Archives of Biochemistry and Biophysics 507 (2011) 262–270

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

Genetic variations within the ERE motif modulate plasticity and energetics of binding of DNA to the ERα nuclear receptor

Brian J. Deegan, Vikas Bhat, Kenneth L. Seldeen, Caleb B. McDonald, Amjad Farooq *

Department of Biochemistry and Molecular Biology and USylvester Braman Family Breast Cancer Institute, Leonard Miller School of Medicine, University of Miami, Miami, FL 33136, United States

ARTICLE INFO

Article history:
Received 26 November 2010
and in revised form 3 January 2011
Available online 7 January 2011

Keywords:
Nuclear receptors
Estrogen receptor α
Estrogen response element
Isothermal titration calorimetry
Circular dichroism
Molecular modeling

ABSTRACT

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 atomic level. Our 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.

Introduction

Nuclear receptors (NRs)[1] act as ligand-modulated transcription factors and orchestrate a plethora of cellular functions central to health and disease [1–4]. 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 [5–7]. 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, 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 [8–10]. While the DB domain recognizes specific promoter response elements within target genes, 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 [11,12]. 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 the MAPK signaling and thus it may further synergize the action of various co-activators and co-repressors recruited by the LB domain at the site of DNA transcription [13,14]. In this manner, nuclear receptors mediate a diverse array of cellular functions from embryonic development to metabolic homeostasis and their aberrant function has been widely implicated in disease [15–19,7].

ERα mediates the action of estrogens such as estradiol and its hyperactivation leads to the genesis of large fractions of breast cancer [20–26]. The DB domain of ERα binds as a homodimer with a twofold axis of symmetry to the ERE motif, containing the AGGTCAnnnTGACCT consensus sequence, located within the promoters of target genes [27]. 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 [28,29]. It is important to note that

* Corresponding author. Fax: +1 305 243 3955.
E-mail address: amjad@farooqlab.net (A. Farooq).

[1] *Abbreviations used:* AR, androgen receptor; CD, circular dichroism; DB, DNA-binding; ERα, estrogen receptor α; ERE, estrogen response element; GR, glucocorticoid receptor; ITC, isothermal titration calorimetry; LB, ligand-binding; MAPK, mitogen-activated protein kinase; MM, molecular modeling; NMR, nuclear magnetic resonance; NR, nuclear receptor; PR, progesterone receptor; SEC, size-exclusion chromatography; TA, trans-activation; TB, terrific broth; Trx, thioredoxin; XRC, X-ray crystallography; ZF, zinc finger.

0003-9861/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.abb.2011.01.002

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, while 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 [27]. 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) [30–32], 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 [33–36]. 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.

Materials and methods

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 50 μM $ZnCl_2$ to an optical density of 0.5 at 600 nm prior to induction with 0.5 mM 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 20 mM imidazole to remove non-specific binding of bacterial proteins to the column. The recombinant protein was subsequently eluted with 200 mM 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 and 10 mg 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 [37]. Results from both methods were in an excellent agreement.

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 Fig. 1. Oligo concentrations were determined spectrophotometrically on the basis of their extinction coefficients derived from their nucleotide sequences using the online software OligoAnalyzer 3.0 (Integrated DNA Technologies) based on the nearest-neighbor model [38]. Double-stranded DNA (dsDNA) oligos were generated as described earlier [39].

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 50 mM sodium phosphate containing 5 mM β-mercaptoethanol at pH 7.0 and degassed using the ThermoVac accessory for 5 min. The experiments were initiated by injecting 25 × 10 μl aliquots of 50–200 μM of a dsDNA oligo containing the ERE motif, or a variant thereof, from the syringe into the calorimetric cell containing 1.8 ml of 5–10 μM 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 [40,39].

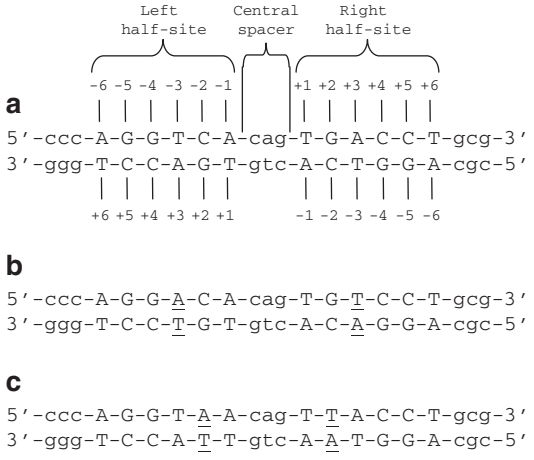

Fig. 1. Nucleotide sequence of dsDNA oligos. (a) Consensus ERE motif. The AGGTCA and TGACCT half-sites are clearly marked. The consensus nucleotides within each half-site are capitalized while the flanking nucleotides and the intervening nucleotides within the central spacer are shown in small letters. 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.

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 a 20 μM of dsDNA oligos containing the ERE motif, or a variant thereof, in 50 mM sodium phosphate at pH 7.0. All experiments were conducted in a quartz cuvette with a 2-mm pathlength in the wavelength range 190–310 nm. Data were recorded with a slit bandwidth of 2 nm at a scan rate of 3 nm/min. Data were normalized against reference spectra to remove the contribution of buffer. The reference spectra were obtained in a similar manner on a 50 mM sodium phosphate at pH 7.0. Each data set represents an average of at least four scans acquired at 1 nm intervals. Data were converted to molar ellipticity, $[\theta]$, as a function of wavelength (λ) of electromagnetic radiation using the equation:

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

where $\Delta\varepsilon$ is the observed ellipticity in mdeg, c is the dsDNA concentration in μM and l is the cuvette pathlength in cm.

Molecular 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 [41]. 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# 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 [42].

Results and discussion

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-site in a symmetrical manner using ITC (Fig. 2 and Table 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 [27]. 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-DNA interface [29]. 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 do-

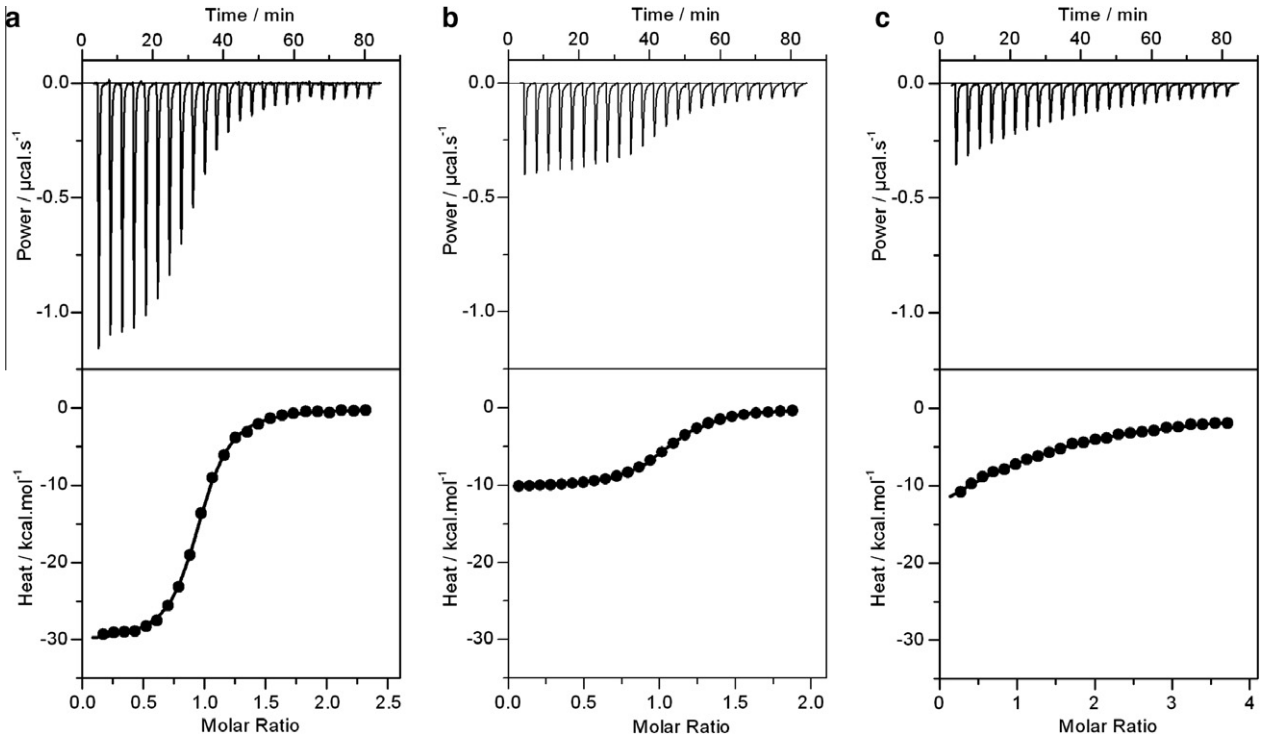

Fig. 2. Representative ITC isotherms for the binding of DB domain of ERα to dsDNA oligos containing the consensus ERE motif (a), the variant Am3Tp3 motif (b) and the variant Am2Tp2 motif (c). The upper panels show the raw ITC data expressed as change in thermal power with respect to time over the period 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 [40,39].

Table 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})	$T\Delta S$ (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	AGGTC**C**cag**G**GACCT	0.79 ± 0.01	−18.25 ± 0.25	−9.91 ± 0.25	−8.34 ± 0.01	Signaling protein WISP2
Gm1Cp1	AGGTC**G**cag**C**GACCT	1.58 ± 0.05	−22.76 ± 0.28	−14.83 ± 0.26	−7.93 ± 0.02	Hemoprotein CYP1B1
Tm1Ap1	AGGTC**T**cag**A**GACCT	0.85 ± 0.06	−22.59 ± 1.08	−14.29 ± 1.08	−8.30 ± 0.04	Transctiption factor HNF3A
Am2Tp2	AGGT**A**AcagT**T**ACCT	7.02 ± 0.04	−25.57 ± 0.03	−18.53 ± 0.03	−7.04 ± 0.01	Nuclear receptor ERα
Gm2Cp2	AGGT**G**AcagT**C**ACCT	2.31 ± 0.11	−20.06 ± 0.22	−12.35 ± 0.19	−7.70 ± 0.03	Mitotic protein PRCC
Tm2Ap2	AGGT**T**AcagT**A**ACCT	0.11 ± 0.01	−24.36 ± 0.42	−14.83 ± 0.45	−9.53 ± 0.02	Nuclear receptor SHP
Am3Tp3	AGG**A**CAcagTG**T**CCT	0.26 ± 0.08	−11.48 ± 0.53	−2.47 ± 0.35	−9.01 ± 0.18	Protease inhibitor CST5
Cm3Gp3	AGG**C**CAcagTG**G**CCT	0.86 ± 0.04	−19.58 ± 0.28	−11.29 ± 0.30	−8.28 ± 0.03	Cytoskeletal protein TNS1
Gm3Cp3	AGG**G**CAcagTG**C**CCT	0.50 ± 0.02	−20.06 ± 0.44	−11.45 ± 0.41	−8.61 ± 0.02	Retinoblastoma protein RBL2
Am4Tp4	AG**A**TCAcagTGA**T**CT	0.31 ± 0.01	−19.20 ± 0.17	−10.30 ± 0.14	−8.89 ± 0.01	Skeletal muscle protein MYOT
Cm4Gp4	AG**C**TCAcagTGA**G**CT	2.56 ± 1.01	−11.77 ± 0.41	−4.85 ± 0.20	−7.66 ± 0.24	Cytoskeletal protein EPS8
Tm4Ap4	AG**T**TCAcagTGA**A**CT	1.17 ± 0.04	−23.87 ± 0.68	−15.76 ± 0.66	−8.10 ± 0.02	GTPase Rab7L1
Am5Tp5	A**A**GTCAcagTGAC**T**T	0.50 ± 0.15	−17.70 ± 0.49	−9.08 ± 0.31	−8.62 ± 0.18	Cytokine IL20
Cm5Gp5	A**C**GTCAcagTGAC**G**T	4.09 ± 0.19	−21.51 ± 0.20	−14.15 ± 0.17	−7.36 ± 0.03	DNA helicase RecQ4
Tm5Ap5	A**T**GTCAcagTGAC**A**T	0.40 ± 0.01	−21.86 ± 0.87	−13.12 ± 0.88	−8.75 ± 0.01	MDR protein MRP5
Cm6Gp6	**C**GGTCAcagTGACC**G**	0.14 ± 0.01	−22.69 ± 0.52	−13.34 ± 0.54	−9.35 ± 0.02	Transcription factor PROP1
Gm6Cp6	**G**GGTCAcagTGACC**C**	0.15 ± 0.01	−26.10 ± 0.20	−16.76 ± 0.17	−9.34 ± 0.03	Caspase CASP7
Tm6Ap6	**T**GGTCAcagTGACC**A**	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 Fig. 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 [63–65]. 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 [40,39]. 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 $T\Delta S = \Delta H - \Delta G$. Binding stoichiometries generally agreed to within ±10%. Errors were calculated from at least three independent measurements. All errors are given to one standard deviation.

main [43], 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 (Fig. 3). Thus, while the binding of Am3Tp3 variant motif to the DB domain is only weaker by about threefold 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 the promoters of target genes 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 orientation [44]. Given that ERα also cooperates with ERβ in binding as a heterodimer to the promoters of target genes [45], 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.

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 (Fig. 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 ($T\Delta S$) contributions for the binding of ERE motif and its variants thereof to the DB domain show poor correlation with the overall free energy (ΔG) (Fig. 4b and c). For example, an increase in favorable ΔH or a decrease in unfavorable $T\Delta S$ 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 [46–52], 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 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 harbor-

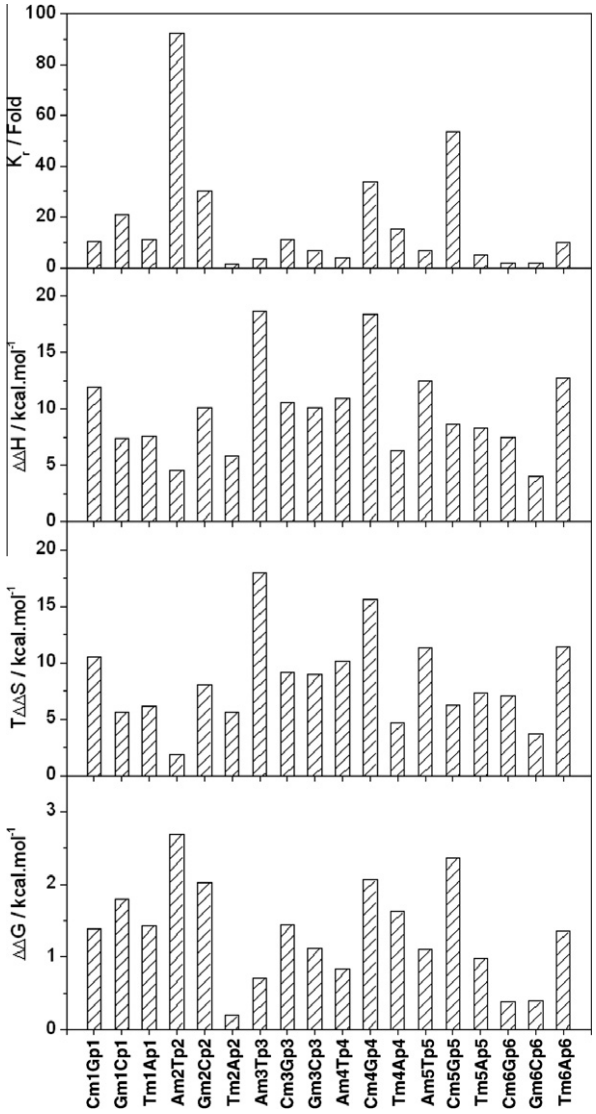

Fig. 3. Analysis of the binding of DB domain of ERα to variant motifs relative to the consensus ERE motif in terms of relative binding affinity (K_r), relative enthalpic contribution ($\Delta\Delta H$) and relative entropic contribution ($T\Delta\Delta S$) to the relative free energy ($\Delta\Delta G$). 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 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 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 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 1).

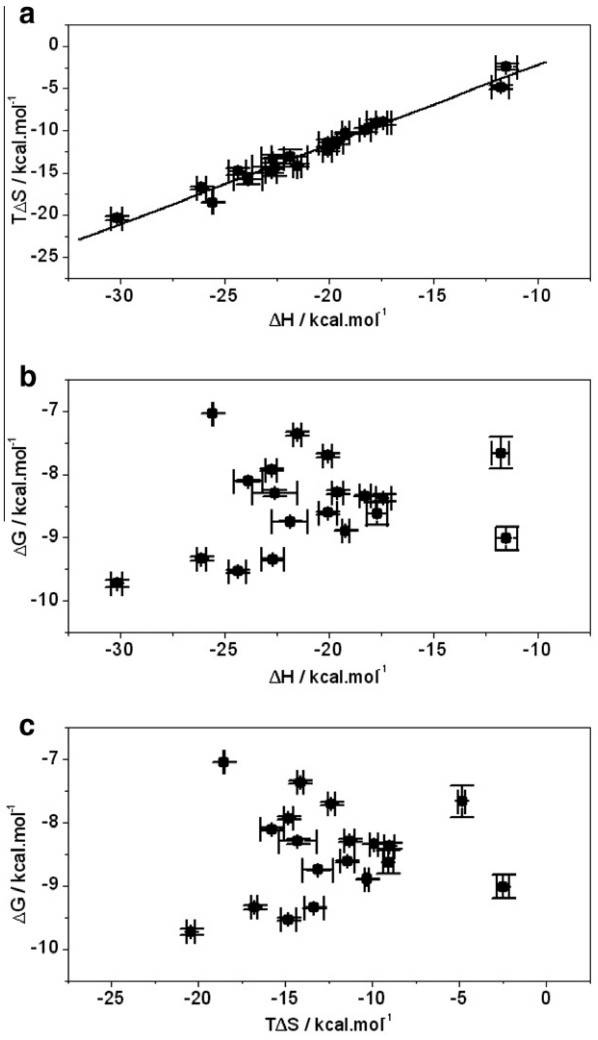

Fig. 4. Interdependence of enthalpic (ΔH) and entropic ($T\Delta S$) contributions to the free energy (ΔG) for the binding of DB domain of ERα to dsDNA oligos containing the consensus ERE and variant motifs. (a) $T\Delta S$ vs. ΔH plot. (b) ΔG vs. ΔH plot. (c) ΔG vs. $T\Delta S$ plot. Note that the solid 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.

ing 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 $T\Delta S$, enthalpically-optimized drugs may not necessarily be effective universally and that entropically-optimized drugs may offer a viable alternative.

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 Fig. 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 [29], 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 note worthy that the optimal contribution of underlying enthalpic and entropic contributions to the overall free energy does not correlate with the least and most tolerable

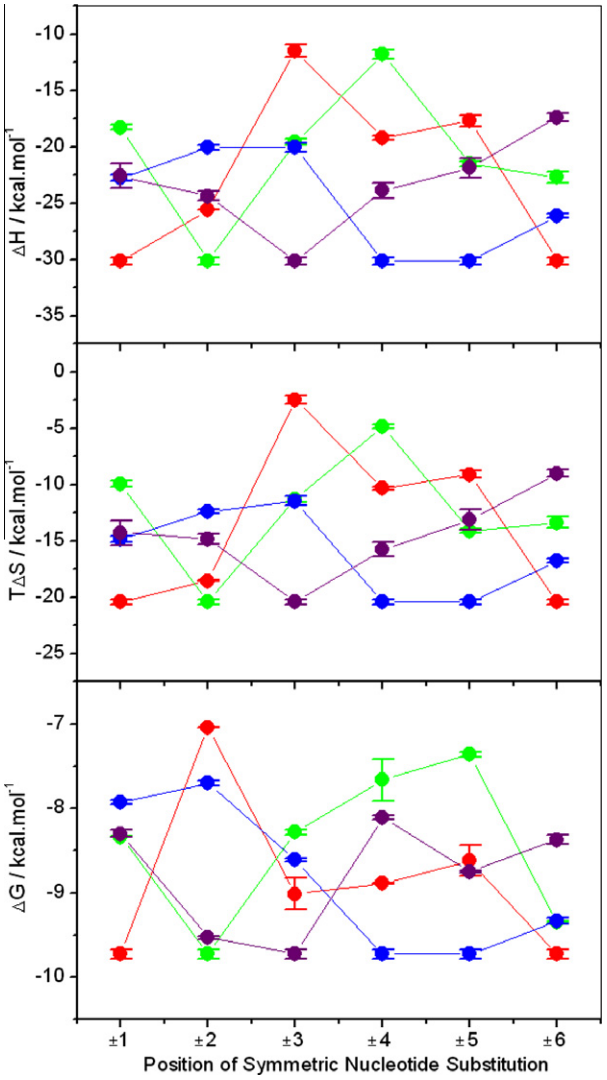

Fig. 5. Dependence of free energy ($\Delta\Delta G$) and the underlying enthalpic (ΔH) and entropic ($T\Delta S$) contributions on the position of symmetric nucleotide substitution within both half-sites of the consensus ERE and variant motifs for the binding of DB domain of ERα. The changes in various 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. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

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 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 Fig. 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.

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 [30,53,31,54–58]. 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

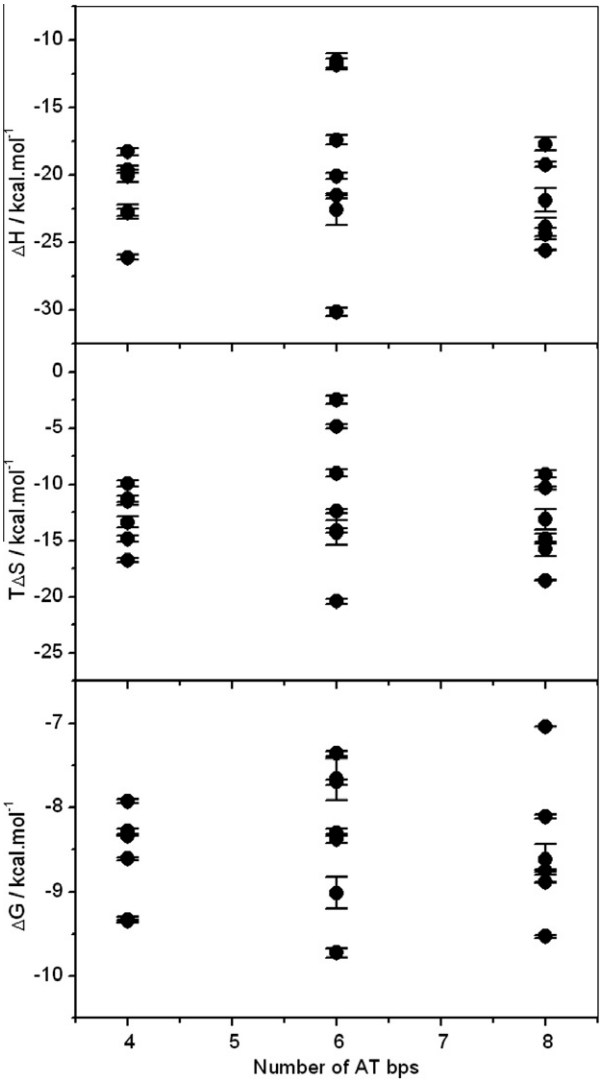

Fig. 6. Dependence of free energy (ΔG) and the underlying enthalpic (ΔH) and entropic ($T\Delta S$) contributions on the total number of AT bps within both half-sites of the consensus ERE and variant motifs for the binding of DB domain of ERα.

pairs are held together by only two hydrogen bonds in lieu of three formed between G and C base pairs. Given that the binding of DB domain of ERα to the ERE motif results in the bending of DNA [43], 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 Fig. 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.

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 [30–32]. 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 (Fig. 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 195, 220 and 280 nm. 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 220 and 280 nm show remarkable heterogeneity and lie as much as more than 10 nm 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 motif tightly governs its conformational flexibility and that the varying flexibility is likely to be an integral feature of their ability to bind to DB domain of ERα with distinct underlying energetics (Table 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.

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 the consensus ERE motif and the Am2Tp2 variant motif (Fig. 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 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 [29], the DB domain binds as a homodimer with a twofold axis of symmetry within the major grooves of DNA duplex. One monomer contacts the antisense strand within the first half-site of the DNA

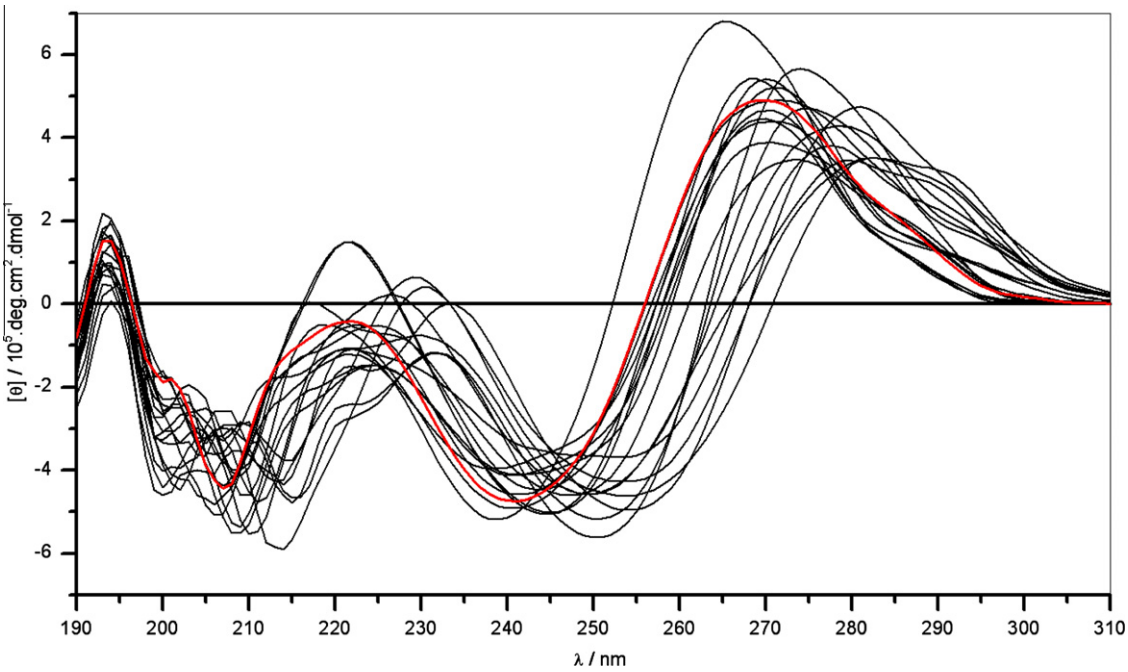

Fig. 7. Representative CD spectra of dsDNA oligos containing the consensus ERE and variant motifs. The spectrum of consensus ERE motif (red) is superimposed onto spectra of variant motifs (black). (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

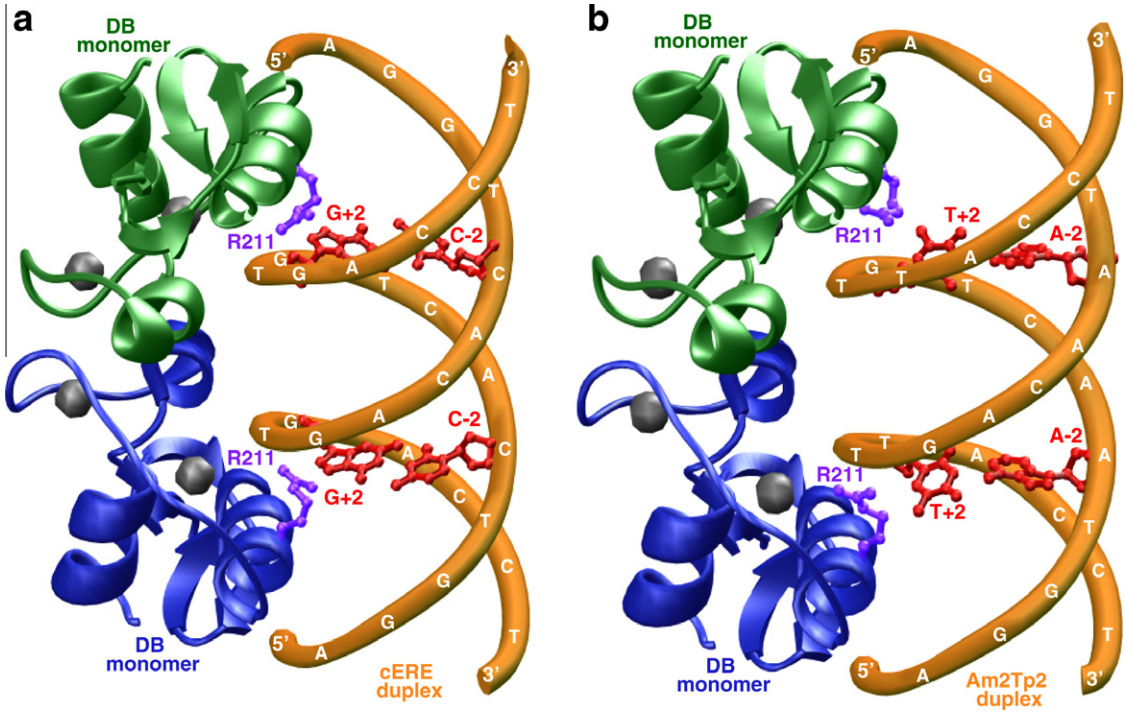

Fig. 8. 3D atomic models of the DB domain of ERα in complex with dsDNA oligos containing the consensus ERE motif (a) and the variant Am2Tp2 motif (b). Note that the DB domain binds to DNA as a homodimer with a twofold axis of symmetry. One monomer of the DB domain is shown in green and the other in blue. 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. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

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.

Conclusions

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 [27]. Despite the knowledge that such genetic variations within the promoters of target genes modulate the transcriptional activity of nuclear receptors [33–36], 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 rationale drug design. Although it is widely believed that AT sequences within DNA account for its intrinsic conformational flexibility such as bending and curvature [30,53,31,54–58], 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 [59–62]. 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.

Acknowledgments

This work was supported by funds from the National Institutes of Health (Grant# R01-GM083897) and the USylvester Braman Family Breast Cancer Institute to A.F. C.B.M. is a recipient of a postdoctoral fellowship from the National Institutes of Health (Award# T32-CA119929). B.J.D. and A.F. are members of the Sheila and David Fuente Graduate Program in Cancer Biology at the Sylvester Comprehensive Cancer Center of the University of Miami.

References

[1] R.M. Evans, Science 240 (1988) 889–895.
[2] J.W. Thornton, Proc. Natl. Acad. Sci. USA 98 (2001) 5671–5676.
[3] H. Escriva, S. Bertrand, V. Laudet, Essays Biochem. 40 (2004) 11–26.
[4] N.J. McKenna, A.J. Cooney, F.J. DeMayo, M. Downes, C.K. Glass, R.B. Lanz, M.A. Lazar, D.J. Mangelsdorf, D.D. Moore, J. Qin, D.L. Steffen, M.J. Tsai, S.Y. Tsai, R. Yu, R.N. Margolis, R.M. Evans, B.W. O'Malley, Mol. Endocrinol. 23 (2009) 740–746.
[5] R. Kumar, E.B. Thompson, Steroids 64 (1999) 310–319.
[6] P. Barnett, H.F. Tabak, E.H. Hettema, Trends Biochem. Sci. 25 (2000) 227–228.
[7] I.J. McEwan, Methods Mol. Biol. 505 (2009) 3–18.
[8] S. Green, P. Chambon, Trends Genet. 4 (1988) 309–314.
[9] P.F. Egea, B.P. Klaholz, D. Moras, FEBS Lett. 476 (2000) 62–67.
[10] F. Claessens, D.T. Gewirth, Essays Biochem. 40 (2004) 59–72.
[11] J. Ham, M.G. Parker, Curr. Opin. Cell Biol. 1 (1989) 503–511.
[12] B.D. Darimont, R.L. Wagner, J.W. Apriletti, M.R. Stallcup, P.J. Kushner, J.D. Baxter, R.J. Fletterick, K.R. Yamamoto, Genes Dev. 12 (1998) 3343–3356.
[13] S. Kato, H. Endoh, Y. Masuhiro, T. Kitamoto, S. Uchiyama, H. Sasaki, S. Masushige, Y. Gotoh, E. Nishida, H. Kawashima, D. Metzger, P. Chambon, Science 270 (1995) 1491–1494.
[14] A. Warnmark, E. Treuter, A.P. Wright, J.A. Gustafsson, Mol. Endocrinol. 17 (2003) 1901–1909.
[15] A.M. Brzozowski, A.C. Pike, Z. Dauter, R.E. Hubbard, T. Bonn, O. Engstrom, L. Ohman, G.L. Greene, J.A. Gustafsson, M. Carlquist, Nature 389 (1997) 753–758.
[16] B. Gottlieb, L.K. Beitel, J. Wu, Y.A. Elhaji, M. Trifiro, Essays Biochem. 40 (2004) 121–136.
[17] M. Gurnell, V.K. Chatterjee, Essays Biochem. 40 (2004) 169–189.
[18] N. Noy, Biochemistry 46 (2007) 13461–13467.
[19] J. Sonoda, L. Pei, R.M. Evans, FEBS Lett. 582 (2008) 2–9.
[20] E.V. Jensen, Perspect. Biol. Med. 6 (1962) 47–59.
[21] E.V. Jensen, H. Jacobson, Recent Prog. Horm. Res. 18 (1962) 318–414.
[22] D. Toft, J. Gorski, Proc. Natl. Acad. Sci. USA 55 (1966) 1574–1581.
[23] D. Toft, G. Shyamala, J. Gorski, Proc. Natl. Acad. Sci. USA 57 (1967) 1740–1743.
[24] V. Kumar, S. Green, G. Stack, M. Berry, J.R. Jin, P. Chambon, Cell 51 (1987) 941–951.
[25] E.V. Jensen, V.C. Jordan, Clin. Cancer Res. 9 (2003) 1980–1989.
[26] N. Heldring, A. Pike, S. Andersson, J. Matthews, G. Cheng, J. Hartman, M. Tujague, A. Strom, E. Treuter, M. Warner, J.A. Gustafsson, Physiol. Rev. 87 (2007) 905–931.
[27] C.M. Klinge, Nucleic Acids Res. 29 (2001) 2905–2919.
[28] J.W. Schwabe, D. Neuhaus, D. Rhodes, Nature 348 (1990) 458–461.
[29] J.W. Schwabe, L. Chapman, J.T. Finch, D. Rhodes, Cell 75 (1993) 567–578.
[30] J.C. Marini, S.D. Levene, D.M. Crothers, P.T. Englund, Proc. Natl. Acad. Sci USA 79 (1982) 7664–7668.
[31] H.M. Wu, D.M. Crothers, Nature 308 (1984) 509–513.
[32] M.G. Munteanu, K. Vlahovicek, S. Parthasarathy, I. Simon, S. Pongor, Trends Biochem. Sci. 23 (1998) 341–347.
[33] S.K. Nordeen, B.J. Suh, B. Kuhnel, C.A. Hutchison 3rd, Mol. Endocrinol. 4 (1990) 1866–1873.
[34] B.A. Lieberman, B.J. Bona, D.P. Edwards, S.K. Nordeen, Mol. Endocrinol. 7 (1993) 515–527.
[35] S.K. Nordeen, C.A. Ogden, L. Taraseviciene, B.A. Lieberman, Mol. Endocrinol. 12 (1998) 891–898.
[36] C.C. Nelson, S.C. Hendy, R.J. Shukin, H. Cheng, N. Bruchovsky, B.F. Koop, P.S. Rennie, Mol. Endocrinol. 13 (1999) 2090–2107.
[37] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, in: J.M. Walker (Ed.), The Proteomics Protocols Handbook, Humana Press, Totowa, New Jersey, USA, 2005, pp. 571–607.
[38] C.R. Cantor, M.M. Warshaw, H. Shapiro, Biopolymers 9 (1970) 1059–1077.
[39] B.J. Deegan, K.L. Seldeen, C.B. McDonald, V. Bhat, A. Farooq, Biochemistry 49 (2010) 5978–5988.
[40] T. Wiseman, S. Williston, J.F. Brandts, L.N. Lin, Anal. Biochem. 179 (1989) 131–137.
[41] M.A. Marti-Renom, A.C. Stuart, A. Fiser, R. Sanchez, F. Melo, A. Sali, Annu. Rev. Biophys. Biomol. Struct. 29 (2000) 291–325.
[42] M. Carson, J. Appl. Crystallogr. 24 (1991) 958–961.
[43] A.M. Nardulli, D.J. Shapiro, Mol. Cell. Biol. 12 (1992) 2037–2042.
[44] K.L. Seldeen, C.B. McDonald, B.J. Deegan, A. Farooq, Biochemistry 48 (2009) 1975–1983.
[45] C.M. Klinge, S.C. Jernigan, S.L. Smith, V.V. Tyulmenkov, P.C. Kulakosky, Mol. Cell. Endocrinol. 174 (2001) 151–166.
[46] D. Foguel, J.L. Silva, Proc. Natl. Acad. Sci. USA 91 (1994) 8244–8247.
[47] J.E. Ladbury, J.G. Wright, J.M. Sturtevant, P.B. Sigler, J. Mol. Biol. 238 (1994) 669–681.
[48] E. Merabet, G.K. Ackers, Biochemistry 34 (1995) 8554–8563.
[49] V. Petri, M. Hsieh, M. Brenowitz, Biochemistry 34 (1995) 9977–9984.
[50] C. Berger, I. Jelesarov, H.R. Bosshard, Biochemistry 35 (1996) 14984–14991.
[51] S. Milev, H.R. Bosshard, I. Jelesarov, Biochemistry 44 (2005) 285–293.
[52] P.L. Privalov, A.I. Dragan, C. Crane-Robinson, Trends Biochem. Sci. 34 (2009) 464–470.
[53] S. Arnott, R. Chandrasekaran, I.H. Hall, L.C. Puigjaner, Nucleic Acids Res. 11 (1983) 4141–4155.
[54] H.S. Koo, H.M. Wu, D.M. Crothers, Nature 320 (1986) 501–506.
[55] D.G. Alexeev, A.A. Lipanov, I. Skuratovskii, Nature 325 (1987) 821–823.
[56] M. Coll, C.A. Frederick, A.H. Wang, A. Rich, Proc. Natl. Acad. Sci. USA 84 (1987) 8385–8389.
[57] H.C. Nelson, J.T. Finch, B.F. Luisi, A. Klug, Nature 330 (1987) 221–226.
[58] A.D. DiGabriele, M.R. Sanderson, T.A. Steitz, Proc. Natl. Acad. Sci USA 86 (1989) 1816–1820.
[59] J.H. Anolik, C.M. Klinge, R.A. Bambara, R. Hilf, J. Steroid Biochem. Mol. Biol. 46 (1993) 713–730.
[60] J.H. Anolik, C.M. Klinge, R. Hilf, R.A. Bambara, Biochemistry 34 (1995) 2511–2520.
[61] J.H. Anolik, C.M. Klinge, C.L. Brolly, R.A. Bambara, R. Hilf, J. Steroid Biochem. Mol. Biol. 59 (1996) 413–429.
[62] M.D. Driscoll, G. Sathya, M. Muyan, C.M. Klinge, R. Hilf, R.A. Bambara, J. Biol. Chem. 273 (1998) 29321–29330.
[63] D. Karolchik, R. Baertsch, M. Diekhans, T.S. Furey, A. Hinrichs, Y.T. Lu, K.M. Roskin, M. Schwartz, C.W. Sugnet, D.J. Thomas, R.J. Weber, D. Haussler, W.J. Kent, Nucleic Acids Res. 31 (2003) 51–54.
[64] J. Laganiere, G. Deblois, C. Lefebvre, A.R. Bataille, F. Robert, V. Giguere, Proc. Natl. Acad. Sci. USA 102 (2005) 11651–11656.
[65] A.S. Hinrichs, D. Karolchik, R. Baertsch, G.P. Barber, G. Bejerano, H. Clawson, M. Diekhans, T.S. Furey, R.A. Harte, F. Hsu, J. Hillman-Jackson, R.M. Kuhn, J.S. Pedersen, A. Pohl, B.J. Raney, K.R. Rosenbloom, A. Siepel, K.E. Smith, C.W. Sugnet, A. Sultan-Qurraie, D.J. Thomas, H. Trumbower, R.J. Weber, M. Weirauch, A.S. Zweig, D. Haussler, W.J. Kent, Nucleic Acids Res. 34 (2006) D590–D598.