Eur Biophys J (1999) 28: 102–111

ARTICLE

Mats A. L. Eriksson · Lennart Nilsson

Structural and dynamic differences of the estrogen receptor DNA-binding domain, binding as a dimer and as a monomer to DNA: molecular dynamics simulation studies

Received: 6 March 1998 / Revised version: 22 June 1998 / Accepted: 2 September 1998

Abstract Molecular dynamics (MD) simulations of the estrogen receptor DNA-binding domain (ERDBD) as a dimer in complex with its DNA response element (ERE) show a significant difference in both structure and dynamics, compared to a MD simulation of monomeric ERDBD bound to its half-site response element (EREH). The C-terminal zinc binding domain (Zn_{II}), including a region (helix II) which is in a helical conformation in ERE-$(ERDBD)_2$, is considerably more flexible in EREH-ERDBD than in the dimeric complex. In EREH-ERDBD, all helical hydrogen bonds in helix II are broken and the entire Zn_{II} region is detached from a hydrogen bonding network that in ERE-$(ERDBD)_2$ connects to other parts of the protein as well as to the DNA. The regions that become flexible in EREH-ERDBD are identical to the regions where the NMR solution structure of free ERDBD is poorly ordered. This strongly suggests that dimerisation of ERDBD is required for ordering of the Zn_{II} region and that monomeric binding to DNA is not sufficient for the ordering. This contrasts to the glucocorticoid receptor DNA-binding domain (GRDBD) which has essentially the same mobility (uniform and limited), regardless of whether it is free as a monomer in solution, bound as a monomer to its half-site response element or in a dimeric complex with the full response element. The hydrogen bonding network that connects Zn_{II} with other parts of the protein and to DNA is almost identical in ERDBD and GRDBD. However, in GRDBD there is also a serine (in the N-terminal zinc coordinating region) with a central role in this network, connecting to the Zn_{II} region. This serine is replaced by a glycine in ERDBD and we suggest that this substitution is sufficient for destabilisation of the network, thus leading to a more flexible Zn_{II} region, which becomes ordered first upon forming a complex with another ERDBD and DNA.

Key words Transcription factors · Glucocorticoid receptor · DNA response elements · Protein-DNA complexes · Cooperative binding

M. A. L. Eriksson [1] · L. Nilsson (✉)
Department of Bioscience at NOVUM, Karolinska Institutet, S-14157 Huddinge, Sweden
e-mail: lennart.nilsson@csb.ki.se

Present address:
[1] Department of Biochemistry, University of Stockholm, S-10691 Stockholm, Sweden

Introduction

The estrogen receptor (ER) belongs to the superfamily of ligand-inducible nuclear receptors transcription factors that also includes the glucocorticoid receptor (GR), thyroid hormone receptors as well as the retinoic acid and vitamin D_3 receptors (Tsai and O'Malley 1994). The ER and GR bind cooperatively (Dahlman-Wright et al. 1990; Härd et al. 1990a; Schwabe et al. 1993b) as homodimers to palindromic DNA-binding sites (response elements, RE), composed of two six-base-pair half-sites, which are separated by a three-base-pair spacer region (Klock et al. 1987; Martinez et al. 1987; Strähle et al. 1987; Beato 1989; Beato et al. 1989; Anolik et al. 1993). The consensus response elements for ER and GR (ERE and GRE, respectively) (Beato et al. 1989; Anolik et al. 1993) differ only in two base pairs per half site, with DNA sequences TGACCT and TGTTCT, respectively (Fig. 1).

The DNA-binding domain [DBD, reviewed in Härd and Gustafsson (1993)] is highly conserved within the family and consists of two subdomains, each composed of a zinc binding domain (Zn_I and Zn_{II}, respectively), an α-helical region (helix I and III, respectively) and a following extended region (ext_I and ext_{II}, respectively, see Fig. 1). The first (N-terminal) subdomain is primarily involved in DNA-DBD interactions, with helix I (the recognition helix) positioned in the DNA major groove and forming the base-specific interactions with DNA. Helix I contains three residues [the "P-box" (Umesono and Evans 1989)] that are essential for the discrimination of ERE and GRE (Danielsen et al. 1989; Mader et al. 1989; Umesono and Evans 1989; Zilliacus et al. 1991, 1992, 1994): Glu25,

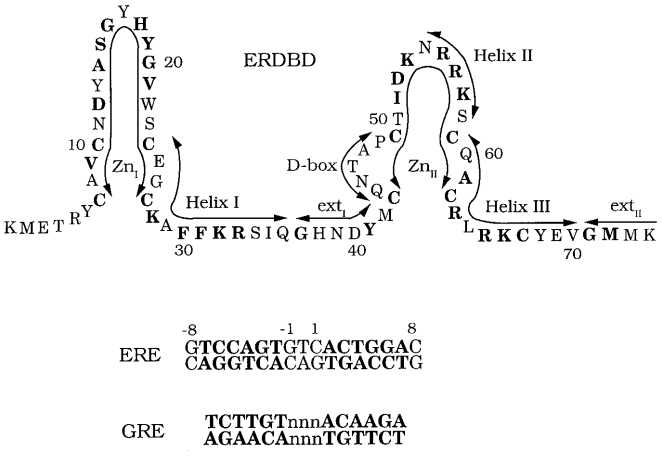


Fig. 1 *Top*, sequence of the estrogen receptor DNA-binding domain (ERDBD). *Bold typed residues* are identical in the glucocorticoid receptor DBD. Abbreviations: *Zn$_I$* and *Zn$_{II}$*, first (N-terminal) and second zinc coordinating subdomains; *helix I* and *helix III*, first and second major α-helical regions; *ext$_I$* and *ext$_{II}$*, two regions with extended conformation; *helix II*, a short piece of distorted helix in the Zn_{II} region [not folded in the solution structure of free ERDBD (Schwabe et al. 1990)]. *Bottom*, DNA-sequence and base pair numbering of the palindromic estrogen receptor response element (ERE) that was present in the simulation. *Below* is the (consensus) glucocorticoid receptor response element (GRE) shown for comparison. The hexameric half-sites are *bold typed*

Gly26 and Ala29 in ERDBD with the corresponding residues in GRDBD, Gly, Ser and Val, respectively. The discriminating roles of residues 25 and 29 are structurally rather obvious, since Glu25 in ERDBD forms a direct hydrogen bond with C5-N_4 as well a number of water-mediated contacts with the bases in the crystal structure (Schwabe et al. 1993a). A replacement of C5 with T5 (in GRE) results in steric conflicts with the methyl group (Zilliacus et al. 1995). Val29 (in GRDBD) is in van der Waals contact with the methyl group of T4 in GRE (Luisi et al. 1991), which is replaced by less favourable, hydrophobic-hydrophilic contacts when T4 is replaced by A4 in ERE. The role of Gly/Ser26 is somewhat different, since it has been shown that the Gly26Ser mutation decreases the affinity to both ERE and GRE, but to a greater extent to ERE (Zilliacus et al. 1992). Since Gly/Ser26 is not involved in any direct interactions with DNA, its discriminating ability is also less obvious than that of the other two P-box residues. In molecular dynamics (MD) simulations (Eriksson and Nilsson 1998) of wild-type GRDBD and of the Ser26Gly mutant ($GRDBD_{ggv}$) in complex with the half-site response elements of GR and ER, GREH and EREH, respectively, it was found that the structural change in $GRDBD_{ggv}$ was larger when bound to EREH than to GREH and that the difference most probably is due to the presence of the T4 methyl group in GREH, which is replaced by A4 in EREH.

The second (C-terminal) subdomain of DBD provides the entire dimerization interface, where one DBD contacts the other in the dimeric binding to the response elements. Most intermolecular contacts are formed in the region between residues Cys43 and Cys49 [the "D-box" (Umesono and Evans 1989)] (Fig. 1), which in GRDBD has a completely different amino acid sequence (Cys-Ala-Gly-Arg-Asn-Asp-Cys). In ERDBD, the dimeric interface consists of two pairs of van der Waals alanines and prolines in van der Waals contact, a pair of N–H···O–C hydrogen bonds, as well as a number of water-mediated contacts (Schwabe et al. 1993a). In the crystal structure of GRE_{S4}-$(GRDBD)_2$ [1] (Luisi et al. 1991) there are more direct protein-protein contacts in the dimeric interface.

The solution structures of free ERDBD (Schwabe et al. 1990) and GRDBD (Härd et al. 1990b, 1990c; Baumann et al. 1993) are overall very similar, with the main differences located in the Zn_{II} region. In GRDBD there is an additional, distorted helix at the tip of this region (helix II), whereas this region is poorly ordered in ERDBD. ^{15}N NMR relaxation measurements (Berglund et al. 1992; Eriksson et al. 1993) as well as an MD simulation (Eriksson et al. 1993) or GRDBD show that the backbone mobility is limited and rather uniform throughout the protein, including the tip of the Zn_{II} region. This contrasts with the fast amide proton exchange rates with the solvent that have been observed for Arg56 and Lys57 in ERDBD (Schwabe et al. 1990), supporting the notion that this region has conformational flexibility. In the crystal structure of GRE_{S4}-$(GRDBD)_2$ (Luisi et al. 1991) the helix II region is well folded and in crystals of ERE-$(ERDBD)_2$ (Schwabe et al. 1993a), in which the asymmetric unit contains two ERE-$(ERDBD)_2$ dimers ("A" and "B"), the helix II region is folded in dimer A, whereas in dimer B it remains extended. Thermodynamic measurements (Lundbäck et al. 1993, 1994) show that the cooperativity of GRDBD when binding to DNA is influenced by a changes response element as well as by P-box mutations in GRDBD. The apparently different structure and dynamics of the Zn_{II} regions of ER- and GRDBD, when free as monomers in solution compared to when DNA-bound as dimers, is very interesting in this context and could be related to the different cooperativities of ERDBD and GRDBD binding to their respective response elements.

In this paper we address the question of why the structure and dynamics in the Zn_{II} region differs in ERDBD and GRDBD. To reveal possible differences in the structure and dynamics of free, as well as DNA-bound, ERDBD we have performed MD simulations of ERE-$(ERDBD)_2$, EREH-ERDBD and free ERDBD. We find a considerably increased flexibility in the Zn_{II} region of EREH-ERDBD, compared to ERE-$(ERDBD)_2$, suggesting that ordering of the Zn_{II} region requires dimerization of ERDBD. This contrasts to a previous MD simulation of GREH-GRDBD (Eriksson and Nilsson 1998), where the structure and dynamics of monomeric GRDBD were similar to the dimeric

[1] In the 2.9 Å crystal structure (Luisi et al. 1991) of GRE-$(GRDBD)_2$ the spacing between the half-sites is four, i.e. one more than in the natural GRE. Because of this extra base-pair spacing (GRE_{S4}), one GRDBD ($GRDBD_{unspec}$) is out of register by one base pair, resulting in a non-specific binding to the DNA, whereas the other GRDBD ($GRDBD_{spec}$) is believed to bind as in the native GRE_{S3}-$(GRDBD)_2$ complex

GRDBD in complex with GRE. We suggest that the different structure and dynamics of ER- an GRDBD are due to a difference in the hydrogen bonding network that connects the DNA with the Zn_I and Zn_{II} regions in GREH-GRDBD and EREH-ERDBD.

Methods

Parameters and protocol

All simulations were performed with the program CHARMM (Brooks et al. 1983) version 25, using the all-atom force field (Chemistry Department, Harvard University, Cambridge, Mass.; MacKerell et al. 1995, 1998), modified for the zinc ions and the side chains of the coordinating cysteins. The zinc ions were treated as formal atoms, covalently attached to the surrounding four cysteines, and force field parameters for the zinc, the coordinating cysteine sulfurs, as well as their C_β atoms, were obtained from MNDO (Dewar and Thiel 1977a, 1977b; Dewar and Merz 1986)/ESP (Besler et al. 1990) calculations with the program MOPAC (Steward 1990). For further details of this parameterization, see Eriksson et al. (1995). Non-bonded interactions were cut off at 10 Å by shifting (Brooks et al. 1983) the potential energy smoothly to zero at 10 Å. A time step of 2 fs was used and the non-bonded list including neighbouring atoms within a 11 Å distance was updated every 40 fs. All bonds were constrained with the SHAKE algorithm (Ryckaert et al. 1977).

Setup and equilibration

All three simulations started with dimer "A" of the 2.4 Å resolution crystal structure of ERDBD as a dimer in complex with the 17 base-pair ERE [(Schwabe et al. 1993a), PDB entry: 1 HCQ]. The side-chains of Gln36, His38, Asp40, Met73 and Lys74 were not present in $ERDBD_{A1}$ of dimer "A" and these were built in extended conformations. Hydrogens were added as previously described (Brünger and Karplus 1988). We removed seven base pairs and one of the monomers ("A"), when simulating one ERDBD monomer bound to its half site (EREH). The starting structure of free ERDBD was obtained by removing ERDBD "A" and the DNA. A subsequent 100 steps of adopted basis Newton-Raphson minimization of the three simulated systems in vacuo removed bad contacts that might have formed in the model building. The protein or the protein-DNA complexes were then surrounded with TIP3P water molecules (Jorgensen et al. 1983) from a sphere of equilibrated water. A sphere of 30 Å radius was used for free ERDBD and EREH-ERDBD, and for ERE-$(ERDBD)_2$ the radius of the sphere was 42 Å. To ensure that cavities were filled with water, the water sphere was rotated several times and new water molecules were added to the protein- or protein-DNA-water sphere if they could be positioned without overlapping existing atoms. The criterion for overlap was a distance closer than 2.8 Å from any atom in the sphere. The water molecules interacted with a stochastic heat bath of 300 K via randomly fluctuating forces and dissipative forces (Brooks and Karplus 1983) in the "buffer region", which constitutes the outer 3 Å shell of water molecules in the respective spheres. The water also interacted with a "deformable boundary force", arising from mean field interactions of fictitious water molecules outside the sphere, to prevent the water from escaping during the simulations. Keeping non-water atoms fixed, the surrounding water molecules were further equilibrated for 40 ps. Counterions [22 Na^+ in ERE-$(ERDBD)_2$, 13 Na^+ in EREH-ERDBD and 5 Cl^- in free ERDBD] were then added to achieve electroneutrality, by replacing the water molecules with the highest (for Na^+) or lowest (for Cl^-) electrostatic energies, and more that 5 Å apart from each other, with ions. A further 40 ps equilibration of the electroneutral systems was performed, still with fixed protein (protein-DNA) atoms. Thereafter, the protein (protein-DNA) was released, except for the first and last base pairs of the DNA in the EREH-ERDBD and ERE-$(ERDBD)_2$ simulations. These base pairs were kept in position with harmonic constraints [force constant: 10 kcal/(mol $Å^2$)], to prevent fraying ends of the DNA that may occur for short DNA fragments in solution. We did not observe fraying ends when removing the corresponding constraints at the end of an MD simulation of GRE-$(GRDBD)_2$ (Eriksson et al. 1995). Nevertheless, we wanted to ensure that fraying ends of DNA do not take place in these simulations, that are run for a longer time, since it does not occur naturally in a longer stretch of DNA. The free ERDBD was then simulated for 1.0 ns, EREH-ERDBD for 1.3 ns and ERE-$(ERDBD)_2$ for 0.5 ns.

Results

ERE-$(ERDBD)_2$

The simulation of the ERE-$(ERDBD)_2$ complex is overall stable, with an initial increase of the root mean square deviation (rmsd) from the crystal structure to around 1.2 Å, where it is stable throughout the simulation (Fig. 2). The averaged structure from the last 50 ps simulation remains very close to the crystal structure and the total interaction energy between ERDBD and DNA (Fig. 2) also remains relatively constant for both monomers throughout the simulations. In the spacer region of DNA between the two hexameric half-sites, the DNA has approached the proteins relative to the crystal structure, resulting in slightly bent DNA, which also has been observed in simulations of GRE-$(GRDBD)_2$ (Bishop and Schulten 1994; Eriksson et al. 1994, 1995) and in recent 100 ps simulations of ERE-$(ERDBD)_2$ (Bishop et al. 1997). The direct and water-mediated protein-DNA contacts (Table 1) are very similar to those found in the crystal structure and to the contacts described by Kosztin et al. (1997), with the majority of the direct hydrogen bonds present during most of the simula-

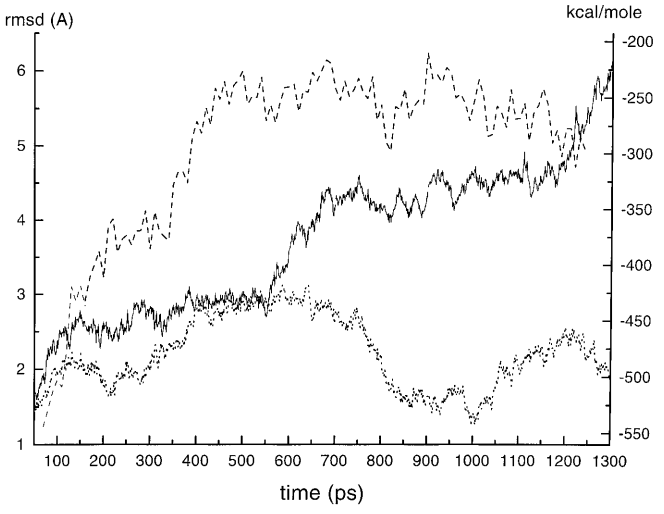


Fig. 2 *Solid line*, time evolution of the root mean square deviation (rmsd, Å) of the non-hydrogen atoms in ERE-$(ERDBD)_2$ from the energy minimized crystal structure; *dashed line*, total interaction energy (kcal/mol) between $ERDBD_{A1}$ and ERE; *dotted line*, total interaction energy (kcal/mol) between $ERDBD_{A2}$ and ERE. There is gap in the curves at around 175 ps, owing to an accidental loss of a piece of trajectory during the equilibration

tion. A few new water bridges [i.e. not found in the crystal structure (Schwabe et al. 1993a)] between the DNA backbone and the proteins have also been established during the simulation (Table 1). In the dimeric interface (i.e. between $ERDBD_{A1}$-$ERDBD_{A2}$), one new symmetric contact between Cys53-O and Arg55-N_ε is established as well as a number of water bridges (Table 2). Only one of the symmetric water bridges between Met42-O and Ser58-O_γ, that is found in the crystal structure, is present also in the simulation (Table 2). The average occupancy time of a water molecule in a bridging position is somewhat lower in the dimeric interface than in the interface between the ERDBDs and DNA, suggesting that the water molecules are more mobile in the dimeric interface. The hydrogen bonding network, formed by residues at the protein-DNA interface, is very similar in the two ERDBDs (Fig. 3, top and middle), with some minor differences for bridging water molecules. We find, not surprisingly, many similarities with the hydrogen bonding network in GRE-$(GRDBD)_2$ (Luisi et al. 1991), where the conserved residues Asp12, Arg56 and Arg63 have nearly identical orientations and connections as in ERE-$(ERDBD)_2$. The flexibility (in terms of by-residue fluctuations around the average structure) is low and relatively uniform, in reasonably good agreement with B-factors from the crystal structure. The ext_I region (Fig. 1) is somewhat more flexible in the MD simulation, and also has several poorly defined residues in the crystal structure (see Methods). Thus, to summarize from the ERE-$(ERDBD)_2$ simulation, we find a stable structure that is very close to the crystal structure and a dynamics in accordance with crystallographic B-factors.

Table 1 Hydrogen bonds[a] and water bridges at the ERDBD-DNA interface (presence time in ps)

Analysed time	220–500 ps		900–1300 ps
	$ERDBD_{A1}$-ERE	$ERDBD_{A2}$-ERE	$ERDBD_{A1}$-EREH
Hydrogen bond			
His18-$N_{\delta1}\cdots$A±7-O_{2p}[b]	280	280	–
Tyr19-N$\cdots$A-7-O_{2p}[b]	–	258	–
Tyr19-$O_\eta\cdots$G-6-O_{2p}[b]	–	241	–
Glu25-$O_{\varepsilon1}\cdots$C±5-N_4[b]	263	277	380
Lys28-$N_\zeta\cdots$G±6-O_6[b]	129	204	–
Lys28-$N_\zeta\cdots$G±6-N_7	157	70	–
Lys32-$N_\zeta\cdots$G±5-N_7[b]	263	165	316
Lys32-$N_\zeta\cdots$T-4-O_4	–	34	–
Arg33-$N_{\eta1}/N_{\eta2}\cdots$G±3-$N_7$[b]	274/0	39/0	288/48
Arg33-$N_{\eta2}\cdots$G 3-O_6	–	–	156
Arg33-$N_{\eta2}\cdots$T±2-O_{2p}[b]	260	277	–
Arg33-$N_\varepsilon\cdots$T2-O_{2p}[b]	112	–	–
Gln36-$N_{\delta2}\cdots$G-5-O_{2p}	–	109	–
Arg56-$N_\varepsilon\cdots$G±3-O_{2p}[b]	274	277	–
Arg56-$N_{\eta2}\cdots$G±3-O_{1p}[b]	87	221	–
Lys57-$N_\zeta\cdots$G3-O_{1p}[b]	277	–	–
Gln60-$N_\varepsilon\cdots$T±2-O_{2p}[b]	193	104	160
Arg63-$N_{\eta1}/N_{\eta2}\cdots$G±3-$O_{2p}$[b]	277/255	274/263	184/240
Lys57-$N_\zeta\cdots$G-6-O_{1p}	–	115	–
Water bridge			
Asp12-$O_{\delta2}\cdots$w$\cdots$A4-O_{2p}	51 (3)[c]	–	–
Tyr17-O$\cdots$w$\cdots$A±7-O_{2p}	161 (7)	174 (1)	35 (5)
Tyr17-N$\cdots$w$\cdots$C-8-O_3'	–	119 (4)	–
Cys24-$S_\gamma\cdots$w$\cdots$A4-O_{2p}	–	–	174 (3)
Glu25-O$\cdots$w$\cdots$A±4-N_6[b]	80 (1)	120 (1)	329 (2)
Glu25-O$\cdots$w$\cdots$A4-N_7	–	–	321 (2)
Glu25-$O_{\varepsilon1}\cdots$w$\cdots$A±4-N_6[b]	55 (1)	129 (1)	–
Gly26-O$\cdots$w$\cdots$G±3-O_{2p}[b]	228 (1)	252 (1)	254 (1)
Gln36-$O_{\varepsilon1}\cdots$w$\cdots$G±5-O_{2p}	43 (4)	87 (7)	–
Arg56-O$\cdots$w$\cdots$T±2-O_{1p}[b]	39 (3)	216 (1)	–
Lys57-O$\cdots$w$\cdots$T-2-O_{1p}	–	158 (1)	–
Gln60-O$\cdots$w$\cdots$T2-O_{1p}	–	–	177 (2)
Arg63-$N_{\eta1}\cdots$w$\cdots$T±2-O_{1p}[b]	35 (1)	215 (1)	–

[a] Criteria for a hydrogen bond between A and H-D: distance A-D < 3,5 Å, angle A-H-D > 135°
[b] Present in the crystal structure (Schwabe et al. 1993a)
[c] Number of different water molecules that occupy the bridging position during the analysed time

EREH-ERDBD

In the simulation of one half-site DNA (EREH) in complex with $ERDBD_{A1}$, we find a considerably less stable structure than for ERE-$(ERDBD)_2$ above. The rmsd from the initial structure for the protein approaches 5 Å after 700 ps (Fig. 4) and increases further after 1.2 ns, and for the DNA rmsd(t) instead reaches a maximum at around 600 ps, after which it returns to 1.2 Å. Apparently the EREH-ERDBD complex has not reached equilibrium in the 1.3 ns

Table 2 Hydrogen bonds and water bridges at the dimeric interface in ERE-$(ERDBD)_2$ during the period 220–500 ps of the simulation

Hydrogen bond ($ERDBD_{A1}$-$ERDBD_{A2}$)	Presence time (ps)
Cys43-O···Arg55-N_ε	182
Met42-O···Ser58-O_γ	123
Pro44-O···Thr50-N[a]	274
Arg55-N_ε···Cys43-O	246
Thr50-N···Pro44-O[a]	266
Water bridge	
Met42-O···w···Ser58-O_γ[a]	35 (3)[b]
Cys43-N···w···Arg55-$N_{\eta 2}$	142 (4)
Thr46-$O_{\gamma 1}$···w···Thr46-$O_{\gamma 1}$[a]	81 (1)
Gln48-O···w···Ala45-O	43 (3)
Gln48-$O_{\varepsilon 1}$···w···Thr46-$O_{\gamma 1}$	33 (3)
Gln60-$O_{\varepsilon 1}$···w···Ser58-O_γ	41 (2)

[a] Present in the crystal structure (Schwabe et al. 1993a)
[b] Number of different water molecules that occupy the bridging position for the analysed time

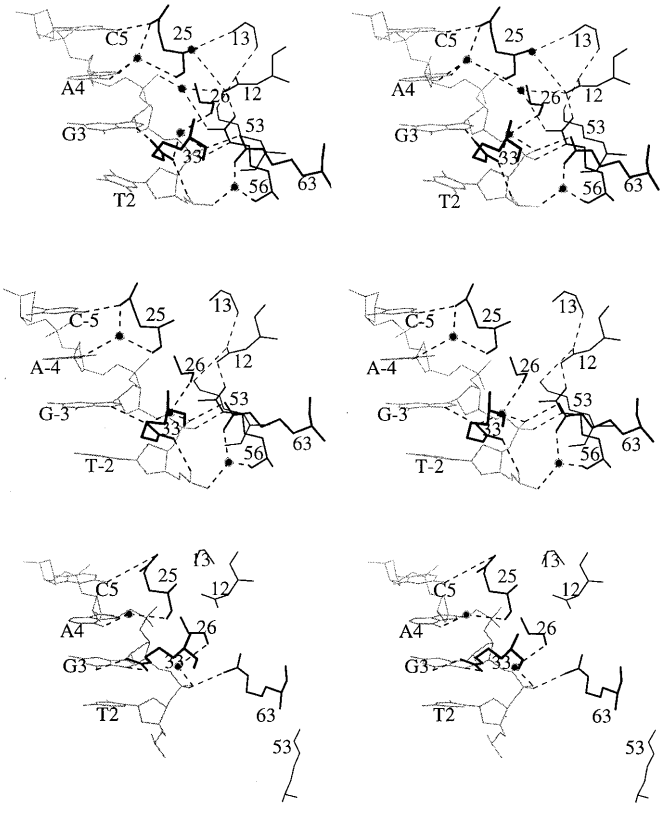


Fig. 3 Some important residues participating in the hydrogen bonding network that connects Zn_I, Zn_{II} and DNA (stereo views): *top*, ERE-$ERDBD_{A1}$; *middle*, ERE-$ERDBD_{A2}$ from averaged structure of the last 50 ps simulation of ERE-$(ERDBD)_2$; *bottom*, EREH-$ERDBD_{A1}$ from averaged structure of the last 50 ps simulation

of simulated time. The total strength of the protein-DNA association as judged from the protein-DNA interaction energy (Fig. 4) has decreased to less than half of that of one protein with the DNA in the dimeric complex (Fig. 2).

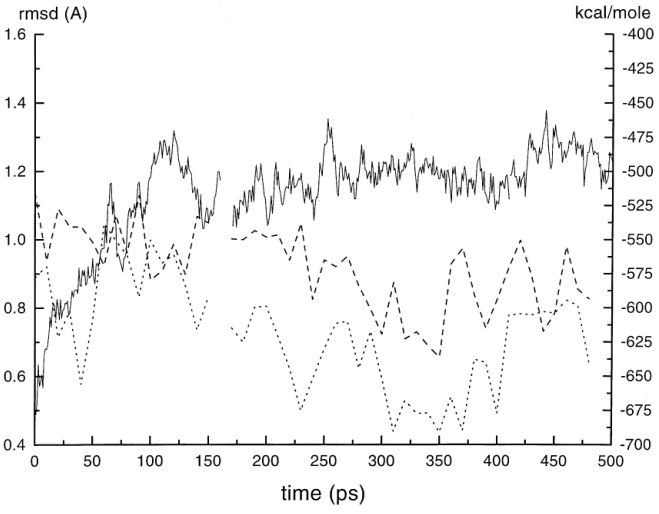


Fig. 4 Time-evolution of the EREH-$ERDBD_{A1}$ simulation. *Solid line*, rmsd (Å) of the non-hydrogen atoms in $ERDBD_{A1}$ from the energy minimized crystal structure; *dotted line*, rmsd (Å) of non-hydrogen atoms in EREH; *dashed line*, total interaction energy (kcal/mol) between $ERDBD_{A1}$ and EREH

During the simulation, the C-terminal of helix I leaves the DNA, unfolds and adopts a conformation which is very similar to that of the NMR solution structure (Schwabe et al. 1990) (Fig. 5). The most striking difference from the crystal structure of ERE-$(ERDBD)_2$ is the orientation of the tip of the Zn_{II} region (Ile51-Ser57), where helix II is located (see Fig. 1). This region has rotated away from the Zn_I region during the simulation and, compared to ERE-$(ERDBD)_2$, helix II has been considerably destabilized (Fig. 6). A comparison of the mean square fluctuations per residue over three different time periods (Fig. 7) reveals a rather constant, high flexibility in the ext_I, D-box and helix II regions. Towards the end of the simulation the residues prior to the Zn_I region also become very flexible. This end of the protein has moved considerably towards the bulk solvent, which should render it more flexible than in its initial position. In more detail (Fig. 3, bottom), residues in the Zn_I and helix I regions have relatively similar positions and contacts with DNA as in the dimeric complex (Table 1). Glu25 forms a very stable hydrogen bond to C5-N_4 and the water-mediated contacts with Glu25-O and the base A4, also found in the crystal structure, are also preserved during almost the entire simulation (Table 1). Further, Lys32 and Arg63 form quite persistent hydrogen bonds to the DNA backbone (Table 1). It is interesting that Arg33 has adopted a similar orientation as in GRE-$(GRDBD)_2$ (Luisi et al. 1991), with the forks of the side chain contacting G3-N_7 and -O_6, respectively. In ERE-$(ERDBD)_2$ there is a stable hydrogen bond between the side chains of Cys10 in the Zn_I region and Arg56 in the Zn_{II} region, which is lost in EREH-ERDBD (Table 3). Furthermore, Asp12, which is hydrogen bonded to the side chain of Lys53, retains this hydrogen bond, and is thus displaced away from the hydrogen bonding network when the

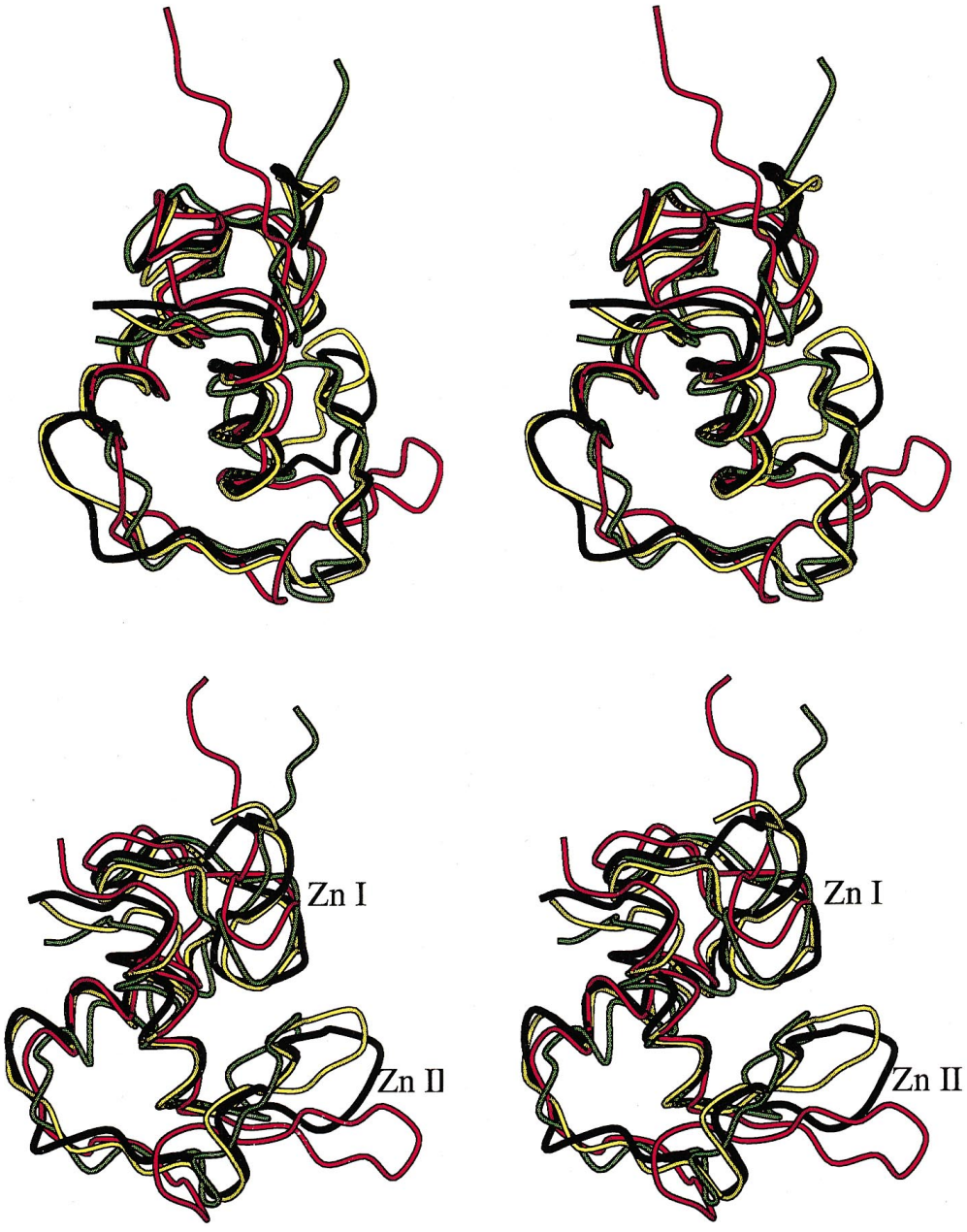


Fig. 5 Two stereo views of an ERDBD monomer; *black*, averaged structure of the last 50 ps simulation of free ERDBD; *green*, averaged NMR solution structure (Schwabe et al. 1990); *red*, averaged structure of the last 50 ps simulation of EREH-ERDBD$_{A1}$; *yellow*, from the crystal structure of ERE-(ERDBD)$_2$ (Schwabe et al. 1993a)

Table 3 Hydrogen bonds that connect the Zn_I and Zn_{II} regions (presence time in ps)

Analysed time	220–500 ps		900–1300 ps	700–1000 ps
Hydrogen bond	ERDBD$_{A1}$-ERE	ERDBD$_{A2}$-ERE	ERDBD$_{A1}$-EREH	Free ERDBD$_{A1}$
Cs10-S$_\gamma$···Arg56-N$_{\eta1}$	258	266	–	–
Asp12-O$_{\delta1/\delta2}$···Lys53-N$_\zeta$	115/154	115/221	148/180	–
Asp12-O$_{\delta2}$···Arg56-N$_{\eta1}$	218	269	–	–

Zn_{II} region is displaced away from the Zn_I region. The final increase in the rmsd(t) (Fig. 4) is due to a further displacement of the Zn_{II} region away from the Zn_I region and this hydrogen bond is completely broken in the averaged structure from the last 50 ps simulation (Fig. 3, bottom).

Free ERDBD

The rmsd from the crystal structure of ERDBD$_{A1}$ (the starting structure) is relatively stable at around 1.6 Å and the protein also remains closer to the crystal structure than to

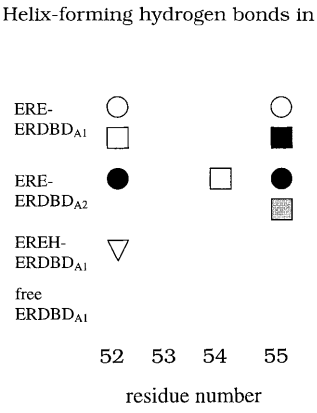


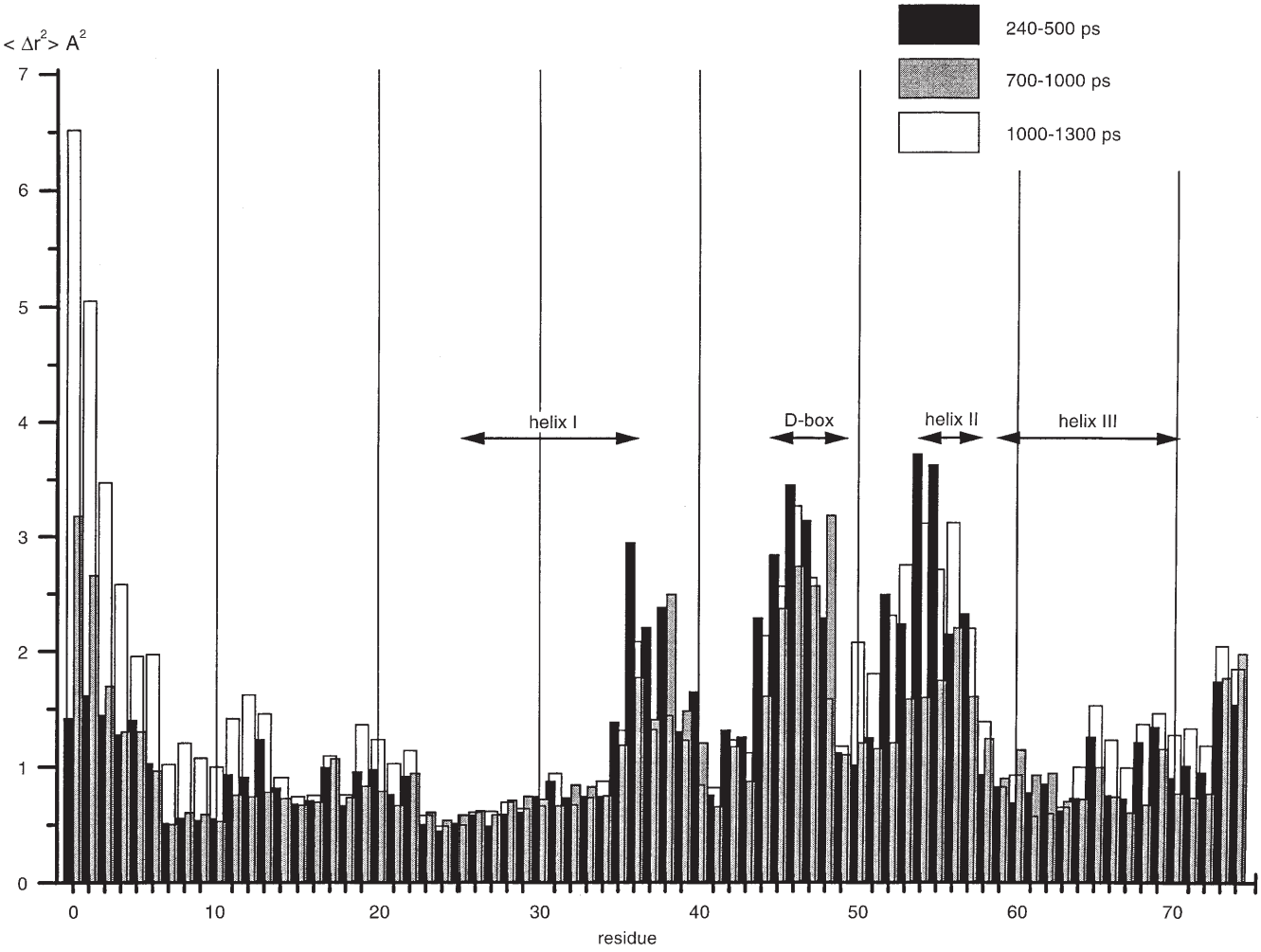
Fig. 6 Occurrence of carbonyl to amide proton hydrogen bonds in the helix II region from the different simulated systems. The symbols indicate (i, i + 4) hydrogen bonds (○), (i, i + 3) hydrogen bonds (□) and (i, i + 2) hydrogen bonds (▽). The symbol is *open* if the hydrogen bond is present in 20–50% of the structures, *shaded* if 50–75% contain the hydrogen bond and *black* if the hydrogen bond is present in more than 75% of the structures

the average NMR solution structure (Figs. 5 and 8). We observe no tendency of unfolding in the C-terminal helix I, as in the solution structure and in the simulation of EREH-ERDBD. However, comparing the by-residue fluctuations between 250–500 ps and 500–1000 ps, respectively, a slightly increased flexibility in the D-box and helix II regions can be seen. These regions have the largest variations within the 30 NMR solution structures (Schwabe et al. 1990) and are also two regions that are very flexible in the EREH-ERDBD$_{A1}$ simulation (Fig. 7). Further, all contacts between the Zn_I and Zn_{II} regions, present in ERE-(ERDBD)$_2$ (Table 3), have broken during the simulation, as have all helical hydrogen bonds in the helix II region (Fig. 6), showing that this region is adopting a more extended conformation, as in the NMR solution structure (Schwabe et al. 1990). These results together might indicate that the ERDBD eventually will rearrange towards the structure in solution in a considerably prolonged simulation.

Discussion

According to our simulations there is a considerable difference in the structure and dynamics of ERDBD binding

Fig. 7 By-residue fluctuations of ERDBD$_{A1}$ in simulated EREH-ERDBD$_{A1}$ around averaged structures over different time periods



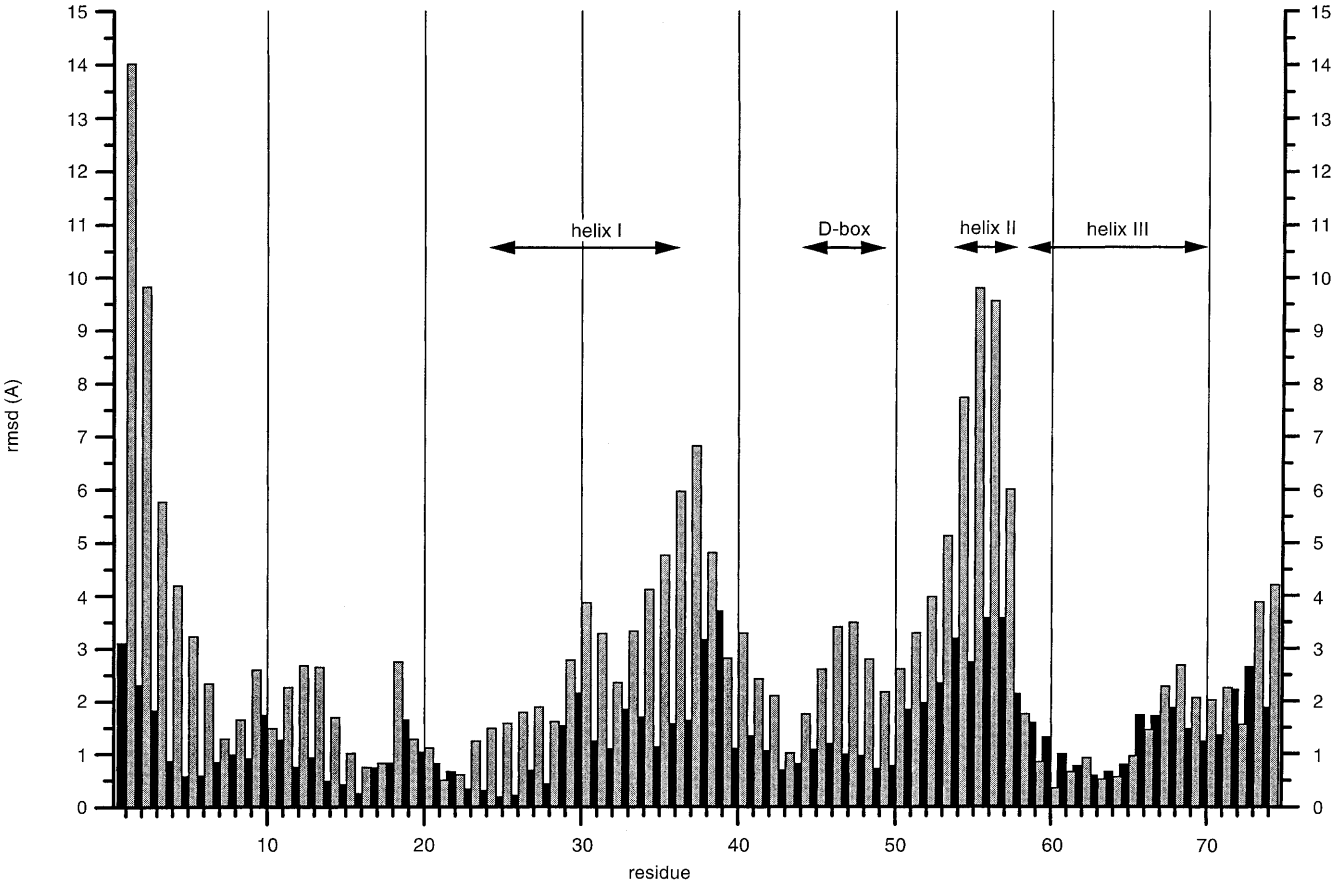


Fig. 8 By-residue rmsd (Å) for the last (averaged) 50 ps simulation of free $ERDBD_{A1}$ from the crystal structure $ERDBD_{A1}$ (*black bars*) and from the NMR solution structure (*grey bars*)

as a monomer to its half-site response element compared to when binding as a dimer to the palindromic response element. These differences evolve during the course of the simulations in such a way that in EREH-ERDBD the rmsd from the initial structure increases to levels which are unusual in simulations of proteins or protein-DNA complexes. In contrast, the simulations of free ERDBD and of the ERE-$(ERDBD)_2$ complex are very well behaved. Even though it is of course always a possibility that at some later time there would be structural changes in these systems too, we interpret the early onset, and large magnitude, of the change in EREH-ERDBD to indicate a significant instability in this structure.

In the dimeric complex, residues from the tip of the Zn_{II} region (including helix II) are involved in an intricate network of direct and water-mediated hydrogen bonds to DNA and to the Zn_I region. In EREH-ERDBD these residues become very flexible and are oriented more towards the solvent as they are detached from the network. It thus seems likely that contact with another ERDBD is a prerequisite for positioning of the Zn_{II} region closer to Zn_I and for contact with DNA. This is consistent with the region between Cys43 and Cys59 being poorly ordered in the NMR solution structure (Schwabe et al. 1990), whereas this region is well defined in ERE-$(ERDBD)_2$ (Schwabe et al. 1993a). Even though the simulated free $ERDBD_{A1}$ remains closer to the crystal structure throughout the simulation, we observe tendencies for the protein to adopt a similar structure and dynamics as in the solution structure. Such tendencies include increased flexibilities in the D-box region, and in particular in helix II during the later part of the simulation, as well as breaking of all hydrogen bonds in this helix. Moreover, the hydrogen bonds that connect the Zn_I ad Zn_{II} regions in ERE-$(ERDBD)_2$ have broken in our simulation of the free $ERDBD_{A1}$.

The residues Arg56 and Lys57 contact the DNA backbone in dimer "A" (see Introduction) of the two ERE-$(ERDBD)_2$ dimers in the asymmetric crystal unit (Schwabe et al. 1993a) and also in recent MD simulations of dimeric ERDBD bound to ERE and to an ERE/GRE chimeric response element (Kosztin et al. 1997). In the present simulation of ERE-$(ERDBD)_2$, Lys57 contacts DNA directly only in $ERDBD_{A1}$, and through a water molecule in $ERDBD_{A2}$ (see Table 1). In dimer "B", these two residues are not in contact with the DNA but are oriented towards the solvent and helix II, which is folded in dimer A and is forming an extended loop (Schwabe et al. 1993a). This structural difference between the two dimers in the asymmetric crystal unit has also been found for dimeric ERDBD in complex with a non-consensus DNA (Schwabe et al. 1995), in spite of the different crystal packing environment compared to the ERE-$(ERDBD)_2$ structure. It was therefore suggested that the structure of ERDBD in dimer B is

a trapped intermediate in the folding process (Schwabe et al. 1995) and this suggestion is further supported by our comparison of ERE-$(ERDBD)_2$ and EREH-ERDBD.

In a recent 500 ps MD simulation of GREH-GRDBD (Eriksson and Nilsson 1998), using the same parameters, protocol and size of the water sphere (30 Å radius) as herein, we found no drifting structure or regions with high flexibility. All protein-DNA contacts that are present in the dimeric complex with GRE (Luisi et al. 1991; Eriksson et al. 1995) were also present in the monomeric, half-site complex (Eriksson and Nilsson 1998). This raises the question why EREH-ERDBD and GREH-GRDBD apparently behave differently in terms of structure and dynamics, compared to their respective dimeric complexes.

A major part of the hydrogen bonding network that connects Zn_I, Zn_{II} and DNA consists of the conserved residues Cys10, Asp12, Arg56, Lys57, the backbone of Glu13 (GRDBD)/Tyr13 (ERDBD), Ser26 (GRDBD)/Gly26 (ERDBD) and of DNA backbone atoms, which all have very similar orientations and connections in the two dimeric complexes. One difference, however, is the presence of the Ser26 hydroxyl group in GRDBD, which also has a central role in this network (Eriksson and Nilsson 1998). The Ser26 side chain connects to the side chains of Arg56 and Arg63 and, through a water bridge, to Asp12. It is possible that the stabilizing effect of Ser26 in GRDBD contributes to preserve the network in GREH-GRDBD as in the dimeric form, and does not require an additional GRDBD to form a stable Zn_{II} region. The absence of a hydroxyl group of residue 26 in ERDBD would then contribute to the lower stability of this network and thus the Zn_{II} region (including helix II) that is found in EREH-ERDBD and in the NMR solution structure of ERDBD, compared to ERE-$(ERDBD)_2$. This suggestion is further supported by the folded helix II in the NMR solution structure of GRDBD (Härd et al. 1990b, 1990c; Baumann et al. 1993). Accordingly, the above-described network would be sufficiently strong to order the residues in the Zn_{II} region and keep the helical conformation in the tip of this region, even in the absence of another GRDBD. We have also analysed this network in EREH-$GRDBD_{ega}$ where the P-box residues Gly25, Ser26 and Val29 in GRDBD where mutated to those in ERDBD, i.e. Glu, Gly and Ala, respectively (Eriksson and Nilsson 1998), and found that Lys57 had lost its strong contact with the DNA, present in GRDBD-GREH. The backbone of Arg56 had also translated away from the DNA, releasing the stable water molecule that connects Arg56-O, Arg63-$N_{\eta 1}$ and the phosphate group of T±2, which is present in GREH-GRDBD, GRE-$(GRDBD)_2$ and ERE-$(ERDBD)_2$. This is also consistent with a reduced stability of the hydrogen bond network connecting to the Zn_{II} region in EREH-$GRDBD_{ega}$. Recent NMR measurements of the $GRDBD_{ega}$ mutant (Berglund et al. 1997) also show an increased flexibility in the Zn_{II} region compared to wild-type GRDBD, which thus further supports the importance of Ser26 for stabilizing the Zn_{II} region. Thermodynamic measurements (Lundbäck et al. 1994) have shown that binding of GRDBD to GREH is stronger than binding of $GRDBD_{ega}$ to EREH, but that the cooperativity in binding dimeric $GRDBD_{ega}$ to ERE is about tenfold larger than when dimeric GRDBD binds to GRE. This, too, agrees with a picture in which binding of the Zn_{II} region to DNA and to the Zn_I region is independent of the presence of another protein in GRDBD, while a binding of this region in $GRDBD_{ega}$ is facilitated by dimerisation.

Acknowledgements We are grateful to Dr. John W. R. Schwabe for providing us with the coordinates of the crystal structure of ERE-$(ERDBD)_2$, and to Dr. K. Schulten for giving us access to his data prior to publication. This work has been supported the Swedish Natural Research Council (NFR) and with computing resources on an IBM SP2 system by the Swedish Council for High Performance Computing (HPDR) and Parallelldatorcentrum (PDC), Royal Institute of Technology.

References

Anolik JH, Klinge CM, Bambara RA, Hilf R (1993) Differential impact of flanking sequences on estradiol- vs 4-hydroxytamoxifen-liganded estrogen receptor binding to estrogen responsive elements. J Steroid Biochem Mol Biol 46:713–730

Baumann H, Paulsen K, Kovács H, Berglund H, Wright APH, Gustafsson J-Å, Härd T (1993) Refined solution structure of the glucocorticoid receptor DNA-binding domain. Biochemistry 32:13463–13471

Beato M (1989) Gene regulation by steroid hormones. Cell 56:335–344

Beato M, Chalepakis G, Schauer M, Slater EP (1989) DNA regulatory elements for steroid hormones. J Steroid Biochem 32:737–747

Berglund H, Kovács H, Dahlman-Wright K, Gustafsson J-Å, Härd T (1992) Backbone dynamics of the glucocorticoid receptor DNA-binding domain. Biochemistry 31:12001–12011

Berglund H, Wolf-Watz M, Lundbäck T, van der Berg S, Härd T (1997) Structure and dynamics of the glucocorticoid receptor DNA-binding domain: comparison of wild type and mutant altered specificity. Biochemistry 36:11188–11197

Besler BH, Merz JKM, Kollman PA (1990) Atomic charges derived from semiempirical methods. J Comput Chem 11:431–439

Bishop T, Schulten K (1994) Molecular dynamics study of a sequence specific protein-DNA interaction. (NATO ASI, vol 426) Kluwer, Dordrecht, pp 419–439

Bishop TC, Kosztin D, Schulten K (1997) How hormone receptor DNA binding affects nucleosomal DNA: the role of symmetry. Biophys J 72:2056–2067

Brooks BR, Bruccoleri RE, Olafsson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy, minimization and dynamics calculations. J Comput Chem 4:187–217

Brooks ICL, Karplus M (1983) Deformable stochastic boundaries in molecular dynamics. J Chem Phys 79:6312–6325

Brünger A, Karplus M (1988) Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. Proteins Struct Funct Genet 4:148–156

Dahlman-Wright K, Siltala-Roos H, Carlstedt-Duke J, Gustafsson J-Å (1990) Protein-protein interactions facilitate DNA binding by the glucocorticoid receptor DNA-binding domain. J Biol Chem 265:14030–14035

Danielsen M, Hinck ML, Ringold GM (1989) Two amino acids within the knuckle of the first zinc finger specify DNA response element activation by the glucocorticoid receptor. Cell 57:1131–1138

Dewar MJS, Merz JKM (1986) MNDO calculations for compounds containing zinc. Organometallics 5:1494–1496

Dewar MJS, Thiel WJ (1977a) Ground states of molecules. 38. The MNDO method. Approximations and parameters. J Am Chem Soc 99:4899–4906

Dewar MJS, Thiel WJ (1977b) Ground states of molecules. 39. MNDO results for molecules containing hydrogen, carbon, nitrogen and oxygen. J Am Chem Soc 99:4907–4917

Eriksson MAL, Nilsson L (1998) Structural and dynamic effects of point mutations in the recognition helix of the glucocorticoid receptor DNA-binding domain. Molecular dynamics simulation studies. Protein Eng 11:589–600

Eriksson MAL, Härd T, Nilsson L (1994) Molecular dynamics simulations of a DNA-binding protein free and in complex with DNA. (NATO ASI, vol 426) Kluwer, Dordrecht, pp 441–456

Eriksson MAL, Berglund H, Härd T, Nilsson L (1993) A comparison of ^{15}N NMR relaxation measurements with a molecular dynamics simulation: backbone dynamics of the glucocorticoid receptor DNA-binding domain. Proteins Struct Funct Genet 17:375–390

Eriksson MAL, Härd T, Nilsson L (1995) Molecular dynamics simulations of the glucocorticoid receptor DNA-binding domain in complex with DNA and free in solution. Biophys J 68:1–25

Härd T, Gustafsson J-Å (1993) Structure and function of the DNA-binding domain of the glucocorticoid receptor and other members of the nuclear receptor supergene family. Acc Chem Res 26:644–650

Härd T, Dahlman K, Carlstedt-Duke J, Gustafsson J-Å, Rigler R (1990a) Cooperativity and specificity in the interactions between DNA and the glucocorticoid receptor DNA-binding domain. Biochemistry 29:5358–5364

Härt T, Kellenbach E, Boelens R, Kaptein R, Dahlman K, Carlstedt-Duke J, Freedman LP, Maler BA, Hyde EI, Gustafsson J-Å, Yamamoto KR (1990b) ^{1}H NMR studies of the glucocorticoid receptor DNA-binding domain: sequential assignments and identification of secondary structure elements. Biochemistry 29:9015–9023

Härd T, Kellenbach E, Boelens R, Maler BA, Dahlman K, Freedman LP, Carlstedt-Duke J, Yamamoto KR, Gustafsson J-Å, Kaptein R (1990c) Solution structure of the glucocorticoid receptor DNA-binding domain. Science 249:157–160

Jorgensen WL, Chandrasekar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935

Klock G, Strähle U, Schütz G (1987) Oestrogen and glucocorticoid response elements are closely related but distinct. Nature 329:734–736

Kosztin D, Bishop TC, Schulten K (1997) Binding of the estrogen receptor to DNA – the role of waters. Biophys J 73:557–570

Luisi BF, Xu WX, Otwinowski Z, Freedman LP, Yamamoto KR, Sigler PB (1991) Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. Nature 352:497–505

Lundbäck T, Cairns C, Gustafsson J-Å, Carlstedt-Duke J, Härd T (1993) Thermodynamics of the glucocorticoid receptor-DNA interaction: binding of wild type GR DBD to different response elements. Biochemistry 32:5074–5082

Lundbäck T, Zilliacus J, Gustafsson J-Å, Carlstedt-Duke J, Härd T (1994) Thermodynamics of sequence-specific glucocorticoid receptor-DNA interactions. Biochemistry 33:5955–5965

MacKerell Jr AD, Wiorkiewicz-Kuczera J, Karplus M (1995) An all-atom empirical energy function for the simulation of nucleic acids. J Am Chem Soc 117:11946–11975

MacKerell Jr AD, Bashford D, Bellott M, Dunbrack Jr RL, Evanseck J, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher IWE, Roux B, Schlenkrich M, Smith J, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M (1998) All-hydrogen empirical potential for molecular modeling and dynamics studies of proteins using the CHARMM22 force field. J Phys Chem B102:3586–3616

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

Martinez E, Givel F, Wahli W (1987) The estrogen-responsive element as an inducible enhancer: DNA sequence requirements and conversion to a glucocorticoid-responsive element. EMBO J 6:3719–3727

Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. J Comput Phys 23:327–341

Schwabe JWR, Neuhaus D, Rhodes D (1990) Solution structure of the DNA-binding domain of the oestrogen receptor. Nature 348:458–461

Schwabe JWR, Chapman L, Finch JT, Rhodes D (1993a) The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. Cell 75:567–578

Schwabe JWR, Chapman L, Finch JT, Rhodes D, Neuhaus D (1993b) DNA recognition by the oestrogen receptor: from solution to the crystal. Structure 1:187–204

Schwabe JWR, Chapman L, Rhodes D (1995) The oestrogen receptor recognizes an imperfectly palindromic response element through an alternative side-chain conformation. Structure 3:201–213

Steward JJP (1990) MOPAC: a semiempirical molecular orbital program. J Comput Aided Mol Des 4:1–45

Strähle U, Klock G, Schutz GA (1987) A DNA sequence of 15 base pairs is sufficient to mediate both glucocorticoid and progesterone induction of gene expression. Proc Natl Acad Sci USA 84:7871–7875

Tsai M, O'Malley BW (1994) Molecular mechanisms of action of steroid/thyroid receptor superfamily members. Annu Rev Biochem 63:451–486

Umesono K, Evans RM (1989) Determinants of target gene specificity for steroid/thyroid hormone receptors. Cell 57:1139–1146

Zilliacus J, Dahlman-Wright K, Wright A, Gustafsson J-Å, Carlstedt-Duke J (1991) DNA binding specificity of mutant glucocorticoid receptor DNA-binding domains. J Biol Chem 266:3101–3106

Zilliacus J, Wright APH, Norinder U, Gustafsson J-Å, Carlstedt-Duke J (1992) Determinants for DNA-binding site by the glucocorticoid receptor. J Biol Chem 267:24941–24947

Zilliacus J, Carlstedt-Duke J, Gustafsson J-Å, Wright APH (1994) Evolution of distinct DNA-binding specificities within the nuclear receptor family of transcription factors. Proc Natl Acad Sci USA 91:4175–4179

Zilliacus J, Wright APH, Carlstedt-Duke J, Nilsson L, Gustafsson J-Å (1995) Modulation of DNA-binding specificity within the nuclear receptor family by substitutions at a single amino acid position. Proteins Struct Funct Genet 21:57–67