) Metabolism and DNA binding studies of 4-hydroxyestradiol and estradiol-3,4-quinone in vitro and in female ACI rat mammary gland in vivo, *Carcinogenesis 25*, 289-297.**

54. **Yue, W., Santen, R. J., Wang, J. P., Li, Y., Verderame, M. F., Bocchinfuso, W. P., Korach, K. S., Devanesan, P., Todorovic, R., Rogan, E. G., and Cavalieri, E. L. (2003) Genotoxic metabolites of estradiol in breast: potential mechanism of estradiol induced carcinogenesis, *J Steroid Biochem Mol Biol 86*, 477-486.**

55. **Beatson, G. T. (1896) On the treatment of inoperable cases of carcinomaof the mamma: suggestions for a new method of treatment with illustrative cases, *Lancet 2*, 104-107.**

56. **Group, E. B. C. T. C. (2005) Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials, *Lancet 365*, 1687-1717.**

57. **Cuzick, J., Powles, T., Veronesi, U., Forbes, J., Edwards, R., Ashley, S., and Boyle, P. (2003) Overview of the main outcomes in breast-cancer prevention trials, *Lancet 361*, 296-300.**

58. **McEwan, I. J. (2009) Nuclear receptors: one big family, *Methods Mol Biol 505*, 3-18.**

59. Barnett, P., Tabak, H. F., and Hettema, E. H. (2000) Nuclear receptors arose from pre-existing protein modules during evolution, *Trends Biochem Sci 25*, 227-228.

60. Kumar, R., and Thompson, E. B. (1999) The structure of the nuclear hormone receptors, *Steroids 64*, 310-319.

61. Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., and Evans, R. M. (1995) The nuclear receptor superfamily: the second decade, *Cell 83*, 835-839.

62. Darimont, B. D., Wagner, R. L., Apriletti, J. W., Stallcup, M. R., Kushner, P. J., Baxter, J. D., Fletterick, R. J., and Yamamoto, K. R. (1998) Structure and specificity of nuclear receptor-coactivator interactions, *Genes Dev 12*, 3343-3356.

63. Ham, J., and Parker, M. G. (1989) Regulation of gene expression by nuclear hormone receptors, *Curr Opin Cell Biol 1*, 503-511.

64. Warnmark, A., Treuter, E., Wright, A. P., and Gustafsson, J. A. (2003) Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation, *Mol Endocrinol 17*, 1901-1909.

65. Klinge, C. M. (2001) Estrogen receptor interaction with estrogen response elements, *Nucleic Acids Res 29*, 2905-2919.

66. Schwabe, J. W., Chapman, L., Finch, J. T., and Rhodes, D. (1993) The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements, *Cell 75*, 567-578.

67. Schwabe, J. W., Neuhaus, D., and Rhodes, D. (1990) Solution structure of the DNA-binding domain of the oestrogen receptor, *Nature 348*, 458-461.

68. Green, S., and Chambon, P. (1987) Oestradiol induction of a glucocorticoid-responsive gene by a chimaeric receptor, *Nature 325*, 75-78.

69. Mader, S., Kumar, V., de Verneuil, H., and Chambon, P. (1989) Three amino acids of the oestrogen receptor are essential to its ability to distinguish an oestrogen from a glucocorticoid-responsive element, *Nature 338*, 271-274.

70. Koszewski, N. J., and Notides, A. C. (1991) Phosphate-sensitive binding of the estrogen receptor to its response elements, *Mol Endocrinol 5*, 1129-1136.

71. Obourn, J. D., Koszewski, N. J., and Notides, A. C. (1993) Hormone- and DNA-binding mechanisms of the recombinant human estrogen receptor, *Biochemistry 32*, 6229-6236.

72. Nordeen, S. K., Suh, B. J., Kuhnel, B., and Hutchison, C. A., 3rd. (1990) Structural determinants of a glucocorticoid receptor recognition element, *Mol Endocrinol 4*, 1866-1873.

73. Nordeen, S. K., Ogden, C. A., Taraseviciene, L., and Lieberman, B. A. (1998) Extreme position dependence of a canonical hormone response element, *Mol Endocrinol 12*, 891-898.

74. Lieberman, B. A., Bona, B. J., Edwards, D. P., and Nordeen, S. K. (1993) The constitution of a progesterone response element, *Mol Endocrinol 7*, 515-527.

75. Nelson, C. C., Hendy, S. C., Shukin, R. J., Cheng, H., Bruchovsky, N., Koop, B. F., and Rennie, P. S. (1999) Determinants of DNA sequence specificity of the androgen, progesterone, and glucocorticoid receptors: evidence for differential steroid receptor response elements, *Mol Endocrinol 13*, 2090-2107.

76. Green, A., Parker, M., Conte, D., and Sarkar, B. (1998) Zinc finger proteins: A bridge between transition metals and gene regulation, *The Journal of Trace Elements in Experimental Medicine 11*, 103-118.

77. Hartwig, A. (2001) Zinc finger proteins as potential targets for toxic metal ions: differential effects on structure and function, *Antioxid Redox Signal 3*, 625-634.

78. Predki, P. F., and Sarkar, B. (1992) Effect of replacement of "zinc finger" zinc on estrogen receptor DNA interactions, *J Biol Chem 267*, 5842-5846.

79. Predki, P. F., and Sarkar, B. (1994) Metal replacement in "zinc finger" and its effect on DNA binding, *Environ Health Perspect 102 Suppl 3*, 195-198.

80. Conte, D., Narindrasorasak, S., and Sarkar, B. (1996) In vivo and in vitro iron-replaced zinc finger generates free radicals and causes DNA damage, *J Biol Chem 271*, 5125-5130.

81. 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* (Walker, J. M., Ed.), pp 571-607, Humana Press, Totowa, New Jersey, USA.

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

83. 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.

84. Zimm, B. H. (1948) The scattering of light and the radial distribution function of high polymer solutions, *J Chem Phys 16*, 1093-1099.

85. Wyatt, P. J. (1993) Light scattering and the absolute characterization of macromolecules, *Anal Chim Acta 272*, 1-40.

86. Berne, B. J., and Pecora, R. (1976) *Dynamic Light Scattering*, Wiley, New York.

87. Chu, B. (1991) *Laser Light Scattering: Basic Principles and Practice*, Academic, Boston.

88. Koppel, D. E. (1972) Analysis of macromolecular polydispersity in intensity correlation spectroscopy, *J. Chem. Phys. 57*, 4814-4820.

89. Bal, W., Schwerdtle, T., and Hartwig, A. (2003) Mechanism of nickel assault on the zinc finger of DNA repair protein XPA, *Chem Res Toxicol 16*, 242-248.

90. Kopera, E., Schwerdtle, T., Hartwig, A., and Bal, W. (2004) Co(II) and Cd(II) substitute for Zn(II) in the zinc finger derived from the DNA repair protein XPA, demonstrating a variety of potential mechanisms of toxicity, *Chem Res Toxicol 17*, 1452-1458.

91. Leavitt, S., and Freire, E. (2001) Direct measurement of protein binding energetics by isothermal titration calorimetry, *Curr. Opin. Struct. Biol. 11*, 560-566.

92. Cliff, M. J., and Ladbury, J. E. (2003) A survey of the year 2002 literature on applications of isothermal titration calorimetry, *J Mol Recognit 16*, 383-391.

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

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

95. 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.

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

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

98. 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.

99. 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.

100. 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.

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

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

103. Noy, N. (2007) Ligand specificity of nuclear hormone receptors: sifting through promiscuity, *Biochemistry 46*, 13461-13467.

104. Gottlieb, B., Beitel, L. K., Wu, J., Elhaji, Y. A., and Trifiro, M. (2004) Nuclear receptors and disease: androgen receptor, *Essays Biochem 40*, 121-136.

105. Gurnell, M., and Chatterjee, V. K. (2004) Nuclear receptors in disease: thyroid receptor beta, peroxisome-proliferator-activated receptor gamma and orphan receptors, *Essays Biochem 40*, 169-189.

106. Brzozowski, A. M., Pike, A. C., Dauter, Z., Hubbard, R. E., Bonn, T., Engstrom, O., Ohman, L., Greene, G. L., Gustafsson, J. A., and Carlquist, M. (1997) Molecular basis of agonism and antagonism in the oestrogen receptor, *Nature 389*, 753-758.

107. Sonoda, J., Pei, L., and Evans, R. M. (2008) Nuclear receptors: decoding metabolic disease, *FEBS Lett 582*, 2-9.

108. Jensen, E. V., and Jacobson, H. (1962) Basic guides to the mechanism of estrogen action, *Recent Prog Horm Res 18*, 318-414.

109. Jensen, E. V. (1962) On the mechanism of estrogen action, *Perspect Biol Med 6*, 47-59.

110. Toft, D., and Gorski, J. (1966) A receptor molecule for estrogens: isolation from the rat uterus and preliminary characterization, *Proc Natl Acad Sci U S A 55*, 1574-1581.

111. Toft, D., Shyamala, G., and Gorski, J. (1967) A receptor molecule for estrogens: studies using a cell-free system, *Proc Natl Acad Sci U S A 57*, 1740-1743.

112. Jensen, E. V., and Jordan, V. C. (2003) The estrogen receptor: a model for molecular medicine, *Clin Cancer Res 9*, 1980-1989.

113. Kumar, V., Green, S., Stack, G., Berry, M., Jin, J. R., and Chambon, P. (1987) Functional domains of the human estrogen receptor, *Cell 51*, 941-951.

114. 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.

115. Fukada, H., and Takahashi, K. (1998) Enthalpy and heat capacity changes for the proton dissociation of various buffer components in 0.1 M potassium chloride, *Proteins 33*, 159-166.

116. Kozlov, A. G., and Lohman, T. M. (2000) Large contributions of coupled protonation equilibria to the observed enthalpy and heat capacity changes for ssDNA binding to Escherichia coli SSB protein, *Proteins Suppl 4*, 8-22.

117. Ortiz-Salmeron, E., Yassin, Z., Clemente-Jimenez, M. J., Las Heras-Vazquez, F. J., Rodriguez-Vico, F., Baron, C., and Garcia-Fuentes, L. (2001) Thermodynamic analysis of the binding of glutathione to glutathione S-transferase over a range of temperatures, *Eur J Biochem 268*, 4307-4314.

118. Lumry, R., and Rajender, S. (1970) Enthalpy-entropy compensation phenomena in water solutions of proteins and small molecules: a ubiquitous property of water, *Biopolymers 9*, 1125-1227.

119. Eftink, M. R., Anusiem, A. C., and Biltonen, R. L. (1983) Enthalpy-entropy compensation and heat capacity changes for protein-ligand interactions: general thermodynamic models and data for the binding of nucleotides to ribonuclease A, *Biochemistry 22*, 3884-3896.

120. Cooper, A., Johnson, C. M., Lakey, J. H., and Nollmann, M. (2001) Heat does not come in different colours: entropy-enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions, *Biophys Chem 93*, 215-230.

121. Sharp, K. (2001) Entropy-enthalpy compensation: fact or artifact?, *Protein Sci 10*, 661-667.

122. Starikov, E. B., and Norden, B. (2007) Enthalpy-entropy compensation: a phantom or something useful?, *J Phys Chem B 111*, 14431-14435.

123. Chen, Y. X., Du, J. T., Zhou, L. X., Liu, X. H., Zhao, Y. F., Nakanishi, H., and Li, Y. M. (2006) Alternative O-GlcNAcylation/O-phosphorylation of Ser16 induce different conformational disturbances to the N terminus of murine estrogen receptor beta, *Chem Biol 13*, 937-944.

124. Faus, H., and Haendler, B. (2006) Post-translational modifications of steroid receptors, *Biomed Pharmacother 60*, 520-528.

125. Popov, V. M., Wang, C., Shirley, L. A., Rosenberg, A., Li, S., Nevalainen, M., Fu, M., and Pestell, R. G. (2007) The functional significance of nuclear receptor acetylation, *Steroids 72*, 221-230.

126. Weigel, N. L., and Moore, N. L. (2007) Steroid receptor phosphorylation: a key modulator of multiple receptor functions, *Mol Endocrinol 21*, 2311-2319.

127. Boron, W. F. (2004) Regulation of intracellular pH, *Adv Physiol Educ 28*, 160-179.

128. Khaled, A. R., Moor, A. N., Li, A., Kim, K., Ferris, D. K., Muegge, K., Fisher, R. J., Fliegel, L., and Durum, S. K. (2001) Trophic factor withdrawal: p38 mitogen-activated protein kinase activates NHE1, which induces intracellular alkalinization, *Mol Cell Biol 21*, 7545-7557.

129. Khaled, A. R., Reynolds, D. A., Young, H. A., Thompson, C. B., Muegge, K., and Durum, S. K. (2001) Interleukin-3 withdrawal induces an early increase in mitochondrial membrane potential unrelated to the Bcl-2 family. Roles of intracellular pH, ADP transport, and F(0)F(1)-ATPase, *J Biol Chem 276*, 6453-6462.

130. Puceat, M., Roche, S., and Vassort, G. (1998) Src family tyrosine kinase regulates intracellular pH in cardiomyocytes, *J Cell Biol 141*, 1637-1646.

131. Harris, T. K., and Turner, G. J. (2002) Structural basis of perturbed pKa values of catalytic groups in enzyme active sites, *IUBMB Life 53*, 85-98.

132. Marini, J. C., Levene, S. D., Crothers, D. M., and Englund, P. T. (1982) Bent helical structure in kinetoplast DNA, *Proc Natl Acad Sci U S A 79*, 7664-7668.

133. Wu, H. M., and Crothers, D. M. (1984) The locus of sequence-directed and protein-induced DNA bending, *Nature 308*, 509-513.

134. Munteanu, M. G., Vlahovicek, K., Parthasarathy, S., Simon, I., and Pongor, S. (1998) Rod models of DNA: sequence-dependent anisotropic elastic modelling of local bending phenomena, *Trends Biochem Sci 23*, 341-347.

135. Deegan, B. J., Seldeen, K. L., McDonald, C. B., Bhat, V., and Farooq, A. (2010) Binding of the ERalpha nuclear receptor to DNA is coupled to proton uptake, *Biochemistry 49*, 5978-5988.

136. Laganiere, J., Deblois, G., Lefebvre, C., Bataille, A. R., Robert, F., and Giguere, V. (2005) Location analysis of estrogen receptor alpha target promoters reveals that FOXA1 defines a domain of the estrogen response, *Proc Natl Acad Sci U S A 102*, 11651-11656.

137. Karolchik, D., Baertsch, R., Diekhans, M., Furey, T. S., Hinrichs, A., Lu, Y. T., Roskin, K. M., Schwartz, M., Sugnet, C. W., Thomas, D. J., Weber, R. J., Haussler, D., and Kent, W. J. (2003) The UCSC Genome Browser Database, *Nucleic Acids Res 31*, 51-54.

138. Hinrichs, A. S., Karolchik, D., Baertsch, R., Barber, G. P., Bejerano, G., Clawson, H., Diekhans, M., Furey, T. S., Harte, R. A., Hsu, F., Hillman-Jackson, J., Kuhn, R. M., Pedersen, J. S., Pohl, A., Raney, B. J., Rosenbloom, K. R., Siepel, A., Smith, K. E., Sugnet, C. W., Sultan-Qurraie, A., Thomas, D. J., Trumbower, H., Weber, R. J., Weirauch, M., Zweig, A. S., Haussler, D., and Kent, W. J. (2006) The UCSC Genome Browser Database: update 2006, *Nucleic Acids Res 34*, D590-598.

139. Nardulli, A. M., and Shapiro, D. J. (1992) Binding of the estrogen receptor DNA-binding domain to the estrogen response element induces DNA bending, *Mol Cell Biol 12*, 2037-2042.

140. Seldeen, K. L., McDonald, C. B., Deegan, B. J., and Farooq, A. (2009) Single Nucleotide Variants of the TGACTCA Motif Modulate Energetics and Orientation of Binding of the Jun-Fos Heterodimeric Transcription Factor, *Biochemistry 48*, 1975-1983.

141. Klinge, C. M., Jernigan, S. C., Smith, S. L., Tyulmenkov, V. V., and Kulakosky, P. C. (2001) Estrogen response element sequence impacts the conformation and transcriptional activity of estrogen receptor alpha, *Mol Cell Endocrinol 174*, 151-166.

142. 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.

143. 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.

144. 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.

145. 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.

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

147. 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.

148. Privalov, P. L., Dragan, A. I., and Crane-Robinson, C. (2009) The cost of DNA bending, *Trends Biochem Sci 34*, 464-470.

149. Koo, H. S., Wu, H. M., and Crothers, D. M. (1986) DNA bending at adenine . thymine tracts, *Nature 320*, 501-506.

150. Alexeev, D. G., Lipanov, A. A., and Skuratovskii, I. (1987) Poly(dA).poly(dT) is a B-type double helix with a distinctively narrow minor groove, *Nature 325*, 821-823.

151. Arnott, S., Chandrasekaran, R., Hall, I. H., and Puigjaner, L. C. (1983) Heteronomous DNA, *Nucleic Acids Res 11*, 4141-4155.

152. Nelson, H. C., Finch, J. T., Luisi, B. F., and Klug, A. (1987) The structure of an oligo(dA).oligo(dT) tract and its biological implications, *Nature 330*, 221-226.

153. Coll, M., Frederick, C. A., Wang, A. H., and Rich, A. (1987) A bifurcated hydrogen-bonded conformation in the d(A.T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin, *Proc Natl Acad Sci U S A 84*, 8385-8389.

154. DiGabriele, A. D., Sanderson, M. R., and Steitz, T. A. (1989) Crystal lattice packing is important in determining the bend of a DNA dodecamer containing an adenine tract, *Proc Natl Acad Sci U S A 86*, 1816-1820.

155. Driscoll, M. D., Sathya, G., Muyan, M., Klinge, C. M., Hilf, R., and Bambara, R. A. (1998) Sequence requirements for estrogen receptor binding to estrogen response elements, *J Biol Chem 273*, 29321-29330.

156. Anolik, J. H., Klinge, C. M., Bambara, R. A., and Hilf, R. (1993) Differential impact of flanking sequences on estradiol- vs 4-hydroxytamoxifen-liganded estrogen receptor binding to estrogen responsive element DNA, *J Steroid Biochem Mol Biol 46*, 713-730.

157. Anolik, J. H., Klinge, C. M., Hilf, R., and Bambara, R. A. (1995) Cooperative binding of estrogen receptor to DNA depends on spacing of binding sites, flanking sequence, and ligand, *Biochemistry 34*, 2511-2520.

158. Anolik, J. H., Klinge, C. M., Brolly, C. L., Bambara, R. A., and Hilf, R. (1996) Stability of the ligand-estrogen receptor interaction depends on estrogen response element flanking sequences and cellular factors, *J Steroid Biochem Mol Biol 59*, 413-429.

159. Miller, J., McLachlan, A. D., and Klug, A. (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes, *Embo J 4*, 1609-1614.

160. Green, A., Parker, M., Conte, D., and Sarkar, B. (1998) Zinc Finger Proteins: A Bridge Between Transition Metals and Gene Regulation, *J Trace Element Exp Med 11*, 103-118.

161. Wolfe, S. A., Nekludova, L., and Pabo, C. O. (2000) DNA recognition by Cys2His2 zinc finger proteins, *Annu Rev Biophys Biomol Struct 29*, 183-212.

162. Thiesen, H. J., and Bach, C. (1991) Transition metals modulate DNA-protein interactions of SP1 zinc finger domains with its cognate target site, *Biochem Biophys Res Commun 176*, 551-557.

163. Freedman, L. P., Luisi, B. F., Korszun, Z. R., Basavappa, R., Sigler, P. B., and Yamamoto, K. R. (1988) The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain, *Nature 334*, 543-546.

164. Hartwig, A., Schwerdtle, T., and Bal, W. (2010) Biophysical analysis of the interaction of toxic metal ions and oxidants with the zinc finger domain of XPA, *Methods Mol Biol 649*, 399-410.

165. Blessing, H., Kraus, S., Heindl, P., Bal, W., and Hartwig, A. (2004) Interaction of selenium compounds with zinc finger proteins involved in DNA repair, *Eur J Biochem 271*, 3190-3199.

166. Deegan, B. J., Bhat, V., Seldeen, K. L., McDonald, C. B., and Farooq, A. (2011) Genetic variations within the ERE motif modulate plasticity and energetics of binding of DNA to the ERalpha nuclear receptor, *Arch Biochem Biophys*, *507*, 262-270.

167. 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.

168. Cooper, A. (2005) Heat capacity effects in protein folding and ligand binding: a re-evaluation of the role of water in biomolecular thermodynamics, *Biophys Chem 115*, 89-97.

169. Rulisek, L., and Vondrasek, J. (1998) Coordination geometries of selected transition metal ions (Co2+, Ni2+, Cu2+, Zn2+, Cd2+, and Hg2+) in metalloproteins, *J Inorg Biochem 71*, 115-127.

170. Marcus, Y. (1988) Ionic Radii in Aqueous Solutions, *Chem. Rev. 88*, 1475-1498.

171. Brandt, M. E., and Vickery, L. E. (1997) Cooperativity and dimerization of recombinant human estrogen receptor hormone-binding domain, *J Biol Chem 272*, 4843-4849.

172. Notides, A. C., Lerner, N., and Hamilton, D. E. (1981) Positive cooperativity of the estrogen receptor, *Proc Natl Acad Sci U S A 78*, 4926-4930.

173. Notides, A. C., and Nielsen, S. (1974) The molecular mechanism of the in vitro 4 S to 5 S transformation of the uterine estrogen receptor, *J Biol Chem 249*, 1866-1873.

174. Hutchens, T. W., Allen, M. H., Li, C. M., and Yip, T. T. (1992) Occupancy of a C2-C2 type 'zinc-finger' protein domain by copper. Direct observation by electrospray ionization mass spectrometry, *FEBS Lett 309*, 170-174.

175. Young, P. C., Cleary, R. E., and Ragan, W. D. (1977) Effect of metal ions on the binding of 17beta-estradiol to human endometrial cytosol, *Fertil Steril 28*, 459-463.

176. Fishman, J. H., and Fishman, J. (1988) Copper and endogenous mediators of estradiol action, *Biochem Biophys Res Commun 152*, 783-788.

177. Darbre, P. D. (2006) Metalloestrogens: an emerging class of inorganic xenoestrogens with potential to add to the oestrogenic burden of the human breast, *J Appl Toxicol 26*, 191-197.

178. Martin, M. B., Reiter, R., Pham, T., Avellanet, Y. R., Camara, J., Lahm, M., Pentecost, E., Pratap, K., Gilmore, B. A., Divekar, S., Dagata, R. S., Bull, J. L., and Stoica, A. (2003) Estrogen-like activity of metals in MCF-7 breast cancer cells, *Endocrinology 144*, 2425-2436.

179. Garcia-Morales, P., Saceda, M., Kenney, N., Kim, N., Salomon, D. S., Gottardis, M. M., Solomon, H. B., Sholler, P. F., Jordan, V. C., and Martin, M. B. (1994) Effect of cadmium on estrogen receptor levels and estrogen-induced responses in human breast cancer cells, *J Biol Chem 269*, 16896-16901.

180. Tyulmenkov, V. V., and Klinge, C. M. (2001) A mathematical approach to predict the affinity of estrogen receptors alpha and beta binding to DNA, *Mol Cell Endocrinol 182*, 109-119.