

Molecular Dynamics Simulations of the Glucocorticoid Receptor DNA-Binding Domain in Complex with DNA and Free in Solution

Mats A. L. Eriksson, Torleif Härd, and Lennart Nilsson

Karolinska Institute Center for Structural Biochemistry, NOVUM S-141 57 Huddinge, Sweden

ABSTRACT Molecular dynamics simulations have been performed on the glucocorticoid receptor DNA binding domain (GR DBD) in aqueous solution as a dimer in complex with DNA and as a free monomer. In the simulated complex, we find a slightly increased bending of the DNA helix axis compared with the crystal structure in the spacer region of DNA between the two half-sites that are recognized by GR DBD. The bend is mainly caused by an increased number of interactions between DNA and the N-terminal extended region of the sequence specifically bound monomer. The recognition helices of GR DBD are pulled further into the DNA major groove leading to a weakening of the intrahelical hydrogen bonds in the middle of the helices. Many ordered water molecules with long residence times are found at the intermolecular interfaces of the complex. The hydrogen-bonding networks (including water bridges) on either side of the DNA major groove involve residues that are highly conserved within the family of nuclear receptors. Very similar hydrogen-bonding networks are found in the estrogen receptor (ER) DBD in complex with DNA, which suggests that this is a common feature for proper positioning of the recognition helix in ER DBD and GR DBD.

INTRODUCTION

The family of ligand-inducible transcription factors includes the receptors for steroid hormones, for example, the glucocorticoid receptor (GR) and the estrogen receptor (ER), as well as receptors for thyroid hormones, vitamin D₃, and retinoic acid. These proteins have a common amino acid sequence organization composed of discrete functional domains for ligand binding, a highly conserved domain for DNA binding (DBD) consisting of about 70 residues, and a domain for transcriptional regulation. The DNA binding sites for the GR are termed glucocorticoid response elements (GREs), and the naturally occurring GREs consist of 15 base-pair partially palindromic sequences comprising two hexameric half-sites with a three-base-pair spacing between the half-sites (Beato, 1989). DBD binds as a dimer to the GRE (see Härd and Gustafsson (1993) for a recent review of the GR DBD). Initial preferential binding to one of the two half-sites has been shown to facilitate binding to the other half-site (Tsai et al., 1988; Härd et al., 1990a). This is consistent with a two-site cooperative model, where DNA binding is dependent on both DBD-DNA interactions and interactions between the monomers at the dimerization interface.

The structure of a (DBD)₂-DNA complex (GRE_{S4}(DBD)₂) has been solved using x-ray crystallography (Luisi et al., 1991) to a resolution of 2.9 Å (Fig. 1). Instead of the naturally

occurring GRE with three basepairs between the half-sites, an idealized response element with a strictly palindromic sequence and a spacing of four basepairs between the half-sites (GRE_{S4}, see Fig. 2 A) was used in the crystal for which 2.9 Å data could be collected. The extra basepair in the spacer region makes it impossible for the two monomers to bind to their specific half-sites and at the same time keep the correct protein-protein interactions (Dahlman-Wright et al., 1991). Thus, one monomer (DBD_{spec}) forms three specific interactions with the bases in one half-site, whereas the other monomer (DBD_{unspec}) appears to bind cooperatively, but in an un-specific manner one basepair out of register with the correct binding site. Whether the interactions observed in DBD_{spec} are the same as those that would form to a native GRE with the correct spacing has not been established unambiguously.

The structure of monomeric GR DBD (Fig. 2) has been determined in solution using NMR spectroscopy (Härd et al., 1990a, b) and has recently been refined using two- and three-dimensional NMR spectroscopy on an ¹⁵N-labeled DBD fragment (Baumann et al., 1993). GR DBD consists of two subdomains (or motifs) that both have the general composition zinc-coordinating domain (Zn_I and Zn_{II})—α-helix (helix I and III)—extended region (ext_I and ext_{II}). The zinc ions are tetrahedrally coordinated to the sulphurs of four cysteine residues (Fig. 2), and the two major helices (I and III) are perpendicular to each other. The hydrophobic side chains of these helices and the following extended regions form the protein core (Härd et al., 1990b; Luisi et al., 1991). The two subdomains of DBD differ from each other both structurally and functionally. The first subdomain is involved in DNA-DBD interactions, with some of the residues in the Zn_I region forming specific interactions with the phosphate groups of DNA and the recognition helix (helix I) positioned in the DNA major groove (Fig. 1). Helix I contains three residues that have been shown to discriminate between glucocorticoid (GRE) and estrogen response elements (ERE)

Received for publication 11 October 1994 and in final form 7 November 1994.

Address reprint requests to Dr. Lennart Nilsson, Ctr. Structural Biology, Karolinska Inst., NOVUM, S-141 57 Huddinge, Sweden. Tel.: 46-8-608-9228; Fax: 46-8-608-9290; E-mail: lennart.nilsson@csb.ki.se.

Abbreviations used: DBD, DNA binding domain; ER, estrogen receptor; ERE, estrogen response element; GR, glucocorticoid receptor; GRE, glucocorticoid response element; MD, molecular dynamics; RMSD, root mean square deviation.

© 1995 by the Biophysical Society

0006-3495/95/02/402/25 \$2.00

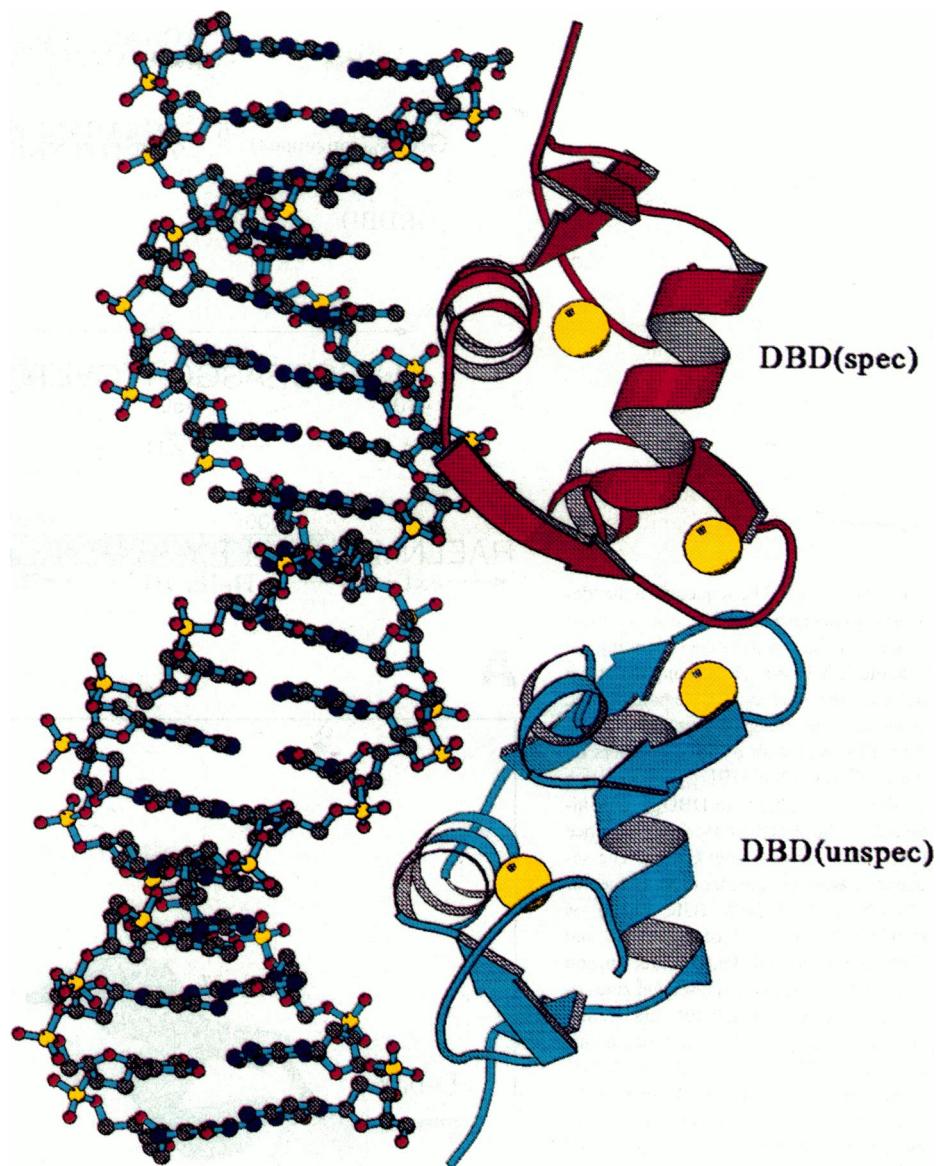


FIGURE 1 Crystal structure (Luisi et al., 1991) of the $(\text{GR DBD})_2\text{-GRE}_{84}$ complex. The specifically bound DBD (DBD_{spec}) is represented as a red ribbon, and the nonspecifically bound DBD ($\text{DBD}_{\text{unspec}}$) is represented as a blue ribbon. Zinc ions are shown as yellow spheres. This figure, and the other molecular graphics figure in this paper were drawn with the program Molscript (Kraulis, 1991).

(Danielsen, et al., 1989; Mader et al., 1989; Umesono and Evans, 1989; Zilliacus et al., 1991, 1992). The second subdomain provides the dimerization interface where most of the intersubunit interactions are made by residues in the loop between Cys-476 and Cys-482, which has been termed the "D-box" (Umesono and Evans, 1989, see Fig. 2). The second zinc-coordinating domain, Zn_{II} , also contains a short piece of distorted helix (helix II).

The agreement between the crystallographically determined DBD structure (Luisi et al., 1991) and the solution structure, determined with NMR (Härd et al., 1990a, b; Baumann et al., 1993), is quite good. There has been some discussion regarding the conformation and flexibility of the Zn_{II} fragment, which is well resolved in the crystal, but less well defined in solution in the earlier NMR-structure determinations (Härd et al., 1990a, b). The corresponding region is also poorly defined in the similar solution structure of the estrogen receptor DBD (ER DBD) (Schwabe et al., 1990,

1993a). It has been speculated that the conformation of this fragment is stabilized upon formation of the dimeric DBD-DNA-complex (Luisi et al., 1991). However, both $\{{}^1\text{H}\}\text{-}{}^{15}\text{N}$ NMR relaxation studies and a molecular dynamics simulation of monomeric GR DBD (Berglund et al., 1992; Eriksson et al., 1993) show that the picosecond mobility is limited and uniform for the entire protein with no pronounced increased mobility in the Zn_{II} fragment. Moreover, in a recent refinement of the NMR structure, the uncomplexed GR DBD is very similar to the DNA-bound form, and the Zn_{II} fragment including helix II is well resolved (Baumann et al., 1993). These results indicate that the additional stabilization of this fragment upon complex formation is of minor importance and that the poor resolution of this region in the initial structure determination (Härd et al., 1990b) was due to a lack of structural constraints rather than to inherently high flexibility. Still, small conformational changes appear to occur upon binding. For instance, the average plane formed by the back-

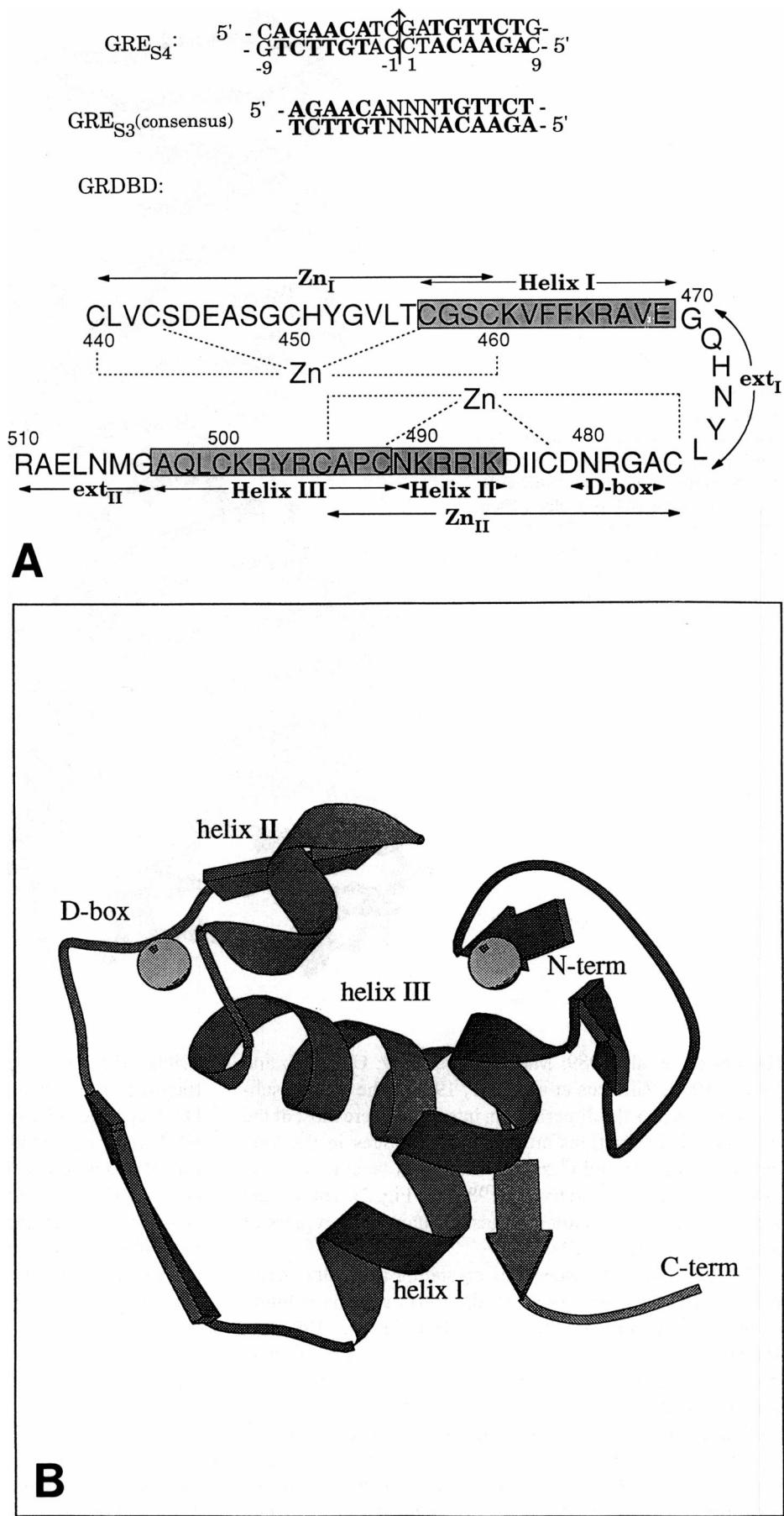


FIGURE 2 (A) The sequence of the idealized glucocorticoid response element (GRE_{S4}). Four basepairs are separating the hexameric half-sites (**bold**), and the arrow indicates the dyad axis of the palindromic sequence. The basepair numbering starts from this dyad with positive numbers on the side that binds to DBD_{spec} and negative on the side that binds to DBD_{unspec}, as indicated. The GRE consensus sequence (GRE_{S3}) is shown below GRE_{S4}. The sequence of the rat glucocorticoid receptor DNA-binding domain (GR DBD) is shown. Abbreviations of fragments and secondary structural elements are: Zn_I and Zn_{II}, first (N-terminal) and second zinc coordinating subdomains; helix I and III, first and second major α -helical subdomains; ext_I and ext_{II}, first and second extended regions and helix II is a short distorted helix in the Zn_{II} region. Helical regions are enclosed in shaded boxes. (B) The crystal structure of the GR DBD drawn in a ribbon representation with the two zinc ions represented by spheres.

bone of residues in the D-box is reoriented by about 90° compared with that of the crystal structure (Baumann et al., 1993). This conformational change within the D-box upon dimerization is not inconceivable because this region seems loosely anchored to other parts of the protein in the monomeric state.

The crystal structure of the estrogen receptor DNA-binding domain (ER DBD) bound as a dimer to the consensus estrogen response element (ERE) was recently solved to a resolution of 2.4 Å (Schwabe et al., 1993b). The two major differences between this complex and the GRE_{S4}-(GR DBD)₂ complex are that the two ER DBD monomers bind to equivalent half-sites on the ERE and that there are more base specific contacts at the ER DBD-ERE interface. These differences led to speculations whether the smaller number of base-specific interactions and ordered water molecules observed in the GRE_{S4}-(GR DBD)₂ complex was a consequence of the lower resolution (2.9 Å), if it was caused by the extra basepair between the half-sites in the GR complex or if it actually reflects different binding modes of the GR and ER DBDs when bound to their natural response elements (Schwabe et al., 1993b).

Molecular dynamics (MD) simulation is a powerful tool for analysis of structural and dynamic features of biomolecules in solution (Brooks et al., 1988), providing information about atomic interactions and their time evolution at a level of detail that is difficult to reach experimentally. Water molecules, which have been shown to be functionally important in other protein-DNA complexes (Aggarwal et al., 1988; Otwinowski et al., 1988; Rosenberg, 1991; Qian et al., 1993; Clark et al., 1993; Clore et al., 1994), are easily incorporated in MD simulations, where they can be characterized in terms of localization and mobility, properties that may be less straightforward to assess by other means (Levitt and Park, 1993). Most simulation studies to date have been concerned with protein systems (Brooks et al., 1988), and to a lesser extent with nucleic acids (for a recent review, see Beveridge and Ravishanker, 1994), and only very few protein-DNA complexes have been studied; preliminary data for the GR DBD-DNA system have been reported by Eriksson et al. (1994), as well as by Bishop and Schulten (1994), and the lac repressor headpiece with its operator was simulated by de Vlieg et al. (1989).

We have used molecular dynamics simulation to study the GR DBD in a dimeric complex with DNA and as a monomer in solution. The work presented here comprises three simulations (Table 1): a simulation of the dimeric complex with DNA starting from the GRE_{S4}-(GR DBD)₂ crystal structure, a simulation of a GR DBD monomer in solution using one of the monomers of the crystal structure as the starting structure, and a simulation of a GR DBD monomer in solution starting from the NMR structure. These simulations are denoted (DBD)₂-DNA, DBD_{mono}, and DBD_{sol}, respectively. All simulations were performed in water with electroneutralizing counterions. GR DBD backbone dynamics as calculated from the DBD_{mono} simulation have previously been reported

TABLE 1 Summary of notations and simulations

Simulation	Simulated species	Starting structure
DNA-(DBD) ₂	Dimeric DBD-GRE complex, with one specifically bound DBD (DBD _{spec}) and one unspecifically bound DBD (DBD _{unspec})	Crystal structure of (GR DBD) ₂ -GRE _{S4} (Luisi et al., 1991)
DBD _{mono}	Monomeric DBD in solution	DBD _{spec} from the (GR DBD) ₂ -GRE _{S4} crystal structure
DBD _{sol}	Monomeric DBD in solution	NMR solution structure (Baumann et al., 1993)

and found to agree well with {¹H}-{¹⁵N} NMR relaxation data (Eriksson et al., 1993).

The objectives of the present studies are to analyze direct and water-mediated interactions at the interfaces of the complex, to compare the structure and flexibility of the proteins in the complex with the monomeric uncomplexed species, and to relate the results to NMR and x-ray experiments. An analysis of differences in structure and flexibility of the sub-domains in relation to monomeric DBD sheds light on the stabilizing role of DNA. We focus on some of the issues discussed above, such as the stability of helix II in the monomeric form of DBD and the conformation of the D-box.

METHODS

Simulation details

Parameters and protocols

All simulations were performed using the program CHARMM (Brooks et al., 1983) version 22/23 and the all-atom version 22 force field (Chemistry Department, Harvard University, Cambridge, MA). Force field parameters for the zinc ions (which were treated as formal atoms in the calculations) were obtained by using MNDO (Dewar and Thiel, 1977a, b; Dewar and Merz, 1986)/ESP (Besler et al., 1990)-calculations with the program package MOPAC (Steward, 1990). Four fragments, each consisting of a cysteine with a methanoyl group at the N-terminal and an N-methyl amino group at the C-terminal, were covalently attached to the zinc ions in a tetrahedral geometry. To evaluate the partial charges of the zinc ion and the atoms in its vicinity, a full geometry optimization was applied in the MNDO/ESP-calculation (Table 2). Partial charges from the CHARMM (Brooks et al., 1983) v. 22 force field were used for atoms further away than C_β from the zinc ions. The partial charge at C_β was adjusted to give a sum of charges = -2 for the tetrahedral complex. The force constants of the bond (k_b) and angles (k_θ) involving the zinc were estimated by varying the actual bond or angle of the optimized structure while keeping all other bonds and angles fixed. The change of heat of formation plotted against the bond distance or the angle was fitted to a quadratic expression, assuming harmonic behavior of the bond-stretching and angle-bending motions. The minima of these

TABLE 2 Partial charges estimated with MNDO/ESP-calculations

Atom/ion	Charge
Zn	0.96
S	-0.72
C _β	-0.02
H (C _β)	0.00

curves were taken as the equilibrium values for the bonds (r_0) and angles (θ_0). The energy (V_n) of distorting the dihedrals around the zinc ion was set to zero, as in previous work on transition metal complexes (Hancock, 1989). The Lennard-Jones parameters (σ and ϵ) for zinc were adapted from Merz (1991), and the parametrization of the sulphur anion, using methyl- and ethylthiolate as model compounds, was based on structural and vibrational ab initio data of ethyl thiolate at the 6-31+G*-level (Department of Chemistry, Harvard University, Cambridge, MA) and the free energy of solvation of *N*-propylthiolate and thiolate (Pearson, 1986) (Table 3).

The simulations were performed with nonbonded interactions cut off by shifting the electrostatic potential to zero at 9.5 Å (Brooks et al., 1983). The nonbonded list including neighboring atoms within a 10.5 Å distance was updated every 50 fs, and a timestep of 2 fs was used. All bonds were constrained with the SHAKE algorithm (Ryckaert et al., 1977).

Simulation of the GRE_{S4} -(GR DBD)₂ complex

We used the crystallographic coordinates of the complex (Luisi et al., 1991), PDB entry 1GLU (Bernstein et al., 1977), with hydrogen atoms added by the CHARMM subroutine HBUILD (Brünger and Karplus, 1988). A short steepest descent minimization (50 steps) was applied to the system in vacuo to remove unfavorable contacts.

The complex was then inserted in the center of a sphere of radius 42 Å filled with TIP3P-water (Jorgensen et al., 1983), and all water molecules located closer than 2.8 Å from the nearest protein or DNA atom were removed. The water molecules interacted with a "deformable boundary force," arising from mean field interactions of water molecules beyond 42 Å (Brooks and Karplus, 1983). In the region between 39 and 42 Å (the buffer region), the water molecules interacted with a stochastic heat bath of 300 K, via randomly fluctuating forces and dissipative forces (Brooks and Karplus, 1983). For the water molecules inside 39 Å, the ordinary MD equations of motion were applied. Being aware of the difficulties to properly treat possible salt effects in this highly charged system, where the concentrations of ionic species close to especially DNA differ considerably from the bulk concentrations (see, for example, Misra et al., 1994), we chose as a first approximation to make the system electroneutral by adding sodium ions. This was done by replacing the 26 water molecules with the highest electrostatic energies of the oxygen atoms and that were located more than 5 Å apart from each other with sodium ions, which resulted in a system of 30,977 total atoms. The system was then minimized and equilibrated for 10 ps keeping the complex rigid. All atoms of the complex were then released, except for the atoms of the four bases at the ends of the DNA-helix, which were held close to their starting positions by harmonic constraints with a force constant of 10 kcal/(mol Å²). The constraints were applied to prevent the relatively short helix from unwinding and to maintain the base-pairing of the highly solvent exposed bases at the ends of DNA. The system was equilibrated for 50 ps, after which a production run of 150 ps was performed; the production run was started when the total potential energy of the system, the temperature, and the root mean square deviation (RMSD) of the atoms from the initial (crystal) structure had taken constant values (Fig. 3). A minor remaining drift of the RMSD values, however, can still be seen throughout the simulation (see Fig. 3 B). Constraints for the terminal DNA basepairs were removed during the last 50 ps of the production run to reveal possible overall bending or other conformational changes of the DNA-helix.

TABLE 3 Intramolecular force constants for atoms in the vicinity of the zinc ions

Bond Zn—S	k_B (kcal Å ⁻² mol ⁻¹) 78.0	r_0 (Å) 2.325
Angle S—Zn—S	k_θ (kcal rad ⁻² mol ⁻¹) 36.0	θ_0 (degrees) 109.5
C_β —S—Zn	36.0	116.3
Dihedral C_β —S—Zn—S	V_n (kcal mol ⁻¹) 0.0	γ (degrees) 0.0
Lennard-Jones parameters	σ (Å) Zn S	ϵ (kcal mol ⁻¹) 1.1 2.2

Simulation of the GR DBD monomers

These two simulations were carried out in a manner very similar to that of the complex, using a 30 Å radius sphere of TIP3P water. As starting structures we used DBD_{spec} from the crystal complex (Luisi et al., 1991) and the refined, energy-minimized NMR solution structure (Baumann et al., 1993), herein referred to as DBD_{mono} and DBD_{sol}, respectively (see Table 1). The monomers were electroneutralized with four chloride counter ions, and the production run was 200 ps for DBD_{mono}. Because of the larger drift in RMSD(t) of DBD_{sol} (Fig. 3 B), we prolonged that simulation to 325 ps and performed the analysis on the last 100 ps when RMSD(t) had leveled off.

During the equilibration phase, both the temperature and potential energy of the system increased for the DNA-(DBD)₂ and DBD_{sol}-simulations (see Fig. 3 A). This reflects that heat is flowing into the system to heat the complex/protein. The reverse is found during the equilibration of the DBD_{mono}-simulation as this system relaxes to a configuration with lower energy, and heat is instead flowing out of the system. This is not surprising because the starting structure for the protein in this case was DBD_{spec} from the DNA complex, simulated as a monomer in solution.

Analysis of the trajectories from the simulations

Hydrogen bonds and water bridges

The criteria used for a hydrogen bond (A···H—D) was that the acceptor (A)-donor (D) distance should be less than 3.5 Å and the angle A—H—D should be larger than 135°. Only those hydrogen bonds that were present during more than 20% of the simulation runs and those water molecules that occupied the same bridging positions for more than 20% of the simulation runs were considered.

DNA structure

The helix axis, helicoidal parameters, and backbone torsional angles were estimated using the program "Dials and Windows" (Ravishankar et al., 1989), which is based on the program "Curves" (Lavery and Sklenar, 1988).

Counterion distribution

The most probable positions for the sodium counterions in the GRE_{S4} -(GR DBD)₂ complex were determined using a three-dimensional histogram with a (1 Å)³ bin size. The histogram was constructed from ion coordinates sampled every picosecond of the production run. Only probabilities larger than 10% were considered in the density map.

Root mean square deviations (RMSD)

Coordinate sets from every picosecond of the production run were rotated to fit the initial structure of the molecule(s) as close as possible by minimizing the mass-weighted root mean square deviation of the heavy atoms from the initial structure. Average RMSD values of the backbone (i.e., C, O, C_α, H_α, N, and HN) and the side chains were then calculated for each residue.

Fluctuations

B-factors from the x-ray study of the complex (Luisi et al., 1991) were compared with the atomic fluctuations in the simulation using the relation

$$\langle \Delta r_i^2 \rangle = 3B_i/(8\pi^2), \quad (1)$$

where Δr_i is the atomic displacement for atom i and B_i is the corresponding B-factor. Fluctuations of the atoms were calculated after applying the rotation of the actual coordinate sets, as described above.

Accessible surface area

Solvent-accessible surface areas (a.s.a.) were calculated using the definition of Lee and Richards (1971) with a probe of radius 1.6 Å. The fractional a.s.a.

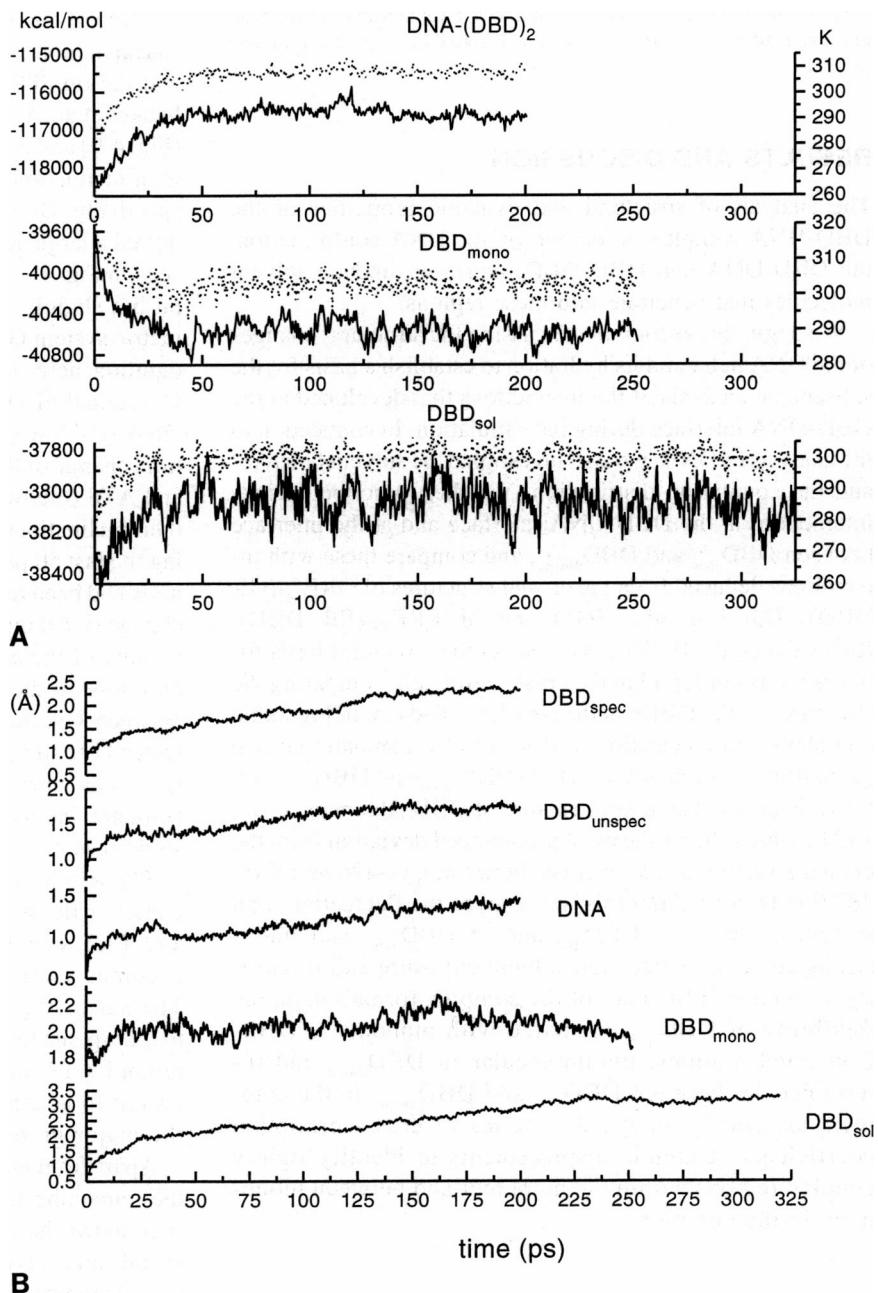


FIGURE 3 (A) Time evolution of the total potential energy (in kcal mol^{-1}) of the systems during the simulations. Dotted lines represent the time evolution of the temperature (in K). (B) Time evolution of the root mean square deviations (RMSD) from the initial structures of the simulations.

was obtained by dividing the backbone or side-chain a.s.a. for each residue with the “free” a.s.a. that was calculated for the atoms of the backbone or the side chains in a residue without taking any neighboring atoms into account.

Cross-correlation coefficients for atomic displacements

Concerted atomic motions can be identified by analyzing the cross-correlation coefficients for atomic displacements (McCammon and Harvey, 1986) defined between two atoms i and k by the expression

$$C_{ik} = \langle \Delta\mathbf{r}_i \cdot \Delta\mathbf{r}_k \rangle / (\langle \Delta\mathbf{r}_i^2 \rangle \langle \Delta\mathbf{r}_k^2 \rangle)^{1/2}, \quad (2)$$

where $\Delta\mathbf{r}_i$ is the displacement from the mean position of the i th atom. For completely correlated motions $C_{ik} = 1$, and for completely anticorrelated motions $C_{ik} = -1$. Deviations from 1 (or -1) implies either that the motions of i and k are less correlated (or anticorrelated) or that the vectors along which they move are nonparallel (or nonantiparallel). Thus, examining C_{ik}

is useful for identifying atoms moving in a correlated manner provided that the angle between the directions of the motions does not deviate too much from 0° or 180° . We estimated C_{ik} values of the backbone of DBD_{mono}, and between the backbones of DBD_{spec} and DBD_{unspec} as the mean of five and four 40 ps block averages, respectively. The atomic displacements of the backbone atoms were replaced by the average displacements of these atoms in each residue.

Diffusion coefficients

We estimated the diffusion coefficients for the water molecules ($D_{\text{H}_2\text{O}}$) and the sodium ions (D_{Na^+}) in the simulated complex using the Einstein relation (Atkins, 1990)

$$\lim(t \rightarrow \infty) \frac{\partial}{\partial t} \langle [\mathbf{R}(t) - \mathbf{R}(0)]^2 \rangle = 6D, \quad (3)$$

where $\langle \rangle$ denotes an average over the water molecules/ions and $\mathbf{R}(t)$ is

the position of a water molecule/ion at time t . To obtain an error estimate of the self-diffusion coefficients, we recalculated Eq. 3 for different time origins.

RESULTS AND DISCUSSION

The analysis of structural and dynamic properties of the DBD-DNA complex is focused on the DNA conformation, the DBD-DNA and DBD-DBD interfaces, and on solvent molecules that penetrate into these regions.

We begin this section by discussing the structural changes of the DNA helix and its hydration to establish a basis for the subsequent analysis of the interactions that developed at the DBD-DNA interface during the simulation. In connection to the analysis of DNA, we also investigate the solvent behavior and the counterion distribution. We then describe various interactions at the DBD-DNA interface and at the interface between DBD_{spec} and DBD_{unspec}, and compare these with interactions deduced from the crystal structures of GRE_{S4}-(GR DBD)₂ (Luisi et al., 1991) and of ERE_{S3}-(ER DBD)₂ (Schwabe et al., 1993b). We discuss the structural basis for interactions not found in the crystal structure, comparing the structures of the DBDs in the simulated and crystallographic complexes. In connection to this, we also compare the two simulations of monomeric DBD (DBD_{mono} and DBD_{sol}) with the refined NMR solution structure of DBD (Baumann et al., 1993), which shows the most pronounced deviation from the crystal structure at the residues between Cys-476 and Cys-482 (the D-box). After this, we analyze the fluctuations and solvent exposure of DBD_{spec} and of DBD_{mono} and find a strong correlation between solvent exposure and flexibility, as well as influences of the complex formation on the flexibility of DBD_{spec} compared with monomeric DBD. Concerted motions, intramolecular in DBD_{mono} and intermolecular between DBD_{spec} and DBD_{unspec} in the complex, are finally analyzed in terms of cross-correlation coefficients of atomic displacements to identify tightly coupled regions within the monomer and between monomers in the complex.

Structure and hydration of DNA in the complex

The helix axis of DNA in the crystal structure is slightly bent in the spacer region between the two half-sites (Fig. 4 A). This bend becomes more pronounced during the simulation, and a second bend is also developing in the region of DNA at the location of the recognition helix of DBD_{unspec} (Fig. 4 A). The DNA helix axis remains more or less straight in the DBD_{spec} half-site region (Fig. 4 A, top) throughout the simulation. The bend in the spacer region (approximately 10° relative to a straight DNA helix axis) is a consequence of new contacts that form between the ext₁ region of DBD_{spec} and the DNA (to be discussed below), pulling the DNA of the spacing region closer toward the dimer. The bend is manifested predominantly as a large increase in the axis tip angle (ATP in Fig. 4 B) to around 20° toward the major groove for the global helix axis segment between CG1 and GC-1. The ori-

entation of the basepairs with respect to the global axis in the spacer region, which is described by the inclination (INC) and the tip (TIP) angle (Lavery and Sklenar, 1988; Ravishankar et al., 1989), also deviates from the starting structure (Fig. 4 B) and is especially pronounced for the basepair GC-1 with inclination and tip angles of around -10° and 10°, respectively. During the simulation this basepair has also acquired a large negative propeller twist (-40°), which can be seen in Fig. 4 A. Bishop and Schulten (1994) observe a similar but slightly larger effect in their simulation of the symmetric system GRE_{S3}-(GR DBD)₂. The bend around the recognition helix of DBD_{unspec} is due to the approach of the C-terminal of DBD_{unspec} (see below) to the last basepair of DNA (the lowest in Fig. 4 A). The most pronounced structural change of the DBD_{spec} half-site is that the basepairs CG9 and AT8 (the two basepairs at the top in Fig. 4 A) increase their inclinations relative to the helix axis to around 20° during the last 50 ps, i.e., when the constraints at the ends of the helix had been removed. The reason for this relatively sudden change is that the helix shows tendencies of unwinding in the vicinity of the base G9, seen as a small displacement of the base toward the bulk water and an increase of the distance between G9 and T8. To maintain the hydrogen bonds between G9 and C9, the basepair starts inclining relative to the helix axis when the end constraints are released, and a concomitant inclination of the basepair AT8 then occurs for steric reasons.

The width of the minor groove (Table 4) remains fairly constant along the helix, except in the region around GC7 and TA2, where the groove is initially more narrow and becomes increasingly narrower during the simulation. The narrowing of the minor groove at TA2 has occurred to compensate for the expansion of the major groove around AT3 and AT-2 (Table 4), which in turn has occurred to accommodate better the recognition helices in the major groove of DNA.

Apart from above mentioned structural changes of DNA, the remaining helicoidal parameters of the basepairs/bases (not shown here) are quite stable and fluctuate around the initial values. Backbone torsional angles are also quite stable (not shown here), and the sugar pucker remains in the initial C2'-endo conformation (Luisi et al., 1991), characteristic for B-DNA (Saenger, 1984). Exceptions to this are both nucleotides in the basepairs GC-1 and GC-4, which have adopted the C3'-endo conformation of the sugar pucker during the simulation.

Crystal-packing forces between adjacent DBD fragments and the formation of end-to-end triple helices with Hoogsteen-like basepairing (Hoogsteen, 1963) of GRE_{S4} in the crystal (Luisi et al., 1991) probably prevented the small structural changes mentioned above to occur in the crystal. In the absence of such forces, as in our simulation, the helix can adjust more easily to facilitate interactions with the DBD-dimer.

Average residence times of bridging water molecules at different interfaces and surfaces are summarized in Table 5. A large fraction (56%) of the water molecules that were

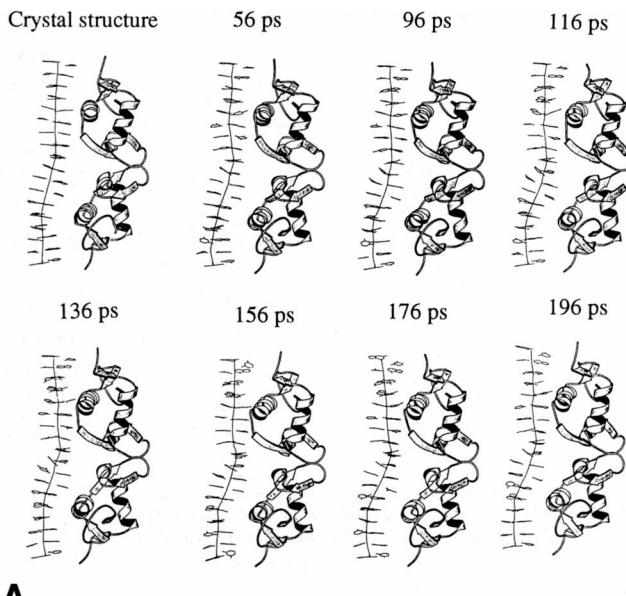
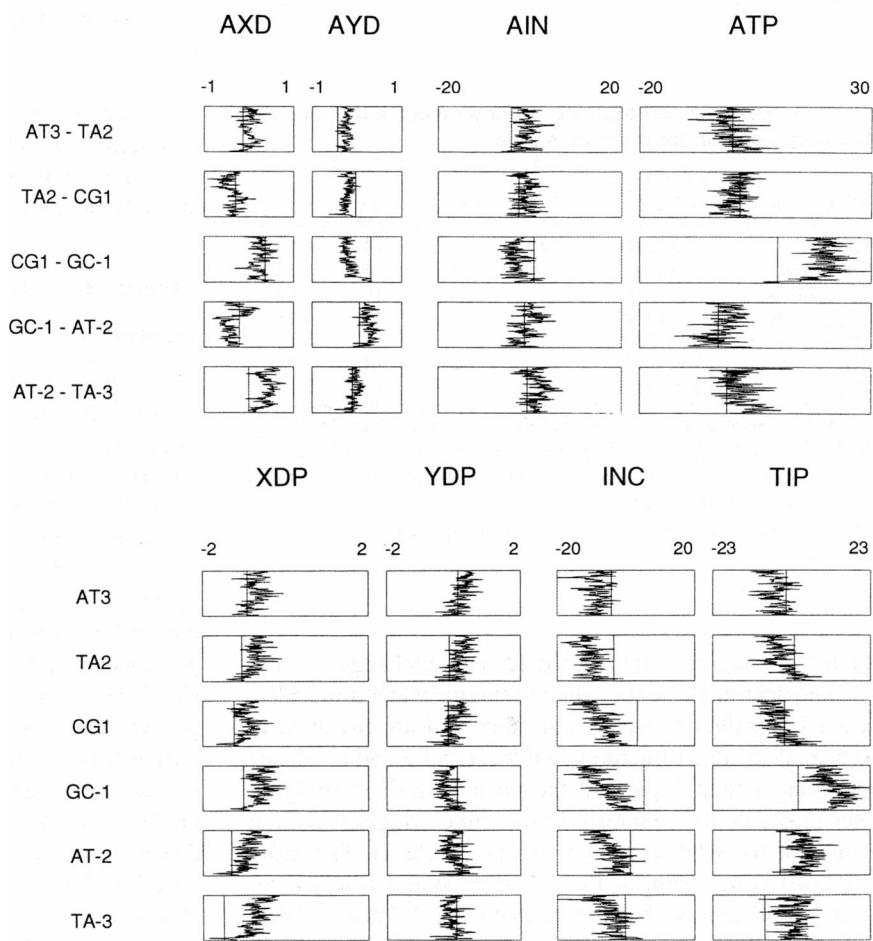
**A**

FIGURE 4 (A) Snapshots of the $(GR DBD)_2$ -DNA-complex during the simulation. The first 50 ps of the total simulated time of 200 ps constitute the equilibration phase. The helix axis of DNA, shown in the figure, was calculated using the program “Dials and Windows” (Ravishanker et al., 1989; Lavery and Sklenar, 1988). (B) Time evolution of the position and orientation of the global axis orientation (*top*) relative to a straight axis and of the position and orientation of the basepair relative to the global helix axis (*bottom*) in the spacer region of DNA according to the definitions by Lavery and Sklenar (1988). AXD = *x* axis displacement (\AA), AYD = *y* axis displacement (\AA), AIN = axis inclination ($^\circ$), ATP = axis tip ($^\circ$), XDP = *x* displacement (\AA), YDP = *y* displacement (\AA), INC = inclination ($^\circ$) and TIP = tip ($^\circ$). The positive *x* axis is directed toward the major groove, and positive angles correspond to right-handed rotation about the *x* axis (inclination) and *y* axis (tip), respectively. The parameters are displayed in “windows” (Ravishanker et al., 1989) where the start of the simulation corresponds to the bottom of a window. The line through the windows are the parameters for the starting (crystal) structure.

**B**

forming bridges between polar groups of DNA had an average residence time less than 2 ps (Table 5), which can be compared with the characteristic hydrogen bond lifetimes of 0.7 ps seen in simulations of TIP3P water (Norberg and

Nilsson, 1994). Water bridges of this type were mainly found between sugar O5' and sugar O3'/phosphate oxygens of the same nucleotide. The short residence times are a reflection of the closeness to the surrounding bulk water molecules with

TABLE 4 Average groove widths in GRE_{s4} during the simulation

Base pair	Minor groove (Å)	Major groove (Å)
GC7	9.0 (9.4)	
AT6	10.3 (9.5)	17.2 (17.7)
AT5	11.5 (11.7)	19.0 (18.3)
CG4	11.7 (10.7)	20.6 (18.0)
AT3	11.0 (11.3)	21.0 (19.2)
TA2	9.4 (10.6)	19.1 (19.8)
CG1	12.5 (13.6)	19.0 (18.4)
GC-1	11.7 (12.9)	19.3 (18.9)
AT-2	12.8 (13.8)	20.8 (19.3)
TA-3	12.4 (13.0)	21.3 (19.3)
GC-4	12.2 (13.1)	15.7 (16.3)
TA-5	12.2 (13.4)	19.4 (19.0)
TA-6	11.7 (12.4)	16.6 (16.9)
CG-7	12.7 (12.4)	
Canonical B-form	11.7	18.1

The groove widths were measured as the minimum distance between phosphorus atoms on either side of the groove. In B-DNA this distance connects the phosphate groups located two base steps away in the 5' direction from the base pair to which the width is assigned (Ravishankar et al., 1989). Base pairs in bold are located within the binding (half) sites. (Base pairs with positive numbers bind to DBD_{spec} and the values within parentheses are those of the crystal structure.)

TABLE 5 Statistics on bridging water molecules that occupy a bridging position for more than 2 ps

Bridges in:	No. of water molecules*	<25%	<50%	<75%	≤100%
DNA	183 (231)	4	7	14	133
DNA-“spine”‡	64 (35)	4	11	24	133
DBD _{spec} -DNA	19 (12)	5	9	41	141
DBD _{unspec} -DNA	25 (24)	6	15	21	137
DBD (complex)	390 (329)	4	8	18	136
DBD _{mono}	231 (234)	4	8	18	107
DBD _{sol}	186 (357)	4	7	16	81

The numbers indicate at what residence time (in ps) less than 25, 50, 75, and 100%, respectively, of the bridging water molecules were found.

*The number of discarded bridging water molecules (with residence times <2ps) is given in parentheses.

†Bridges between a purine N3/pyrimidine O2 and a sugar O4' of the next nucleotide.

which the bridging water molecules can exchange readily. When neglecting these short-lived bridging water molecules, we are left with 183 water molecules that are involved in more stable bridge formations. Of these, 50% had residence times longer than 7 ps, and the most long-lived bridging water molecule has remained in the same bridging position for almost the whole production run (Table 5). The more long-lived water bridges include intra-strand bridges between A6-N7 and A5-N6 and between G-1-N7 and A-2-N6. However, the majority of the more long-lived water bridges in DNA were formed between purine atom N3/pyrimidine atom O2 and the sugar oxygen O4' of the next nucleotide on the same strand in the minor groove. The one-water bridges in the minor groove are one of the characteristic features of the B-form of DNA (the “spine of hydration”) and have been found in crystal structures of B-DNA (Drew and Dickerson, 1981; Kopka et al., 1983) and at the ends of the helix in the

crystal complex of ERE-ER DBD (Schwabe et al., 1993b) as tightly bound water molecules. 50% of the water molecules forming “spine” bridges had residence times longer than 11 ps (Table 5), and 25% of the water molecules in the “spine” remained in the same bridging position for more than 24 ps. Comparing these residence times with the corresponding times for all bridging water molecules in DNA, 7 ps (<50%) and 11 ps (<75%), respectively, shows that the subset of water molecules involved in the “spine of hydration” is forming more stable bridges. Another indication of the higher stability is the smaller fraction of water bridges with residence times less than 2 ps, 35% in the subset of “spine” water compared with 56% of all bridging water molecules in DNA. The main purpose of the residence times given in Table 5 is to reveal differences in the distribution of residence times between different interfaces and molecules of the complex and in the monomer, rather than to estimate the lifetimes of the bridges. Nevertheless, a comparison can be made with results from an MD simulation of the left-handed Z-form of DNA (Eriksson and Laaksonen, 1992), where the lifetime, i.e., the time when 50% of the bridging positions had been replaced with a new water molecule, for the one-water bridges was estimated to 15 ps. This is clearly larger than would be obtained for the water bridges in the B-DNA in our simulation, and the difference probably reflects the fact that the water molecules in the vicinity of the DNA are more strongly held by the Z-DNA (Westhof, 1987).

Solvent distribution and diffusion

Counterion distribution in the DNA-(DBD)₂-complex

There were 11 positions in the counterion density map with a probability of >10% to contain a sodium ion (Fig. 5). All of these positions are close to the DNA, and the majority are located close to or inside the DNA minor groove. Several of these locations are at a considerable distance (>10 Å) from the initial ion positions. The higher mobility of the counterions in the major groove and the more distinct positions in the minor groove result from the higher density of negatively charged phosphate oxygens around the minor groove and the presence of more electronegative atoms (pyrimidine-O2/purine-N3) in the minor groove compared with the major groove. Three of the ions were closer to the pyrimidine-O2/purine-N3 (around 4.4 Å) than to the closest phosphate group (around 6 Å). Because the distance between Na⁺ and the O2/N3-atoms coincides with the second peak in the radial distribution function g(Na⁺-O) of sodium ions in water (~4.5 Å, Impey et al., 1983), it is likely that these ions are attached to the minor groove together with their first hydration shell. The ion-coordinating waters are then hydrogen-bonded to pyrimidine-O2/purine-N3 atoms and thereby lock the position of the ion inside the groove. The other Na⁺ ions (except for one, which is located further out from the DNA) are instead coordinating the phosphate groups along the backbone of the minor groove. In an x-ray study of the Cs⁺-distribution around CsDNA in the B-form (Bartenev et al., 1983), the cesium ions were forming strings along both the

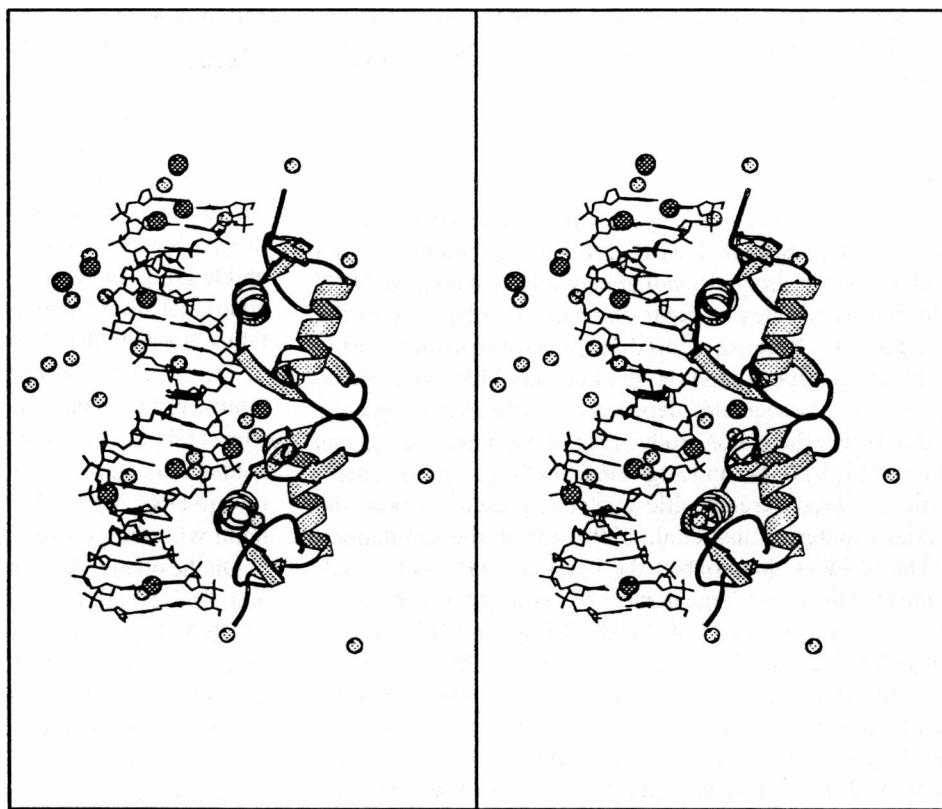


FIGURE 5 Stereo view of the counterion (Na^+) distribution (dark spheres) around the DNA-(DBD)₂ complex, according to the probability density map. The initial positions of the counterions are shown as lighter spheres.

minor and major groove, with the ions in the minor groove coordinating pyrimidine-O2/purine-N3 and -N2 of the bases. In contrast, the cesium ions in the major groove were located predominantly at the DNA surface and closer to the phosphate groups. Bartenev et al. (1983) conclude that, *in solution*, the ions in the major groove would be as mobile as the other cations in the vicinity of the phosphate groups, and their counterion coordination is therefore in good agreement with the most probable positions of the sodium ions in our simulation. Furthermore, a mixture of base- and phosphate-coordinating counterions was also observed in an MD simulation of a five-basepair fragment of B-DNA with Na^+ counterions (Seibel et al., 1985). Base-coordinating counterions (Li^+ , Na^+ , and K^+) have also been found in Monte Carlo studies of B- and Z-DNA (Clementi and Corongiu, 1983) and in an MD simulation of the Z-form of DNA (Laaksonen and Nilsson, 1989; Eriksson and Laaksonen, 1992).

It should be pointed out that the relatively short simulated time of the DNA-(DBD)₂-complex only allows us to sample a small fraction of the total counterion configuration space. To make more definite and statistically more reliable statements about the locations of the counterions would require a considerably longer simulation time.

Diffusion of water and counterions

From the mean square displacement (Eq. 3) of the water molecules and Na^+ counterions in the simulated complex, we estimated the diffusion coefficients, $D_{\text{H}_2\text{O}}$ to $1.7 (\pm 0.05) \cdot 10^{-9} \text{ m}^2/\text{s}$ and D_{Na^+} to $0.5 (\pm 0.1) \cdot 10^{-9} \text{ m}^2/\text{s}$, respectively.

The experimental value for $D_{\text{H}_2\text{O}}$ at 300 K is $2.3 \cdot 10^{-9} \text{ m}^2/\text{s}$ (Mills, 1973), and in pure TIP3P-water (Jorgensen et al., 1983) at 300 K $D_{\text{H}_2\text{O}}$ has been estimated to $1.3\text{--}4.2 \cdot 10^{-9} \text{ m}^2/\text{s}$ (Tasaki et al., 1993; Norberg and Nilsson, 1994). The sodium ion diffusion is low because of the preference for Na^+ to be located inside or at the surface of the minor groove (see above) and is probably also caused by the slow diffusion of the counterions in a direction perpendicular to the helix axis. In a 80 ps MD simulation of an eight-basepair fragment of NaDNA in the B-form (van Gunsteren et al., 1986), D_{Na^+} values in the range $0\text{--}5.0 \cdot 10^{-9} \text{ m}^2/\text{s}$ were obtained depending on the location of the counterion. In longer simulations of Na^+ and a dinucleotide in TIP3P water D_{Na^+} was calculated to $1.3 \cdot 10^{-9} \text{ m}^2/\text{s}$ (Norberg and Nilsson, 1994). Further, in MD-simulations of various alkali-halide salts in water (Impey et al., 1983) D_{Na^+} was estimated to $1.0 \cdot 10^{-9} \text{ m}^2/\text{s}$, and the experimental value at 300 K is $1.33 \cdot 10^{-9} \text{ m}^2/\text{s}$ (Atkins, 1990). Our estimate of D_{Na^+} therefore appears reasonable.

Hydrogen-bonding network at the DBD-DNA interface

The overall hydrogen-bonding patterns between protein and DNA are similar in the simulated and crystallographic structures, with certain differences between the two half-sites. A number of new, mainly water-mediated hydrogen bonds show up very clearly in the simulation as water molecules find good bridging locations at the interface; there are more

interfacial waters in the simulation than in the starting structure. These findings are discussed below, focusing on the water bridges.

DBD_{spec}-DNA

The interface between the specifically bound DBD and the corresponding half-site in GRE_{S4} (Fig. 6, *top* and Fig. 7, *left*) contains two hydrogen-bonding networks that connect DBD with the phosphates or sugar oxygens on either side of the major groove. Because of this division, DBD can measure the width of the groove, and the binding should in principle be sensitive to any sequence-dependent variations in the groove width. Inside the major groove (Fig. 6, *top*), the side chains of Lys-461, Val-462, and Arg-466 from the recognition helix form the three base-specific contacts, present in both the crystal structure (Luisi et al., 1991) and in the simulation.

The residues that contact the bases T5, G4, and T3 are connected through a large number of hydrogen bonds to residues from various parts of the DBD fragment (Fig. 6, *bottom*). We note that all residues that participate in this network of hydrogen bonds are either conserved throughout the family of nuclear receptors (Laudet et al., 1992) or essential for the discrimination between a glucocorticoid response element and an estrogen response element (Ser-459 and Val-462) (Danielsen et al., 1989; Mader et al., 1989; Umesono and Evans, 1989; Zilliacus et al., 1991; Zilliacus et al., 1992). The central residues in this hydrogen-bonding network are Asp-445 in the Zn_I region, Arg-489 in helix II, and Arg-496 in helix III, which connect to the phosphates of GRE and to the other residues in this network via hydrogen bonds and water bridges from both the backbone and the side chains. Arg-489—O is connected via a water molecule to the phosphate group of T3 and to Arg-496-N_{η1}. This water bridge was found to be rather stable during the simulation, with a mean residence time of 52 ps (Table 6), and one of the bridging water molecules was a crystal water. Arg-489—N_{η1} forms a stable (Table 7) hydrogen bond to the phosphate group of G4 and water-mediated contacts with Ser-459—O_γ and with Asp-445—O_{δ2} (Fig. 6, *bottom*). The same water molecule that is bridging Arg-489—N_{η1} and Ser-459—O_γ is also connected to the phosphate group of T5 and to Asp-445—O_{δ2}, so this water molecule is probably important for giving these residues the proper orientation upon binding to GRE. Arg-489—N_{η2} forms hydrogen bonds to Asp-445—O_{δ2} and to Cys-443—S_γ. The position of Ser-459 is fixed with a hydrogen bond from the side-chain hydroxyl to the sulphur of Cys-457 and with a water-mediated contact between its carbonyl group and the phosphate group of G4. Arg-496 is also one of the most central residues in this hydrogen-bonding network with its side chain hydrogen-bonded to the phosphate of G4, the sulphur of Cys-460, and via a trifurcated water bridge to the phosphate group of T3 and the carbonyl oxygen of Arg-496.

The residues Arg-489 and Arg-496 have strikingly similar hydrogen-bonding patterns with N_{η2} hydrogen-bonded to cysteine sulphurs, N_{η1} to phosphate groups, and with the

carbonyl group fixed by a water bridge in Arg-489 and by Cys-500—S_γ in Arg-496. It is notable that three of four zinc coordinating cysteines in the Zn_I domain (Cys-440 being the exception) participate with their sulphurs to the stabilization of the hydrogen-bonding network.

Other contacts on the same side of the major groove that are absent in the crystal structure are strong (Table 7) hydrogen bonds from the side chain of His-472 and from the amide group of Asn-473 to the phosphate group of A2. The additional water-mediated contacts from Asn-473—O and Gln-471—O to the phosphate group of A2 thus involve most of the residues from the ext_I region at the DBD_{spec}-DNA interface, according to our results. The side chain of Lys-490 from helix II, which in the crystal structure forms a water-mediated contact to the phosphate group of C1, has turned toward DNA during the simulation to form a stable hydrogen bond with the phosphate group of G-1.

On the other side of the groove, the side chains of His-451 and Tyr-452 are hydrogen-bonded to the phosphate groups of DNA. These contacts are also present in the crystal structure, whereas the relatively weak (Table 6) water-mediated contact between Cys-450—O and the phosphate group of A8 was found in the simulation only.

The observation of rather few ordered water molecules at the DNA-DBD_{spec} interface of the crystal structure compared with the present simulation might be a consequence of the relatively low 2.9 Å resolution (Luisi et al., 1991), and possibly because the crystal packing prevents the DBD fragments to interact close enough to allow the formation of as many water bridges at the interface as was found in the simulation.

DBD_{unspec}-DNA

Because of the extra spacing basepair in GRE_{S4}, the base sequence of the second half-site is shifted one basepair of register compared with the natural GRE. This shift naturally results in drastic differences in the binding of DBD_{unspec} compared with DBD_{spec}. Only one of the three base-specific interactions involving the recognition helix of DBD_{spec} in its specific site is formed also by DBD_{unspec}, namely, the hydrogen bond between the side chain of Lys-461 and atom N7 of G-7, which is present in both the crystal structure and in the simulation (Figs. 8 and 9). All other interactions are formed with the phosphate groups/sugar oxygens of the DNA backbone. As was the case with DBD_{spec}, many water-mediated interactions that were not found in the crystal structure appear in the simulation. The only new direct hydrogen bond (or, rather, salt bridge) that had developed during the later part of the simulation was between the side chain of Lys-465 and the phosphate group of A-6. This interaction, which is also observed for both DBD monomers in the simulation of the GRE_{S3}-(DBD)₂ system (Bishop and Schulten, 1994), was achieved by distorting the recognition helix to allow the side chain of Lys-465 to protrude into the groove, discussed more below.

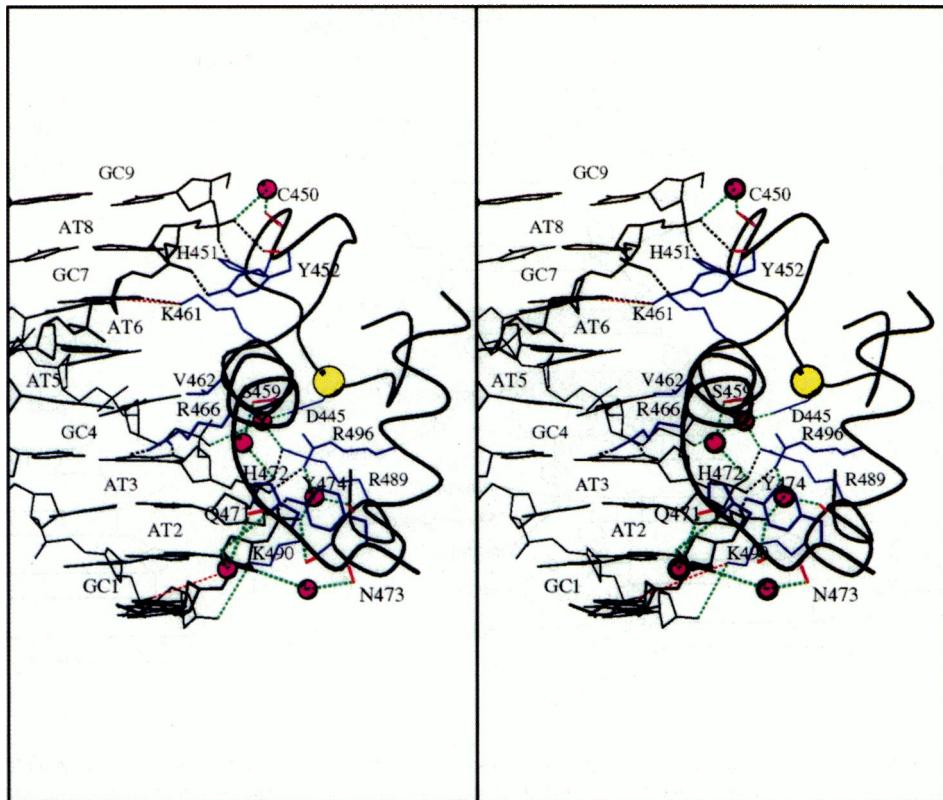
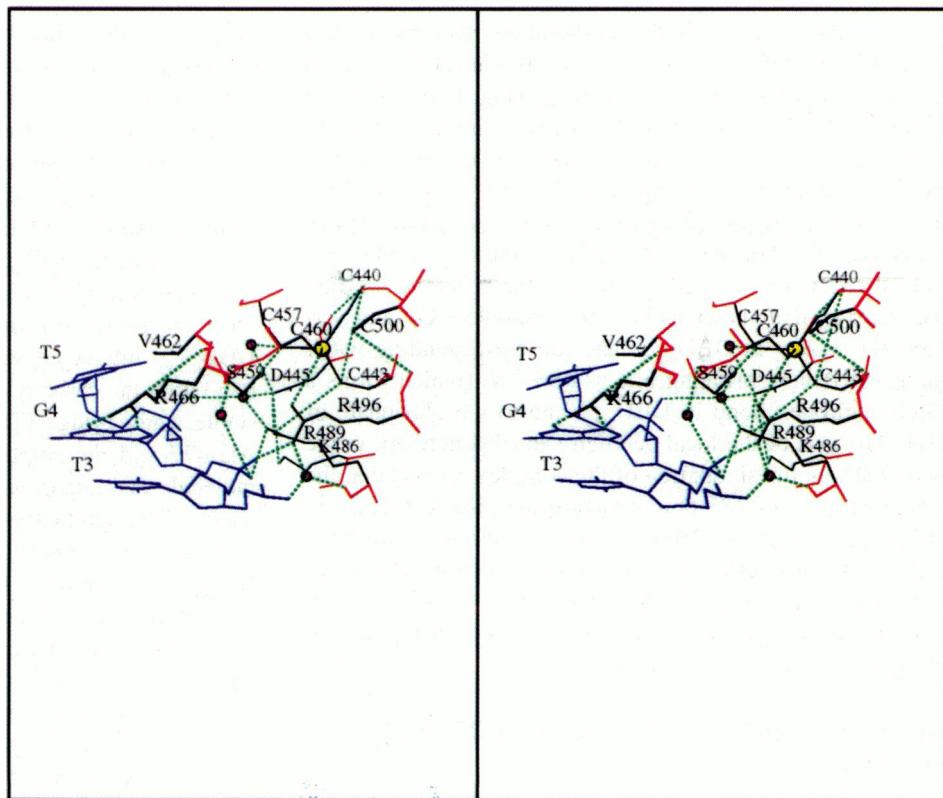


FIGURE 6 (top) Stereo view of the DBD_{spec}-DNA interface (crystal structure). Bridging water molecules (present for more than 20% of the production run) are shown as magenta spheres. Black dotted lines are interactions that are found in both the crystal structure and in the simulation, red dotted lines are contacts present only in the crystal structure, and green dotted lines represent contacts that are found in the simulation only. (bottom) A detailed stereo view of the hydrogen-bonding network (green dotted lines) between DBD_{spec} and DNA at the side of the major groove that is closest to the spacer region. Backbone atoms of the protein are connected with red lines, side-chain atoms with black lines, and the DNA fragment is colored blue. Magenta spheres are bridging water molecules, and the yellow sphere represents a zinc ion.



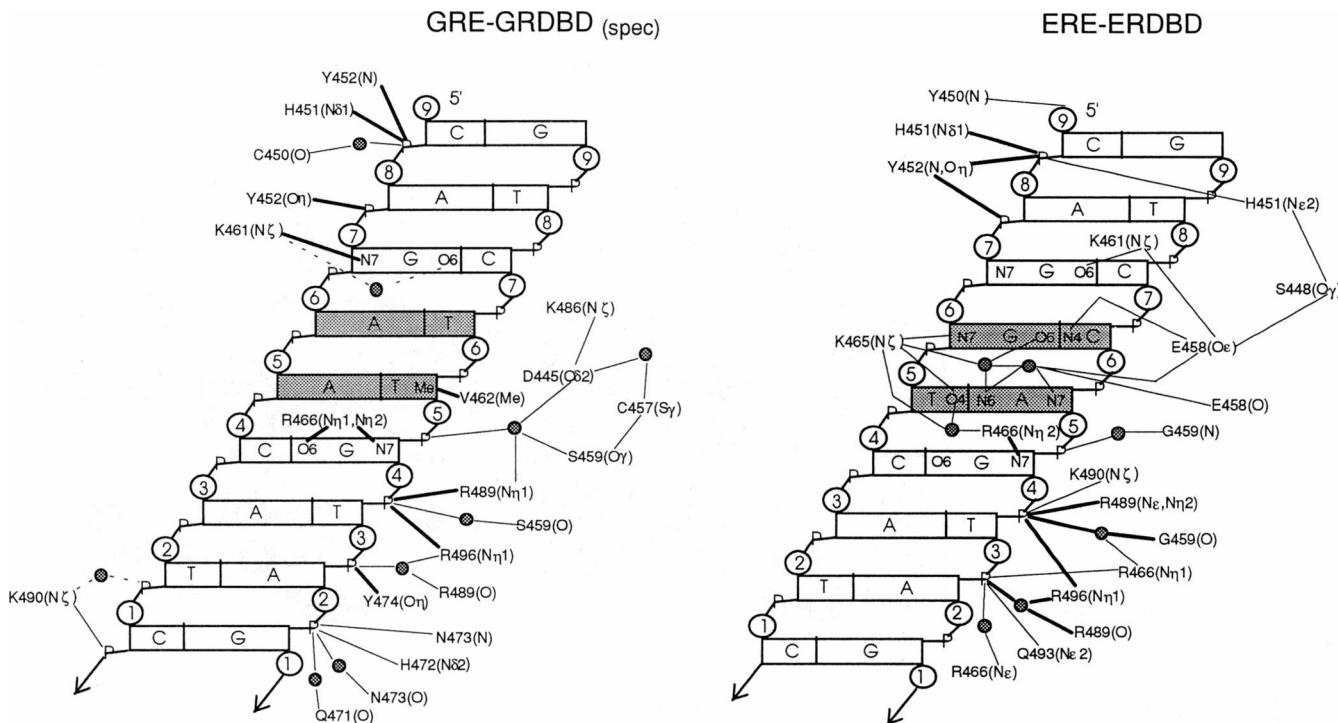


FIGURE 7 (left) A schematic drawing of the contacts formed at the DBD_{spec}-DNA interface of the simulated complex. Thicker lines represent contacts found in both the simulation and in the crystal structure, thinner lines are contacts found in the simulation only, and dotted lines are contacts that are lost during the simulation but that are present in the crystal structure. The circles represent bridging water molecules, present at least 20% of the production run. (right) A corresponding drawing of the contacts that are found in the crystal complex between the estrogen response element (ERE) and the DNA-binding domain of the estrogen receptor (ER DBD, Schwabe et al., 1993b). Bold lines in this drawing represent contacts that are also found in our simulation and/or in the crystal structure of the GR DBD-DNA complex. The two shaded basepair are those that differ between the GRE and ERE half-sites.

If we compare the hydrogen-bonding network around Asn-445, Ser-459, Arg-489, and Arg-496 in DBD_{unspec} with the corresponding network in DBD_{spec} (Fig. 6, bottom), we find many similarities. Arg-496 has a very similar hydrogen-bonding structure in DBD_{unspec} with its carbonyl group hydrogen-bonded to the sulphur of Cys-500. Arg-496—N_{η1} interacts with the phosphate of T-3, via a water bridge with the carbonyl of Arg-489, and with N_{η2} to the sulphur of C460, as is also the case for DBD_{spec}. One of the differences is that the water-mediated contact between Asp-445—O_{δ2} and Ser-459—O_γ is lost in DBD_{unspec} and the corresponding water molecule is now bridging to Lys-486—N_ζ (which forms a direct hydrogen bond in DBD_{spec}) and to the phosphate of G-4. Analysis of the local geometry for this network in the two DBDs in the simulation of the complex reveals that the side chain of Lys-486 points toward the protein surface in DBD_{unspec}, whereas in DBD_{spec} it is directed more toward the center of the network than in the crystal structure. This small structural change is probably sufficient to weaken the network in DBD_{unspec}, thereby causing the minor differences in the networks of the two DBDs.

Hydrogen-bonding network at the DBD-DBD interface

Hydrogen bonds that can be observed in both the crystal structure and in the simulation (Fig. 10) include a symmetri-

cal pair of salt bridges between Arg-479—N_{η1}/N_{η2} and Asn-481—O_{δ1}/O_{δ2} and a symmetrical pair of backbone contacts between Ala-477—O and Ile-483—NH. Both of these pairs of interactions were strong during the simulation and present during almost the entire production run (Table 7). However, the symmetrical pair of hydrogen bonds between the side chain of Asn-491 in DBD_{spec} and the carbonyl oxygen of the same residue in DBD_{unspec}, also present in the crystal structure, were much weaker in the simulation (present for 24 and 45 ps, respectively). Instead, the side chain of Asn-491 in DBD_{unspec} interacted with DBD_{spec} via a strong hydrogen bond with the carbonyl oxygen of Leu-475 and via a rather stable water bridge (Table 6) to both the amide group of Ala-477 and the sulphur of Cys-492 (Fig. 10); the water molecule that occupied this position was one of the crystal waters. The symmetry-related hydrogen bond between the side chain of Asn-491 in DBD_{spec} and Leu-475—O in DBD_{unspec}, absent in the crystal structure, was also found in the simulation and had a similar strength (Table 7). Another symmetrical pair of weaker hydrogen bonds (present during 75 and 39 ps, respectively) not observed in the crystal structure was found between the side chain of Arg-488 and the carbonyl oxygen of Cys-476. The reasons for these slightly changed interactions at the dimer interface in the simulation compared with the crystal structure are discussed in the section on structural changes in the DBDs.

TABLE 6 Bridging positions at the DBD-DNA interface and at the DBD-DBD interface that are occupied by water molecules for more than 30 ps during the simulation

Bridge	Mean residence time*
	(ps)
DBD _{spec} -DNA	
Asp-445-O _{δ2} ...w...T5-O2P	31 (4)
Cys-450-O...w...A8-O1P	22 (2)
Ser-459-O...w...G4-O2P	110 (1)
Ser-459-O _γ ...w...T5-O2P	47 (1)
Gln-471-O...w...A2-O2P	12 (4)
Asn-473-O...w...A2-O1P	141 (1)
Arg-489-O...w...T3-O1P	52 (2) [†]
Arg-489-N _{η1} ...w...T5-O2P	114 (1)
Arg-496-N _{η1} ...w...T3-O1P	81 (1) [‡]
DBD _{unspec} -DNA	
Asp-445-O _{δ2} ...w...G4-O1P	47 (3)
Cys-450-O...w...A8-O1P	93 (1)
Ser-459-O...w...T3-O2P	73 (1)
Glu-469-O _{ε2} ...w...A6-O1P	14 (6)
Lys-486-O...w...T3-O1P	17 (2)
Lys-486-N _δ ...w...G4-O1P	53 (1)
Leu-507-O...w...G7-O2P	66 (2)
DBD _{spec} -DBD _{unspec}	
Asn-473-O...w...Ile-487-O	95 (1)
Asn-473-O...w...Lys-490-N _ε	50 (1)
Ala-477-N...w...Asn-491-N _{δ2}	103 (1) [‡]
Asp-481-O...w...Ala-477-O	24 (2)
Lys-490-O...w...Lys-490-N _ε	73 (1)
Cys-492-S _γ ...w...Asn-491-N _{δ2}	92 (1) [‡]

*Calculated as the total time a certain bridge was present divided by the number of water molecules occupying this position at some time during the 150 ps production run (in parentheses).

[†]The water bridge is formed by a water molecule found in the crystal structure.

In addition to these hydrogen bonds, a few water bridges were found at the dimer interface (Fig. 10). Of the total of 25 water molecules that occupied bridging positions for more than 2 ps at the dimer interface during the production run, 50% had a residence time longer than 16 ps (Table 5). These bridges are thus as stable as those at the DBD-DNA interface, and considerably more stable than the intramolecular bridging water molecules within DBD. This is reasonable, considering that most of the intramolecular water bridges are formed on the surface of the protein where water molecules (on average) are more easily exchanged with the surrounding bulk waters than at the more buried dimer interface.

From Table 5 it is also clear that the length of the production run limits the calculated residence time of the most stable water bridges, because they are almost as long as the production run. This observation is substantiated by a recent 1 ns MD-simulation of BPTI (Brunne et al., 1993), where residence times for water molecules near a given BPTI atom vary between 10 and 500 ps. However, one should note that the coordinates from this simulation were saved only every 10 ps, so that shorter residence times could not be observed. The residence times of hydrating water molecules found in this simulation are consistent with NMR results of BPTI at 4°C (Otting et al., 1991), in which average residence times were concluded to be less than 500 ps.

TABLE 7 Hydrogen bonds at the DBD-DNA interface and at the dimeric interface that are present during more than 30 ps of the simulation

	Hydrogen bond	Presence time (ps)
DBD _{spec} -DNA		
His-451-N _{δ1} ...A8-O2P	150*	
Tyr-452-N...A8-O2P	150*	
Tyr-452-O _η ...G7-O2P	150*	
Lys-461-N _ε ...G7-N7	102*	
Arg-466-N _{η1} ...G4-O6	150*	
Arg-466-N _{η2} ...G4-N7	150*	
His-472-N _{δ1} ...A2-O2P	150*	
Asn-473-N...A2-O1P	123	
Tyr-474-O _η ...T3-O2P	150*	
Arg-489-N _{η1} ...G4-O2P	148*	
Lys-490-N _ε ...G1-O1P	141	
Arg-496-N _{η1} ...G4-O2P	50*	
DBD _{unspec} -DNA		
Tyr-452-N...A8-O2P	150*	
Tyr-452-O _η ...G7-O2P	150*	
Lys-461-N _ε ...G7-N7	110*	
Lys-465-N _ε ...A6-O1P	47	
Arg-466-N _{η1} ...A2-O2P	150*	
Arg-489-N _ε ...T3-O2P	150*	
Arg-496-N _{η2} ...T3-O2P	150*	
DBD _{spec} -DBD _{unspec}		
Leu-475-O...Asn-491-N _{δ2}	112	
Cys-476-O...Arg-488-N _{η1}	75	
Ala-477-O...Ile-483-N	147*	
Arg-479-N _{η1/η2} ...Asp-481-O _{δ1/δ2}	150*	
Asp-481-O _{δ1/δ2} ...Arg-479-N _{η1/η2}	150*	
Ile-483-N...Ala-477-O	130*	
Arg-488-N _{η1} ...Cys-476-O	39	
(Asn-491-O...Asn-491-N _{δ2})	24*)	
Asn-491-N _{δ2} ...Asn-491-O	45*	
Asn-491-N _{δ2} ...Leu-475-O	92	

*The hydrogen bond is present in the crystal structure.

Comparisons to the ER DBD-ERE-interface

The estrogen response element (ERE) differs from the GRE only in two basepairs in each half-site (Fig. 7): AT6 has been replaced by GC6 and AT5 by TA5. Furthermore, only three amino acids in the GR and ER DBDs appear to be essential for the discrimination between GRE and ERE, respectively (Danielson et al., 1989; Mader et al., 1989; Umesono and Evans, 1989; Ziliacus et al., 1991, 1992): Gly-458 (Glu-458), Ser-459 (Gly-459), and Val-462 (Ala-462), where the residues in parentheses are the amino acids of the ER DBD. The side chain of Glu-458 in ER DBD forms a specific interaction to the base C6 (N4) and water-mediated specific contacts to A5 (N6 and N7) (Schwabe et al., 1993b), and the importance of this residue for the discrimination between the response elements is rather obvious (Fig. 7, right). Val-462 in the GR DBD is in van der Waals contact with the methyl group of T5. The corresponding residue in the ERE DBD is an alanine, for which the side chain might not be able to interact with the methyl group of T5 in GRE. Moreover, T5 in the ERE is replaced by A5, which lacks the methyl group, and this probably reduces the affinity of GR DBD for the ERE. The third essential residue for the discrimination is Ser-459 in GR DBD. The role of this residue is more difficult to understand because it is not involved in any specific interaction with the bases. It has been suggested (Luisi et al.,

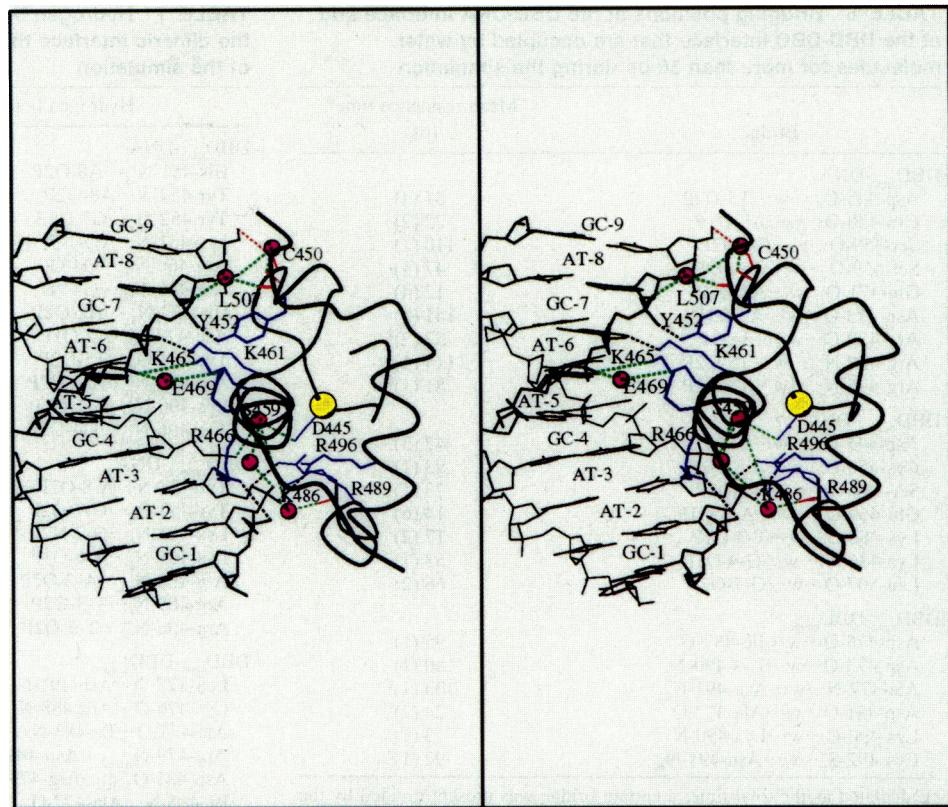


FIGURE 8 Stereo view of the DBD_{unspec}-DNA interface (crystal structure). The coloring of the lines and atoms is the same as in Fig. 6 (top).

1991) that this residue contributes indirectly to the discrimination between the response elements by stabilizing the hydrogen bond network to give the involved residues their proper orientation. This notion is supported by the observation that Ser-459—O_γ forms an integral part of the network described above for GR DBD. However, it is not obvious why these interactions are not needed for the ER DBD to stabilize a corresponding network of hydrogen bonds.

Apart from the base-specific interactions, which differ in the ER and GR DBDs, the overall features of the interactions between DBD and the phosphate groups on either side of the major groove are very similar in the simulation reported here, and in the ER DBD crystal complex (Schwabe et al., 1993b). Invariant interactions (*bold lines* in Fig. 7, right) are made by His-451 and Tyr-452 to one side of the groove and by Arg-489, Arg-496, and the carbonyl group of Ser(Gly)-459 on the other side of the groove. The latter three residues form hydrogen-bonding patterns that are remarkably similar in the two complexes, with two almost identically interacting water molecules. The rather invariant interactions to the phosphate groups thus seem to serve as a common recognition network of hydrogen bonds for measuring the width of the groove and for proper positioning of the recognition helix.

Structural changes of DBD in the simulations

DBD_{spec} versus the crystal structure

There is a sudden increase in the RMSD of DBD_{spec} after about 125 ps of simulation (Fig. 3 B), and the total potential

energy of the system also rises temporarily at this time (Fig. 3 A). The major structural differences between the average structures before (50–100 ps) and after (150–200 ps) this event are located at the last residues of the Zn_I domain, around helix II, and at the ext_{II} region (Fig. 11). The ext_{II} region has moved 4 Å toward the DNA compared with the crystal structure, and the two last residues have turned around 180° and adopt a conformation very similar to that of the ext_{II} region of DBD_{unspec} in the crystal structure. The closest interactions between protein fragments from different unit cells in the crystal are between the side chain of Glu-508 in DBD_{spec} and the side chain of Lys-510 in DBD_{unspec} in a neighboring unit cell, which are about 5 Å apart. Because of the opposite charges of the two residues, they are likely to attract each other in the crystal. It should be noted that the GR DBD fragment used for x-ray diffraction studies contained an additional 15 residues at the C-terminal compared with that simulated here. These residues were poorly ordered in both DBD_{spec} and DBD_{unspec}, and their positions are not included in the reported crystal structure (Luisi et al., 1991). Thus, it is very likely that these omitted residues affect the orientation of the last (ordered) residues in the crystal. Interestingly, the binding affinity of the ER DBD to an ERE is greatly reduced if the DBD fragment is truncated at residue 508 even though the residues beyond 508 do not form any interactions with DNA in the crystal (Schwabe et al., 1993b).

The translation and rotation of the ext_{II} region, which proceed during the whole simulation, induce tensions in helix III that eventually result in the insertion of a water molecule

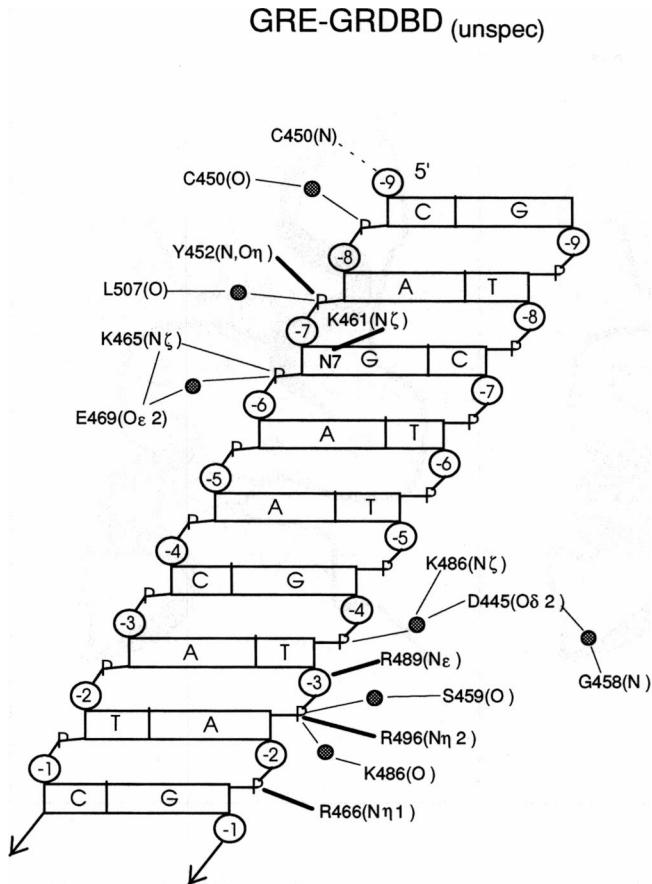


FIGURE 9 Schematic drawing of the contacts formed at the DBD_{unspec}-DNA interface of the simulated complex. Different types of lines have the same meaning as in Fig. 7 (left).

between Cys-500—O and Gly-504—NH after 125 ps. This insertion can be detected as a change in the Cys-500—O···Gly-504—NH distance from about 2.0 Å to 4.2 Å during the later part of the production run and is probably also causing the temporary rise of the potential energy (Fig. 3A), which occurs simultaneously. Before the insertion of the water molecule, which is present for the last 69 ps of the simulation, there are large fluctuations in the Cys-500—O···Gly-504—NH distance, indicating that the C-terminal part of helix III is fraying in the simulation because of the displacement of the last residues of the ext_{II} region. The fraying of the C-terminal of helix III is also reflected in weaker hydrogen bonds (Fig. 12) and larger structural differences when comparing this part of helix III with that of the crystal structure (Figs. 11 and 13). The last part of the Zn_I region (Gly-453—Cys-460) has moved concerted with the ext_{II} region because of the strong hydrogen bond between Tyr-452—O and Asn-506—NH, which was present for 140 ps in the simulation, as well as in the crystal structure, whereas the residues Cys-450—Tyr-452 are anchored by the contacts to the DNA backbone. There is also a translation of helix II toward DNA in the later part of the simulation (Fig. 11), which is caused by the approach of the side chain of Lys-490 to the phosphate group of G-1.

Structural rearrangements that occur before the insertion of the water molecule in helix III include a movement (about 1.5 Å) of the last residues of the helix I and ext_I regions toward the backbone of DNA, which result in the contacts between residues Gln-471—Asn-473 and the phosphate groups of DNA (Fig. 6, top and 7, left). The side chain of Lys-465, which is deeply involved in interactions at the ERE-ER DBD interface (Schwabe et al., 1993b), approaches the backbone of DNA during the simulation. The simultaneous opening of the major groove around the recognition helix (see above) results in a displacement of the phosphate backbone in the region around AT5 away from Lys-465, and no contact is established during the simulation. The insertion of the C-terminal part of helix I deeper into the major groove during the simulation makes this helix slightly distorted as indicated by the weaker hydrogen bonds in the middle of the helix (Fig. 12). The residues in the D-boxes (Cys-476—Cys-482) of DBD_{spec} and DBD_{unspec} move concertedly about 1.5 Å toward helix II of DBD_{spec}, because of the strong symmetrical salt bridges between the side chains of Arg-479 and Asp-481 (Table 7). Finally, the side chains of Glu-446 and Ser-448 move about 4 Å from the crystal structure and the backbone of Glu-446—Ser-448 moves away from the center of the complex (Fig. 11). The reasons for this last structural change are not clear; it cannot be due to loss of crystal contacts in the simulation because there are no closely interacting residues from other unit cells in the crystal to those residues.

DBD_{unspec} versus the crystal structure

Apart from the concerted motion of the D-box regions of the two DBDs, described above, there is a small translation (about 1.3 Å) of the ext_{II} region leading to the formation of the water-mediated contact between the carbonyl group of Leu-507 and the DNA backbone (Figs. 8 and 9). In contrast to DBD_{spec}, this small translation does not weaken helix III and its hydrogens bonds remain similar to those in the crystal structure (Fig. 12). The central part of helix I (Fig. 12) becomes distorted compared with the crystal structure because of the approach of the side chains of Lys-465 and Glu-469 to the phosphate group (Figs. 8 and 9). The approach of these residues toward the DNA backbone is facilitated by a movement of the DNA backbone toward the recognition helix and also results in a displacement of the C-terminal part of this helix away from the spacer region (i.e., downwards in Fig. 4A). The translation of helix I causes the concerted translation of the D-box regions of the DBDs, because the ext_I and D-box regions follow the translation of helix I in DBD_{unspec}. The concerted translation of the D-box regions is also related to the approach of the ext_I region of DBD_{spec} toward DNA, so that the D-box is pushed toward helix II in DBD_{spec}, in conjunction with a motion of the DBD_{unspec} D-box away from its helix II. Lys-490 of DBD_{unspec} is also approaching the backbone of DNA (as in DBD_{spec}) with about 3 Å, which induces distortions of helix II in DBD_{unspec} and its hydrogen bonds disappear (or become very weak, see Fig. 12). The

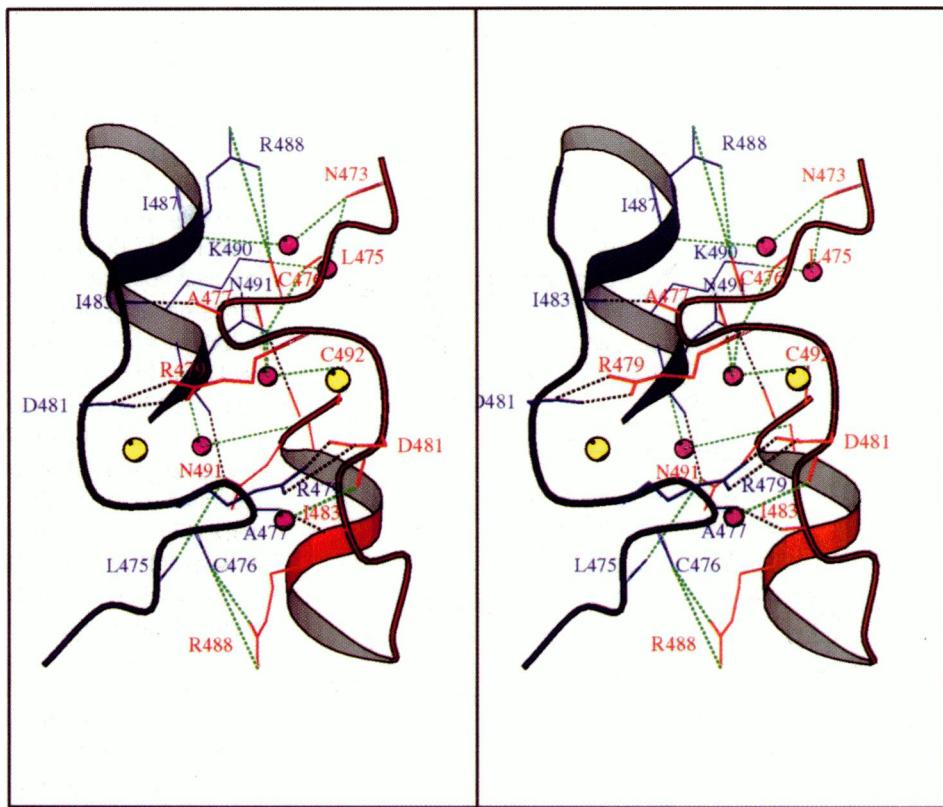


FIGURE 10 Stereo view of the dimeric interface between the two DBD-fragments (crystal structure). The coloring of the lines and atoms is the same as in Fig. 6 (top). The red ribbon represents DBD_{spec} , and the blue ribbon represents DBD_{unspec} .

distortions of the residues in the regions around helix II give rise to the slightly changed interactions at the dimer interface, described above.

DBD_{mono} and DBD_{sol} versus the NMR solution structure

The agreement between the two simulated structures of monomeric GR DBD and the NMR structure (Baumann et al., 1993) is quite good (Fig. 14) for the residues in the Zn_I region, except for the region around Glu-446. The side chain of Glu-446 rotates $\sim 60^\circ$ (DBD_{sol}) and $\sim 90^\circ$ (DBD_{mono}) away from the bulk solvent toward the N-terminal of helix I compared with the NMR structure. This change leads to distortions of the backbone in the vicinity of Glu-446. The C-terminal residues in helix I of DBD_{sol} and DBD_{mono} are translated toward the ext_{II} region with about 3 Å compared with the NMR structure (Fig. 14), resulting in more stable and regular hydrogen bonds for the last turn of helix I (Fig. 12). The regions that show the largest deviations both from the NMR structure and between the two simulated monomers are the residues in the ext_I region and the D-box region (Figs. 13 and 14). This is not surprising, because pairwise RMSDs between the 24 refined NMR structures are quite high (2–4 Å, Baumann et al., 1993) for residues Gly-470–Asn-480, i.e., the conformation of these residues is less well defined. The D-box region of DBD_{sol} moves with about 4 Å toward the bulk solvent, and the angle between the average planes formed by the backbone of these residues in DBD_{sol} and in the crystal structure is slightly smaller than the 90° found (Baumann et al., 1993) when comparing the average NMR

structure with the crystal structure. In DBD_{mono} , however, the plane formed by the backbone of the D-box is almost parallel to that of the crystal structure. Because both DBD_{sol} and DBD_{mono} have relatively low fluctuations of the backbone for these residues, it seems plausible that there is an energy barrier for the 90° reorientation of the average plane formed by the backbone atoms in the D-box when going from the crystal complex to the solution structure. This energy barrier is apparently too high to be overcome for DBD_{mono} during the 250 ps simulation.

Both simulated monomers adopt a more extended conformation in helix II than is observed in the NMR structure. The hydrogen bond between Arg-488—O and Asn-491—NH occurs in 15 of the 24 resulting NMR structures, whereas the hydrogen bond between Ile-487—O and Lys-490—NH occurs in only 7 structures (Fig. 12). The former hydrogen bond (Arg-488—O \cdots Asn-491—NH) is rather weak during the simulation, and the latter hydrogen bond has disappeared in the simulation of DBD_{mono} and is replaced by a strong (130 ps of occurrence during the production run) water bridge. The relatively weak hydrogen bonds in helix II and the water bridge found in the simulation could indicate that this helix is more labile than the two other helices in the monomeric form and that it “breathes” between a partially unfolded helix and a distorted helix. MD simulations (DiCapua et al., 1991; De Loof et al., 1992) have shown that the insertion of water molecules or bridges between the hydrogen bonds in an α -helix might initiate a destabilization of the helix. Helix II in DBD_{sol} is even more open, and

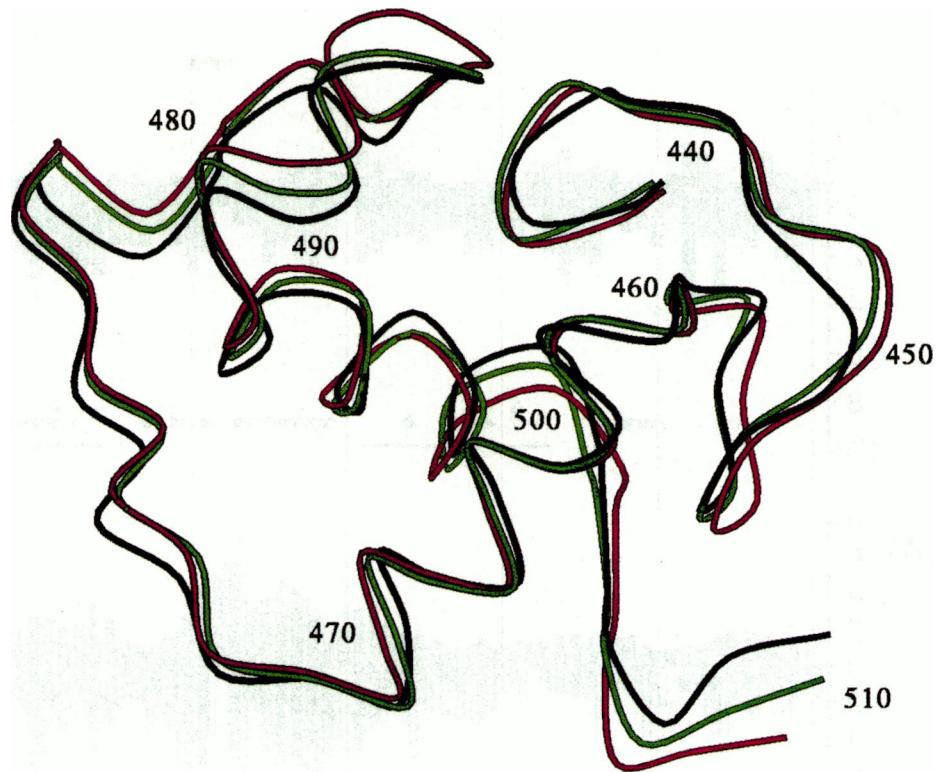
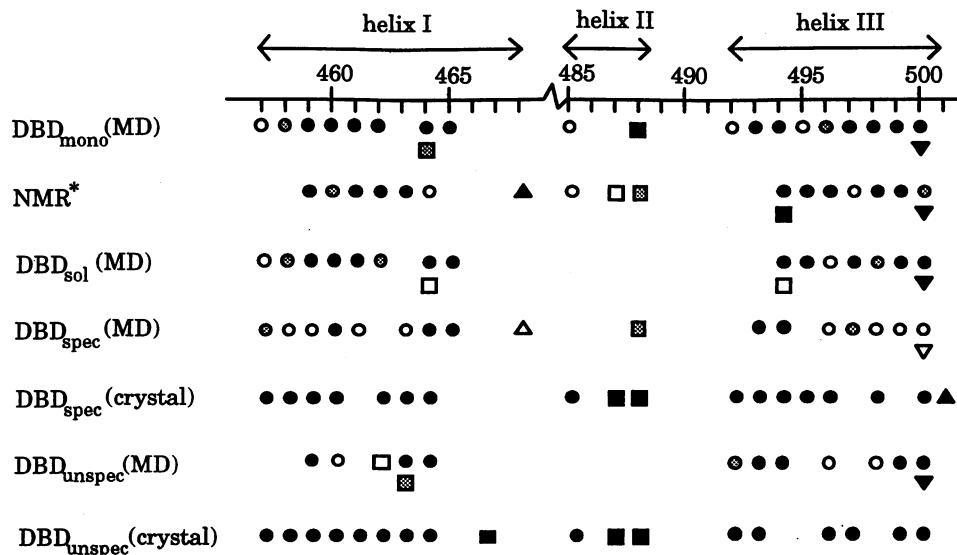


FIGURE 11 Average structures of DBD_{spec} during 50–100 ps (green) and 150–200 ps (red) simulated time, respectively, superimposed on the crystal structure of DBD_{spec} (black) (Luisi et al., 1991).

the distances Ile-487—O···Lys-490—NH and Arg-488—O···Asn-491—NH are around 4.6 and 5.9 Å, respectively. These two hydrogen bonds were initially weak (2.7 and 2.3 Å, respectively) because we used the average NMR structure as the starting structure of the DBD_{sol} simulation. The unfolding of helix II in DBD_{sol} compared with DBD_{mono}, therefore, is not surprising, because the corresponding initial distances of DBD_{mono} (i.e., the crystal structure) were 2.6 and 2.0 Å, respectively. The opening of helix II results in a shift

of about 2 Å of the N-terminal residues of helix III toward the core in both simulations of the monomers (Fig. 14). The hydrogen bond between Arg-496—O and Cys-500—NH in helix III is weakened in all simulations (Fig. 12), because of the formation of a rather strong hydrogen bond between Arg-496 and Cys-500—S, found in all simulated forms of DBD but neither in the crystal structure nor in the NMR structure. The hydrogen bond between Arg-496—O and Cys-500—S, occurred during about 50% of the production runs and could

FIGURE 12 Stability and types of hydrogen bonds in the three helical regions determined from the simulation, the crystal structure (Luisi et al., 1991), and from the 24 refined NMR solution structures (Baumann et al., 1993). The symbols indicate (i,i+4) hydrogen bonds (○), (i,i+3) hydrogen bonds (□), (i,i+2)-hydrogen bonds (Δ), and (i,i+5) hydrogen bonds (▽). The symbol is open if the hydrogen bond was found in 20–50% of the structures, shaded if 50–75% of the structures contained the hydrogen bond, and black if the hydrogen bond was present in more than 75% of the structures.



* From the 24 refined NMR solution structures (Baumann et al., 1993)

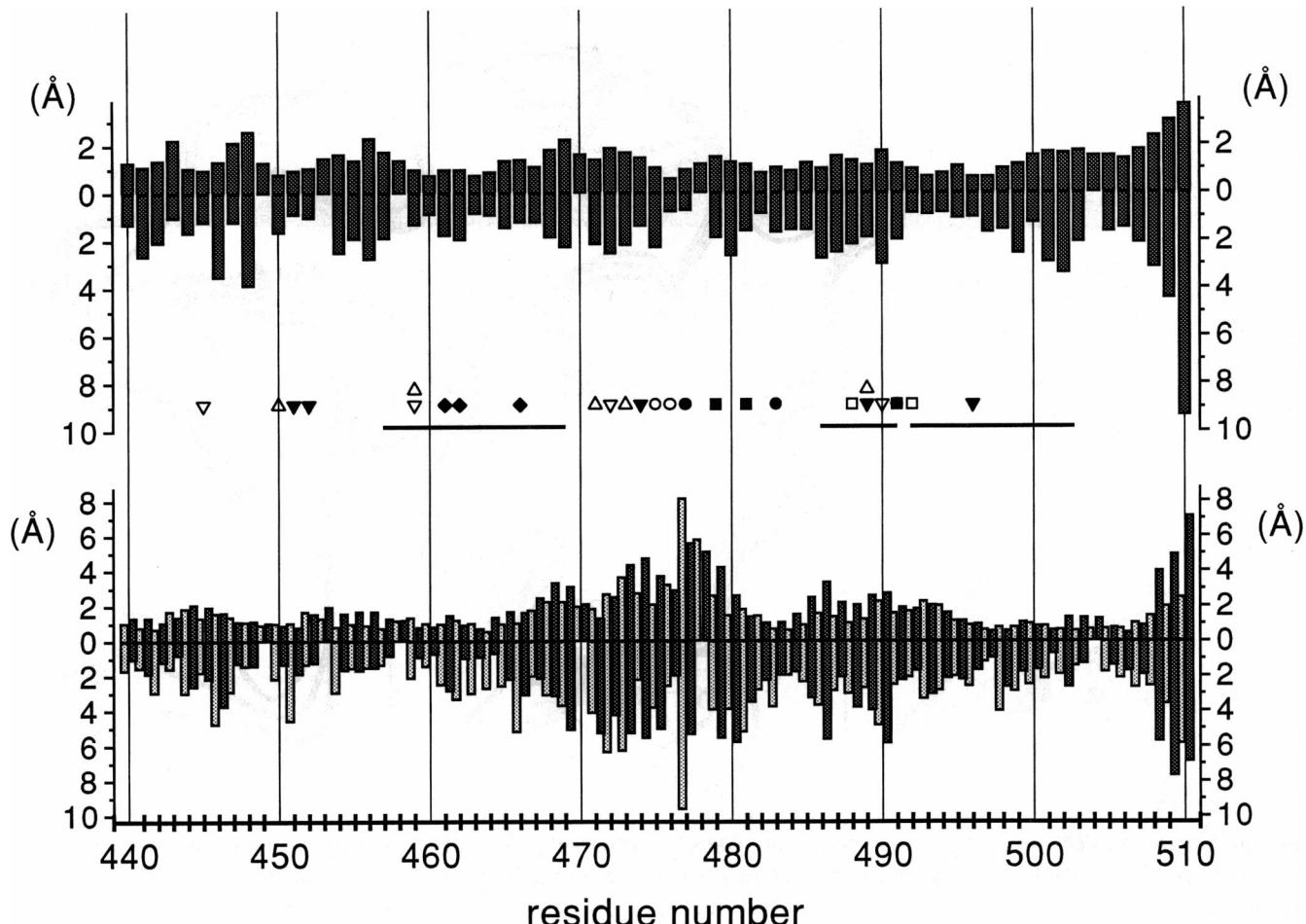


FIGURE 13 (top) Residue-averaged root mean square deviations (RMSD) of DBD_{spec} compared with the crystal structure for the backbone (bars above the abscissa) and side chains (bars below the abscissa). (bottom) RMSD of DBD_{sol} (darker bars) and of DBD_{mono} (lighter bars) compared with the average NMR solution structure, with the same direction of the bars as above. Intermolecular interactions are denoted as follows: (○) interactions between DBD_{spec} and DBD_{unspec} involving backbone atoms, (□) interactions between DBD_{spec} and DBD_{unspec} involving side-chain atoms, (Δ) backbone atoms of DBD_{spec} that interact with DNA phosphates or sugar oxygens, (▽) side-chain atoms of DBD_{spec} that interact with DNA phosphates or sugar oxygens, (◇) interactions between side chains of DBD_{spec} and DNA bases. Filled symbols indicate that the interactions are present in the crystal structure, whereas open symbols represent contacts that are found in the simulation only. The three thick lines represent the location of the helices of DBD.

be established with minor structural changes of the protein. Finally, the flexible residues of the ext_{II} region have, not surprisingly, different average structures in the NMR structure, in DBD_{sol}, and in DBD_{mono}.

Internal mobility in DNA-complexed and free DBDs

We now compare and discuss intramolecular fluctuations and accessible surface areas in the specifically bound DBD of the complex (DBD_{spec}) and monomeric DBD in solution (DBD_{mono}). These properties are very similar in DBD_{mono} and DBD_{sol}, and the latter simulation, therefore, is not treated explicitly. The only significant difference between the two monomer simulations is that the residues in the D-box region are slightly more ($\sim 0.5 \text{ \AA}^2$) flexible in DBD_{sol} than in DBD_{mono}.

In general, the mobility is limited and uniform throughout the DBD backbone in both DBD_{spec} and DBD_{mono}, with an

increased flexibility at the C-terminal of the fragments. This uniform mobility was also observed in ${}^1\text{H}$ - ${}^{15}\text{N}$ NMR relaxation studies (Berglund et al., 1992; Eriksson et al., 1993) of DBD in solution and in a previous investigation of the flexibility of the amide bonds in the simulated DBD_{mono} (Eriksson et al., 1993). A comparison of the fluctuations of the backbone of DBD_{spec} and DBD_{mono} with their fractional exposure (% a.s.a.) to the solvent (Fig. 15) shows that the solvent exposure correlates with the mobility. A “spline” of the surface exposure from one residue to another gives a very good correlation with the mobility of the backbone for DBD_{mono} and DBD_{spec} because the motions of the backbone atoms are highly correlated to backbone motions of sequential neighbors. The fluctuations of helix I are slightly lower in DBD_{spec} than in monomeric DBD_{mono}, most likely due to the stabilizing effect of DNA on helix I, because this region of DBD_{spec} is positioned in the DNA major groove. Helix I shows tendencies of fraying at the N-terminal (Fig. 12) in both monomeric simulations, possibly as a consequence of

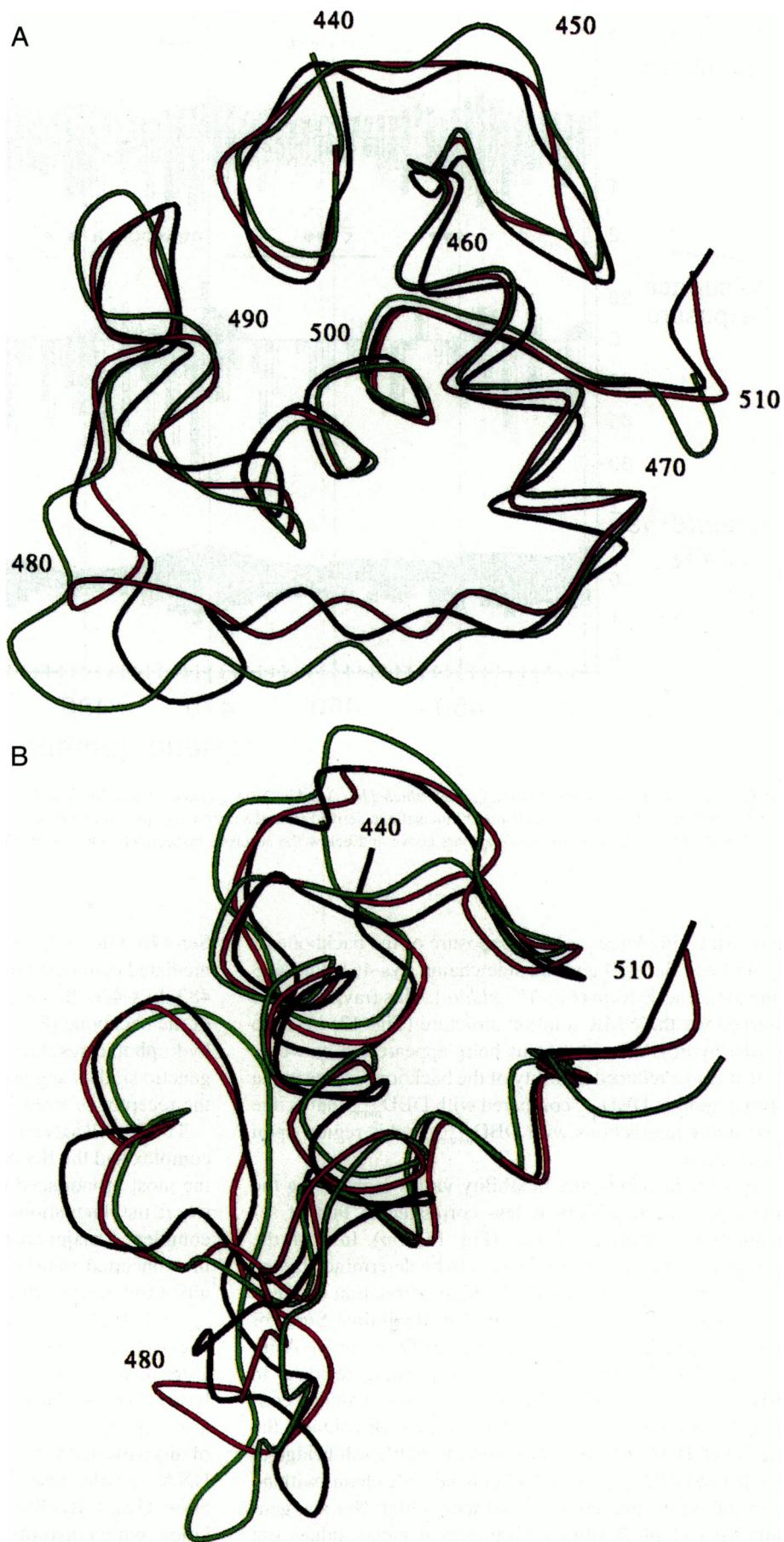


FIGURE 14 (A) Average structures of DBD_{sol} (green) and DBD_{mono} (red) superimposed on the average NMR solution structure (black) (Baumann et al., 1993). (B) A rotated view highlighting different conformations of the D-box.

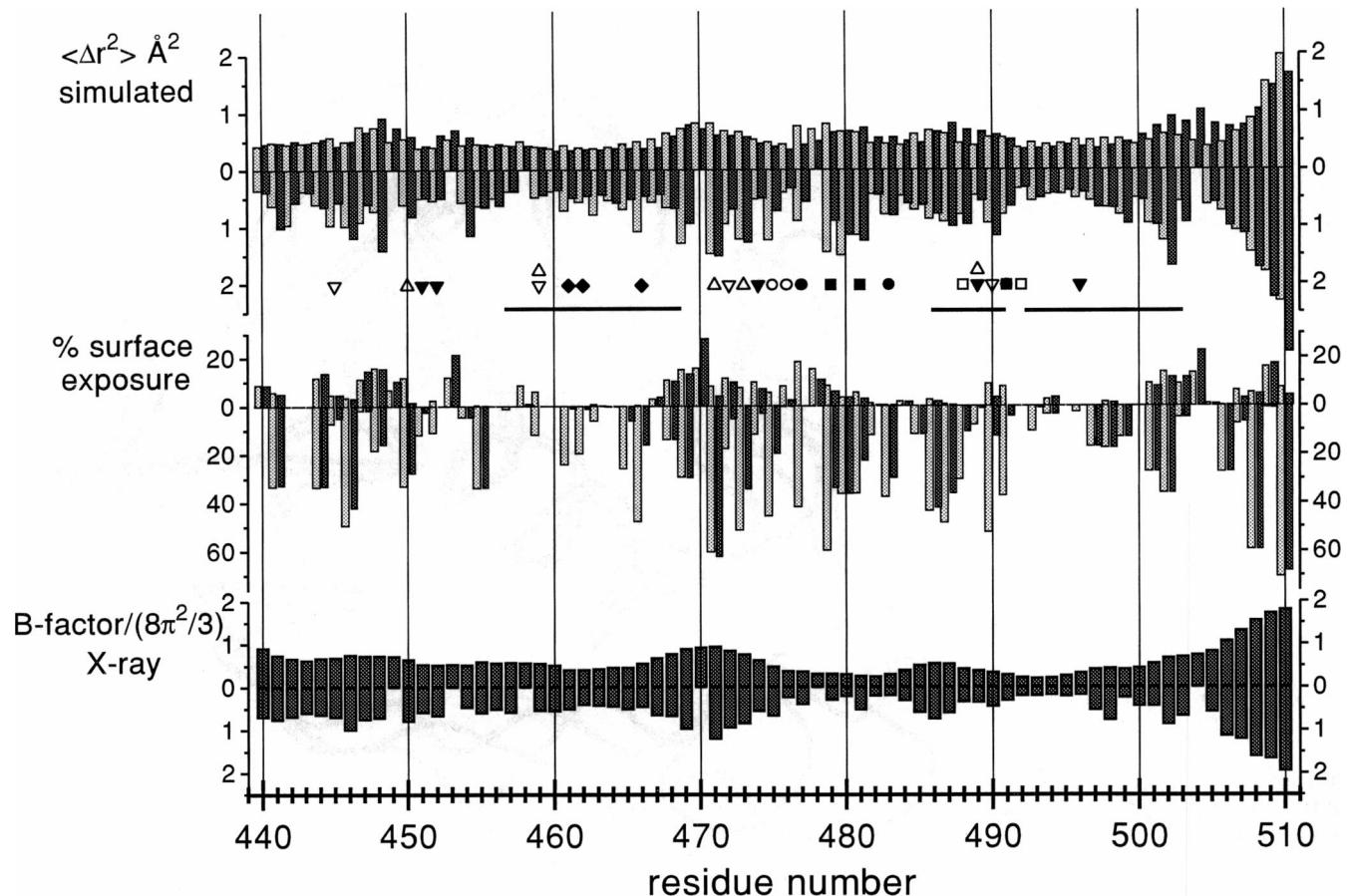


FIGURE 15 (top) Fluctuations in DBD_{mono} (light shaded bars) and in DBD_{spec} (darker bars). (middle) Residual surface exposure in DBD_{mono} (light shaded bars) and in DBD_{spec} (darker bars). (bottom) Fluctuations derived from the crystallographic (Luisi et al., 1991) B-factors ($3B/8\pi^2$) of DBD_{spec}. Values for the backbone and side chains are shown as bars above and below the abscissa, respectively. Other symbols have the same meanings as in Fig. 13.

the considerably larger solvent exposure of the backbone of Gly-458 and Ser-459 and the side chains Lys-461–Phe-463 in the monomeric form (Fig. 15, middle). This fraying is also observed for the NMR solution structure (Fig. 12) because the first hydrogen bonds of this helix appeared to be weak. There is also a reduced mobility of the backbone atoms in the D-box region in DBD_{spec} compared with DBD_{mono} that is due to the many interactions with DBD_{unspec} in this region upon dimerization.

For the side chains, the flexibility varies more along the protein and the flexibility is less correlated to that of sequentially neighboring residues (Fig. 15, top). Instead, the flexibility of the side chains appears to be determined by the solvent exposure as is seen by the high correlation between the % a.s.a. for the side chains and their flexibility. Some of the side chains that interact with either DBD_{unspec} or DNA are of course less flexible than the corresponding residues in DBD_{mono}; e.g., Asp-445, which forms a water-mediated contact to the backbone of DNA, Arg-466, which contacts the base G4 of DNA, and Arg-479 with the stable salt bridge to Asp-481 of DBD_{unspec}. On the other hand, side chains with no direct intermolecular contacts and with a high % a.s.a. generally have a high flexibility. Examples of such residues are:

Ser-448, Glu-469, Gln-471 (Gln-471 forms a weak water-mediated contact with the backbone of DNA), Asn-480, Ile-483, Lys-486, Ile-487, Gln-502, and the C-terminal residues of the backbone (Fig. 15). The high surface exposure of the hydrophobic residues Ile-483 and Ile-487 is unusual, and genetic studies suggest that these contact other domains of the receptor or even other proteins (Schena et al., 1989).

The overall agreement between the B-factors of the crystal complex and the fluctuations of DBD_{spec} is rather good, with the most pronounced differences around the D-box region, where the fluctuations are larger in the simulation than in the complex. A major contribution to this difference comes from the concerted translations of the residues in the D-box regions and the distortions of residues around helix II in DBD_{spec} and DBD_{unspec} in the simulation. Another difference between B-factors and the simulated fluctuations is found at the C-terminal of helix III, where the fluctuations in DBD_{spec} are larger than the fluctuations deduced from the B-factors, and also larger than in free DBD. This mobility reflects fraying of this part caused by the translation of the ext_{II} region toward DNA and the insertion of a water bridge between Cys-500—O and Gly-504—NH during the later part of the simulation, which disrupts the helix and makes it more flexible.

Concerted intra- and intermolecular motions of the DBD backbone

Cross-correlation coefficients for the backbone atoms in DBD_{mono} are mapped in Fig. 16 (*top*). Residues in the Zn_I region seem to fluctuate in a very concerted manner because they are almost all significantly correlated with one another. This is quite reasonable considering that most of the residues in this region are connected by backbone hydrogen bonds. The motions of the residues in helix I and III are correlated with sequentially nearby residues, and the correlation reaches about seven residues away in both the N- and C-terminal directions. This is, of course, a reflection of the stabilizing effect of the α -helical hydrogen bonds, which give rise to the concerted motions. Motions of residues in the extended regions (ext_I and ext_{II}) and the fragment Ile-484-Lys-486 all appear to be uncorrelated to those of nearby regions, indicating less rigid structures. This motion is in good agreement with observed backbone fluctuations (Fig. 15), which are somewhat higher for these three fragments. Parts of the protein that are spatially close (see Fig. 2 *B*) often move in a correlated manner, for example, the first five or six residues of the Zn_I region correlate with most of the residues of helix III. The region centred around Tyr-452 moves concerted with the C-terminal residues of helix III and the beginning of the ext_{II} region, reflecting the strong hydrogen bond between Tyr-452—O and Asn-506—NH as discussed above.

The backbone motions in the two monomers in the (DBD)₂-DNA simulation (Fig. 16, *bottom*) are, not surprisingly, most concerted in the D-box regions (Cys-476–Cys-482), where many strong interactions are formed between the monomers. Another region of concerted motion between residues of the two monomers is found at the end of helix II, where Lys-490 and Asn-491 in the monomers are interacting (see above). Residues from the D-box of DBD_{spec} are also moving concertedly with the first part of helix III of DBD_{unspec}, and vice versa. This correlation appears to be caused by intramolecular concerted motions (Fig. 16, *top*) of the residues from the D-box and the N-terminal of helix III. One can also observe an almost “symmetrical” concerted motion between Cys-443–Asp-445 in one monomer and at residues around Cys-476 in the other monomer. These residues are not spatially close, but the strong intramolecular salt bridges between the side chains of Asp-445 and Arg-489 found in both monomers of the complex and the symmetrical intermolecular interactions between the Arg-488 side chain and Cys-476—O (see above) cause the concerted motion.

SUMMARY AND CONCLUSIONS

Overall, the integrity of the initial structure of the (GR DBD)₂-GRE_{S4} complex is well maintained during the simulation, albeit with some local rearrangements both in the DNA and protein units, primarily leading to more pronounced interactions between protein and DNA. A number of stable hydrogen bonds, direct as well as water-mediated, form during the simulation. Many of the interfacial water

molecules are immobilized, with calculated residence times limited by the length of the simulation.

The major structural changes that occur for DNA during the simulation of the complex are as follows: 1) a bend of the helix axis of approximately 10° in the spacer region between the half-sites that appear to be caused mainly by the new interactions between the ext_I region of DBD_{spec} and the backbone of DNA; 2) parts of the major groove that accommodate the recognition helices expand slightly in conjunction with a narrowing of the minor groove in the spacer region and at the end of the DNA-helix where DBD_{spec} binds.

The most stable intramolecular water bridges in DNA are found in the minor groove, and the majority of these bridges are located between a purine N3/pyrimidine O2 and the sugar oxygen O4' of the next nucleotide. These water molecules in the minor groove are about twice as stable as other DNA-bound waters. The sodium counterions are preferentially located either close to the phosphate groups of the minor groove or inside the minor groove. In the latter case, the sodium ions are coordinating purine N3/pyrimidine O2 through water molecules from the first hydration shell of the ion.

Many new, stable hydrogen bonds and water bridges are established between DBD_{spec} and the backbone of DNA during the simulation. Most of the interactions that develop at the DBD_{spec}-DNA interface involve residues in the ext_I region, but also the residues Asn-445, Arg-489, and Arg-496. The three latter residues are involved, in DBD_{spec} as well as in DBD_{unspec}, in stable hydrogen-bonding networks that connect many parts of the DBD fragments. Comparing this hydrogen-bonding network at the specific site with the corresponding network of the similar estrogen receptor DNA complex (Schwabe et al., 1993b) reveals striking similarities between our simulation of the (GR DBD)₂-GRE_{S4} and the crystal structure of the (ER DBD)₂-ERE complex. The interactions that form between DBD and DNA on the other side of the major groove are also similar in the two complexes, and the vast majority of residues that are involved in these hydrogen bonds are highly conserved within the family of nuclear hormone receptors (Laudet et al., 1992). These observations suggest that these hydrogen-bonding networks are a common feature in ER and GR (and possibly in other members of this family), which is used to ensure proper positioning of the recognition helix.

The dimer interface is slightly distorted compared with the starting structure in the regions around helix II of DBD_{spec} and DBD_{unspec}, because of movements of the Lys-490 side chain toward the backbone of DNA. This results in two symmetry-related pairs of interactions between the side chains of Asn-491 and Leu-475—O and a weaker interacting pair between the side chains of Arg-488 and Cys-476—O. The D-box regions (Cys-476–Cys-482) of DBD_{spec} and DBD_{unspec}, which are interacting strongly through a symmetrical pair of salt bridges, move concertedly about 1.5 Å toward helix II of DBD_{spec}. This movement is related to correlated movements of the ext_I region DBD_{spec} and the C-terminal part of helix I in DBD_{unspec}.

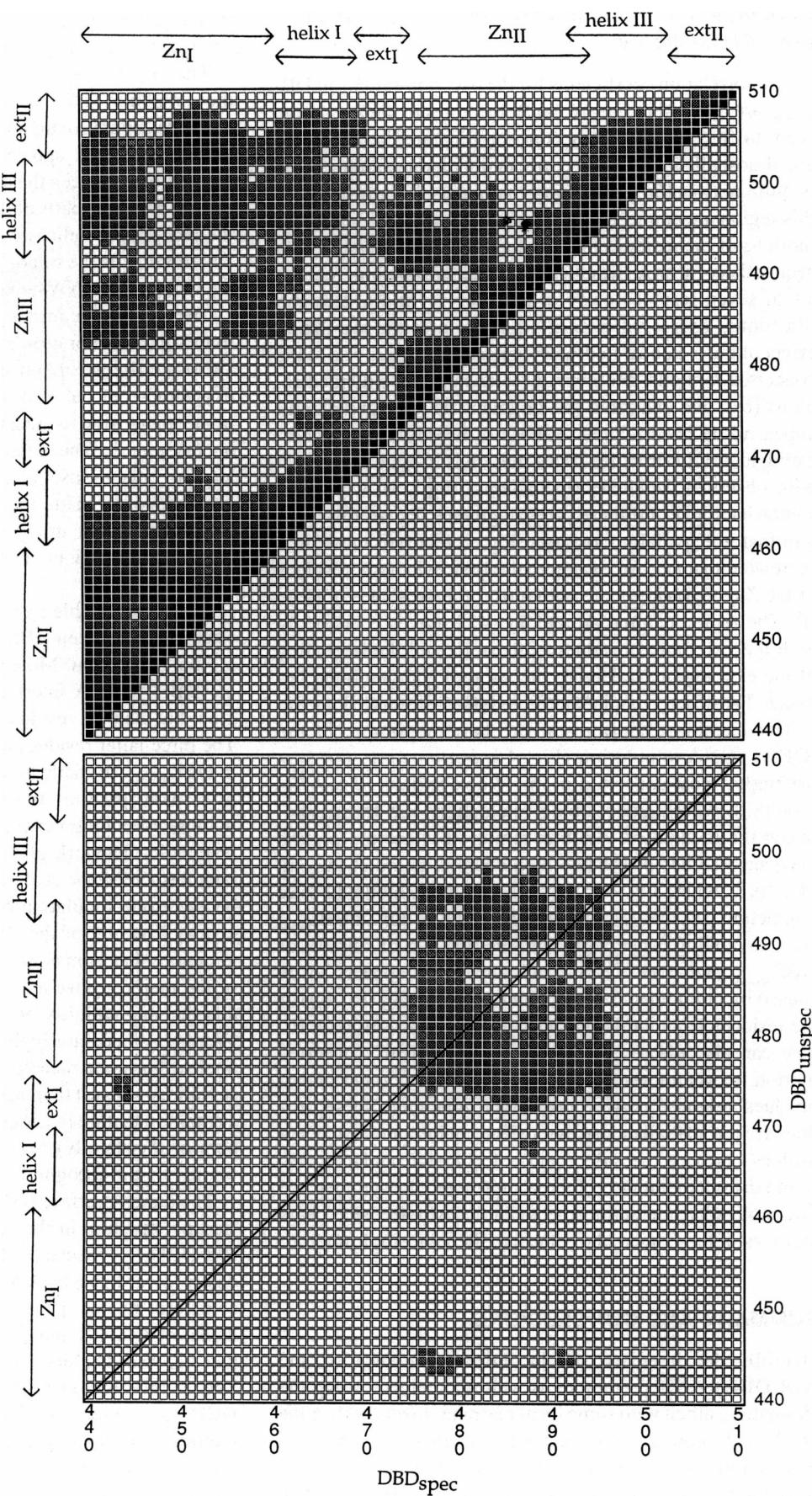


FIGURE 16 Cross-correlation map for the displacements of the backbone atoms in DBD, averaged over each residue. Cross-correlation coefficients (C_{ij}) are shaded if they are >0.4 and become gradually darker to $C_{ij} = 1$, for which they are black. No significant negative cross-correlation coefficients are observed. (top) Intramolecular C_{ij} values from the DBD_{mono}-simulation. (bottom) Intermolecular C_{ij} values between DBD_{spec} and DBD_{unspec} in the simulation of the complex.

The recognition helices of DBD_{spec} and DBD_{unspec} move deeper into the DNA major groove, resulting in distortions and in a weakening of the helical hydrogen-bonding patterns in the middle of the helices. Helix III of DBD_{spec} is also weakened in the simulation, because of the approach of the last residues of the ext_{II} region to the backbone of DNA for closer interactions. The movement of the ext_{II} region results in an insertion of a water molecule at the C-terminal of helix III.

A comparison of the average NMR solution structure (Baumann et al., 1993) with the two simulations of the monomeric form of DBD (DBD_{mono} and DBD_{sol}) shows structural differences in the ext_I and in the D-box regions, and the differences correlate with pairwise RMSD values between the 24 refined NMR structures. The conformation of the D-box region differs in the complex and in the free monomer (Baumann et al., 1993). In the two monomer simulations, the conformation of this region remains close to that of the starting structures, suggesting that there is an energy barrier to the conformational change of the D-box.

We are grateful to Prof. Paul Sigler for providing us with the coordinates for the crystal structure of the GRE_{S4}-(DBD)₂ complex.

This work has been supported by the Swedish Natural Science Research Council.

REFERENCES

- Aggarwal, A. K., D. W. Rodgers, M. Drottar, M. Ptashne, and S. C. Harrison. 1988. Recognition of a DNA operator by the repressor of phage 434: a view at high resolution. *Science*. 242:899–907.
- Atkins, P. W. 1990. Physical Chemistry, 4th ed. W. H. Freeman & Co., New York.
- Bartenev, V. N., Eu. I. Golovamov, M. A. Kapitonova, M. A. Mokulskii, L. I. Volkova, and I. Ya. Skuratovskii. 1983. Structure of the B-DNA cationic shell as revealed by an x-ray diffraction study of CsDNA. *J. Mol. Biol.* 169:217–234.
- Baumann, H., K. Paulsen, H. Kovács, H. Berglund, A. P. H. Wright, J.-Å. Gustafsson, and T. Härd. 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.
- Berglund, H., H. Kovács, K. Dahlman-Wright, J.-Å. Gustafsson, and T. Härd. 1992. Backbone dynamics of the glucocorticoid receptor DNA-binding domain. *Biochemistry*. 31:12001–12011.
- Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, E. F. Mayer, J. M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. 1977. The protein data bank. A computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535–542.
- Besler, B. H., K. M. Merz, Jr., and P. A. Kollman. 1990. Atomic charges derived from semiempirical methods. *J. Comp. Chem.* 11: 431–439.
- Beveridge, D., and G. Ravishanker. 1994. Molecular dynamics studies of DNA. *Curr. Opin. Struct. Biol.* 4:246–255.
- Bishop, T., and K. Schulten. 1994. Molecular dynamics study of a sequence specific protein-DNA interaction. In Computational Approaches in Supramolecular Chemistry. G. Wipff, editor. NATO ASI Series Vol. 426. Kluwer, Dordrecht. 419–439.
- Brooks, B. R., R. E. Brucolieri, B. D. Olafsen, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.* 4:187–217.
- Brooks III, C. L., and M. Karplus. 1983. Deformable stochastic boundaries in molecular dynamics. *J. Chem. Phys.* 79:6312–6325.
- Brooks III, C. L., B. M. Pettitt, and M. Karplus. 1988. Proteins: a theoretical perspective of dynamics, structure and thermodynamics. *Adv. In. Chem. Phys.* Vol. 71, Wiley, New York.
- Brunne, R. M., E. Liepinsh, G. Otting, K. Wüthrich, and W. F. van Gunsteren. 1993. Hydration of proteins. A comparison of experimental residence times of water molecules solvating the bovine pancreatic trypsin inhibitor with theoretical model calculations. *J. Mol. Biol.* 231: 1040–1048.
- Brünger, A., and M. Karplus. 1988. Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. *Proteins Struct. Funct. Genet.* 4:148–156.
- Clark, K. L., E. D. Halay, E. Lai, and S. K. Burley. 1993. Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature*. 364:412–420.
- Clore, G. M., A. Bax, J. G. Omichinski, and A. M. Gronenborn. 1994. Localization of bound water in the solution structure of a complex of the erythroid transcription factor GATA-1 with DNA. *Structure*. 2:89–94.
- Clementi, E., and G. Corongiu. 1983. Structure of aggregates of water and Li⁺, Na⁺ and K⁺ counterions with nucleic acid in solution. *J. Biol. Phys.* 11:33–42.
- Dahlman-Wright, K., A. Wright, J.-Å. Gustafsson, and J. Carlstedt-Duke. 1991. Interaction of the glucocorticoid receptor DNA-binding domain with DNA as a dimer is mediated by a short segment of five amino acids. *J. Biol. Chem.* 266:3107–3112.
- Danielsen, M., L. Hinck, and G. M. Ringold. 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.
- De Loof, H., L. Nilsson, and R. Rigler. 1992. Molecular dynamics simulation of galanin in an aqueous and nonaqueous solution. *J. Am. Chem. Soc.* 114:4028–4035.
- de Vlieg, J., H. J. C. Berendsen, and W. F. van Gunsteren. 1989. An NMR-based molecular dynamics simulation of the interaction of the lac repressor headpiece and its operator in aqueous solution. *Proteins Struct. Funct. Genet.* 6:104–127.
- Dewar, M. J. S., and K. M. Merz, Jr. 1986. MNDO Calculations for compounds containing zinc. *Organometallics*. 5:1494–1496.
- Dewar, M. J. S., and W. J. Thiel. 1977a. Ground states of molecules. 38. The MNDO method. Approximations and parameters. *J. Am. Chem. Soc.* 99:4899–4906.
- Dewar, M. J. S., and W. J. Thiel. 1977b. Ground states of molecules. 39. MNDO Results for molecules containing hydrogen, carbon, nitrogen and oxygen. *J. Am. Chem. Soc.* 99:4907–4917.
- DiCapua, F. M., S. Swaminathan, and D. L. Beveridge. 1991. Theoretical evidence for water insertion in α -helix bending: molecular dynamics of Gly₃₀ and Ala₃₀ in vacuo and in solution. *J. Am. Chem. Soc.* 113: 6145–6155.
- Drew, H. R., and R. E. Dickerson. 1981. Structure of a B-DNA Dodecamer. III. Geometry of hydration. *J. Mol. Biol.* 151:535–556.
- Eriksson, M. A. L., H. Berglund, T. Härd, and L. Nilsson. 1993. A comparison of ¹⁵N 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, M., T. Härd, and L. Nilsson. 1994. Molecular dynamics simulation of a DNA binding protein free and in complex with DNA. In Computational Approaches in Supramolecular Chemistry. G. Wipff, editor. NATO ASI Series Vol. 426. Kluwer, Dordrecht. 441–456.
- Eriksson, M. A. L., and A. Laaksonen. 1992. A molecular dynamics study of conformational changes and hydration of left-handed d(CGCGCGCGCG)₂ in a nonsalt solution. *Biopolymers*. 32: 1035–1059.
- Hancock, R. D. 1989. Molecular mechanics calculations as a tool in coordination chemistry. *Prog. Inorg. Chem.* 37:187–291.
- Härd, T., and J.-Å. Gustafsson. 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., E. Kellenbach, R. Boelens, R. Kaptein, K. Dahlman, J. Carlstedt-Duke, L. P. Freedman, B. A. Maler, E. I. Hyde, J.-Å. Gustafsson, and K. R. Yamamoto. 1990a. ¹H NMR studies of the glucocorticoid receptor DNA-binding domain: sequential assignments and identification of secondary structure elements. *Biochemistry*. 29:9015–9023.

- Härd, T., E. Kellenbach, R. Boelens, B. A. Maler, K. Dahlman, L. P. Freedman, J. Carlstedt-Duke, K. R. Yamamoto, J.-Å. Gustafsson, and R. Kaptein. 1990b. ^1H NMR studies of the glucocorticoid receptor DNA-binding domain: sequential assignments and identification of secondary structure elements. *Science*. 249:157–160.
- Hoogsteen, K. 1963. The crystal and molecular structure of a hydrogen-bonded complex between 1-methyl thymine and 9-methyl adenine. *Acta Cryst.* 16:907–916.
- Impey, R. W., P. A. Madden, and I. R. McDonald. 1983. Hydration and mobility of ions in solution. *J. Phys. Chem.* 87:5071–5083.
- Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
- Kopka, M. L., A. V. Fratini, H. R. Drew, and R. E. Dickerson. 1983. Ordered water structure around a B-DNA dodecamer. A quantitative study. *J. Mol. Biol.* 163:129–146.
- Kraulis, P. J. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24:946–950.
- Laudet, V., C. Hänni, J. Coll, F. Catzeffis, and D. Stéhelin. 1992. Evolution of the nuclear receptor gene superfamily. *EMBO J.* 11:1003–1013.
- Lavery, R., and H. Sklenar. 1988. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* 6:63–91.
- Laaksonen, A., and L. G. Nilsson. 1989. Molecular dynamics simulation of double helix Z-DNA in solution. *Chem. Phys.* 129:175–183.
- Lee, B., and F. M. Richards. 1971. The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* 55:379–400.
- Levitt, M., and B. H. Park. 1993. Water: now you see it, now you don't. *Structure*. 1:223–226.
- Luisi, B. F., W. X. Xu, Z. Otwinowski, L. P. Freedman, K. R. Yamamoto, and P. B. Sigler. 1991. Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature*. 352:497–505.
- Mader, S., V. Kumar, H. de Verneuil, and P. Chambon. 1989. Three amino acids of the oestrogen receptor are essential to its ability to distinguish an oestrogen from a glucocorticoid response element. *Nature*. 338:271–274.
- McCammon, A. J., and S. C. Harvey. 1986. Dynamics of Proteins and Nucleic Acids. Cambridge University press, Cambridge U.K. 87–91.
- Merz, K. M., Jr. 1991. CO₂ binding to human carbonic anhydrase II. *J. Am. Chem. Soc.* 113:406–411.
- Mills, R. 1973. Self-diffusion in normal and heavy water in the range 1–45°. *J. Phys. Chem.* 77:685–688.
- Misra, V. K., K. A. Sharp, R. A. Friedman, and B. Honig. 1994. Salt effects on ligand-DNA binding. Minor groove binding antibiotics. *J. Mol. Biol.* 238:245–263.
- Norberg, J., and L. Nilsson. 1994. High pressure molecular dynamics of a nucleic acid fragment. *Chem. Phys. Lett.* 224:219–224.
- Otting, G., E. Liepinsh, and K. Wüthrich. 1991. Protein hydration in aqueous solution. *Science*. 254:974–980.
- Otwinowski, Z., R. W. Schevitz, R.-G. Zhang, C. L. Lawson, A. Joachimiak, R. Q. Marmorstein, B. F. Luisi, and P. B. Sigler. 1988. Crystal structure of the trp repressor/operator complex at atomic resolution. *Nature*. 335:321–329.
- Pearson, R. G. 1986. Ionization potentials and electron affinities in aqueous solution. *J. Am. Chem. Soc.* 108:6109–6114.
- Qian, Y. Q., G. Otting, and K. Wüthrich. 1993. NMR detection of hydration water in the intermolecular interface of a protein-DNA complex. *J. Am. Chem. Soc.* 115:1189–1190.
- Ravishanker, G., S. Swaminathan, D. L. Beveridge, R. Lavery, and H. Sklenar. 1989. Conformational and helicoidal analysis of 30 ps of molecular dynamics on the d(CGCAGATTGCG) double helix. *J. Biomol. Struct. Dyn.* 6:669–699.
- Rosenberg, J. M. 1991. Structure and function of restriction endonucleases. *Curr. Opin. Struct. Biol.* 1:104–113.
- Ryckaert, J. P., G. Cicotti, and H. J. C. Berendsen. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comp. Phys.* 23:327–341.
- Saenger W. 1984. Principles of Nucleic Acid Structure. Springer-Verlag, New York.
- Schwabe, J. W. R., L. Chapman, J. T. Finch, D. Rhodes, and D. Neuhaus. 1993a. DNA recognition by the oestrogen receptor: from solution to the crystal. *Structure*. 1:187–204.
- Schwabe, J. W. R., L. Chapman, J. T. Finch, and D. Rhodes. 1993b. The crystal structure of the oestrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. *Cell*. 75:567–578.
- Schwabe, J. W. R., D. Neuhaus, and D. Rhodes. 1990. Solution structure of the DNA-binding domain of the oestrogen receptor. *Nature*. 348:458–461.
- Schenk, M., L. P. Freedman, and K. R. Yamamoto. 1989. Mutations in the glucocorticoid receptor zinc finger region that distinguish interdigitated DNA binding and transcriptional enhancement activities. *Genes Dev.* 3:1590–1601.
- Seibel G. L., U. C. Singh, and P. Kollman. 1985. A molecular dynamics simulation of double-helical B-DNA including counterions and water. *Proc. Natl. Acad. Sci. USA*. 82:6537–6540.
- Steward, J. J. P. 1990. MOPAC: a semiempirical molecular orbital program. *J. Computer Aided Mol. Des.* 4:1–45.
- Tasaki, K., S. McDonald, and J. W. Brady. 1993. Observations concerning the treatment of long-range interactions in molecular dynamics simulations. *J. Comp. Chem.* 14:278–284.
- Tsai, S. Y., J. Carlstedt-Duke, N. L. Weigel, K. Dahlman, J.-Å. Gustafsson, M.-J. Tsai, and B. W. O'Malley. 1988. Molecular interactions of steroid hormone receptor with its enhancer element: evidence for receptor dimer formation. *Cell*. 55:361–369.
- Umesono, K., and R. M. Evans. 1989. Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell*. 57:1139–1146.
- van Gunsteren, W. F., H. J. C. Berendsen, R. G. Geurtsen, and H. R. J. Zwinderan. 1986. A molecular dynamics computer simulation of an eight-base-pair DNA fragment in aqueous solution: comparison with experimental two-dimensional NMR data. *Ann. N. Y. Acad. Sci.* 482:287–303.
- Westhof, E. 1987. Hydration of oligonucleotides in crystals. *Int. J. Biol. Macromol.* 9:185–192.
- Zilliacus, J., K. Dahlman-Wright, A. Wright, J.-Å. Gustafsson, and J. Carlstedt-Duke. 1991. DNA binding specificity of mutant glucocorticoid receptor DNA-binding domains. *J. Biol. Chem.* 266:3101–3106.
- Zilliacus, J., A. P. H. Wright, U. Norinder, J.-Å. Gustafsson, and J. Carlstedt-Duke. 1992. Determinants for DNA-binding site recognition by the glucocorticoid receptor. *J. Biol. Chem.* 267:24941–24947.