MOLECULAR DYNAMICS STUDY OF A SEQUENCE SPECIFIC PROTEIN-DNA INTERACTION*

T. BISHOP and K. SCHULTEN [a]

Beckman Institute and Departments of Chemistry and Physics
University of Illinois at Urbana-Champaign
405 North Matthews, Urbana, Illinois 61801

* This work was supported by a grant from the National Institutes of Health (P41RR05969). [a] To whom correspondance should be addressed.

ABSTRACT. The crystal structure of the DNA-binding domain of the glucocorticoid receptor complexed with DNA has been made available to us by Sigler *et al* [1]. The glucocorticoid receptor binds as a dimer to DNA, and the receptor dimer recognizes a specific sequence of DNA through nucleotide interactions in the major groove. The target DNA consists of a palindrome or near palindrome of hexameric half-sites separated by a three base pair sequence of DNA. To increase the stability of the system for the following molecular dynamics study, the protein and DNA have been encapsulated in an ellipsoid of water consisting of approximatively 3,000 water molecules. Energy minimization, equilibration and dynamics have been conducted and the following properties of the system have been evaluated : DNA parameters (including major and minor groove width, twist and bending of the helical axis); interactions between the receptors and DNA, and the correlation to bending; and orientation of the alpha helix axes of the receptor in comparison to DNA helical axis.

Introduction

Nuclear receptors form a class of ligand-inducible transcription activators effecting such diverse tissues as skin, bone, and behavioral centers in the brain, in addition to regulating reproductive and secondary sex tissues. This class of proteins includes the steroid hormone receptors that respond to estrogens, progestins, androgens, glucocorticoids, mineralcorticoids, ecdysteroids, and vitamin D [2–5].

The steroid hormone receptors have been characterized according to six functional domains which are schematically presented for three receptors in Fig. 1. The

G. Wipff (ed.), Computational Approaches in Supramolecular Chemistry, 419–439.

DNA-binding(DBD) domain is a localized subdomain of the receptor which alone is responsible for sequence specificity and exhibits a high degree of structural and sequence homology between the various receptors in this family. The DBD, whether

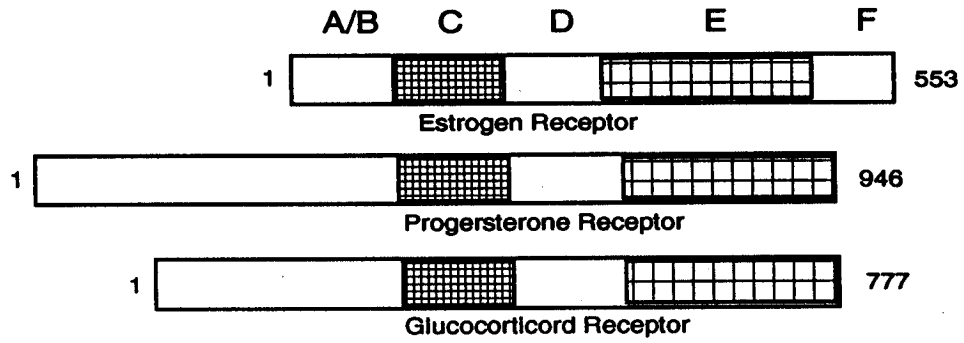

Figure 1: Functional regions of three steroid hormone receptor proteins. A/B, F are the modulating regions; C is the DNA-binding domain (DBD); D is the hinge region; E is the ligand binding region.

included as a subdomain of the whole receptor or expressed as the only domain of a truncated receptor, enables recognition and binding to a specific sequence of DNA, termed the response element [6]. The structure of the DBD is somewhat similar to the zinc finger motif in which four cysteine residues are tetrahedrally coordinated to a zinc ion and serve to orient two alpha-helices nearly perpendicular to one another [1,7,8].

The response element for the steroid hormone receptors is a palindromic or nearly palindromic DNA strand of hexameric half-sites separated by three base pairs (GRE, figure 2). The three centrally located base pairs are non-specific allowing, for example, a perfectly palindromic DNA strand to serve as a response element. The latter, labelled GRE3, is shown together with a consensus glucocorticoid response element,GRE. Once bound to DNA, one of the above mentioned alpha-helical segments of the DBD, the recognition helix, fits into the major groove of the DNA and makes specific contacts with nucleotides, while the other alpha-helix is oriented parallel to the DNA axis and makes contacts with the DNA backbone. The corresponding helical arrangement is presented in figure 3.

It has been demonstrated that the complete receptor, as well as proteins consisting of just the DBD subdomain of the steroid and thyroid hormone receptors, exist as monomers in solution but dimerize upon binding to DNA [9,10]. The method of gel-assay analysis which determines the degree of bending based upon mobility [11,12], indicates that the DBD of the estrogen receptor will induce a bend in the DNA of 34deg (deg will denote an angle measured in degrees.) and an even greater bend for the complete receptor [13,14].

The crystal structure of the DNA binding domain of the glucocorticoid receptor bound to a segment of DNA has been determined by Sigler et. al. [1]. This structure

Figure 2: Response element sequences. GRE4 is the non-consensus response element used for the crystallographic analysis, with axis of symmetry passing inbetween base pairs. GRE3 is a perfect palendrome response element for the glucocorticoid receptor which was created from the crystal structure. The consensus sequence,GRE, contains several variable base pairs labeled N and displays the numbering scheme used for this paper.

displays only minor deviations from linearity; however, in order to obtain crystals with an axis of symmetry passing between base pairs, the three base pair spacing at the center of the glucocorticoid response element had been replaced in [1] by four base pairs. The resulting DNA base pair sequence, labelled GRE4, is also shown in figure 2. The GRE4-protein-dimer complex exhibits for one protein monomer specific contacts with the consensus sequence and for the other monomer only contacts with a non-consensus sequence, i.e., in the observed crystal structure one monomer appears to bind tightly to the DNA strand, whereas the second monomer appears to exhibit weaker interactions with the DNA strand. The strong dimer interactions are presumed to outweigh the tendency of the second monomer to make contacts with the DNA as evidenced by the crystal structure.

In the present molecular dynamics study we have altered the DNA strand of the crystal structure back to the GRE3 form and adapted the atomic coordinates of the protein monomer subunits accordingly. Water was added to the resulting structure and we carried out molecular dynamics simulations to investigate the protein–DNA interactions for the natural response element GRE3. The simulations revealed, as shown below, that the two monomer subunits develop energetically comparable interactions with the response element and that the DNA strand assumes a bend of a magnitude which is in excellent agreement with bending angles observed for dimers of the DBD of the estrogen receptor [13,14].

Methods

In this section we will describe three structures which defined initial states for three molecular dynamics simulations: the complex of the GRE3 strand with a dimer of glucocortcoid receptor DBD's encapsulated in water (referred to as DNA–GR–DBD dimer), the GRE3 strand alone encapsulated in water (referred to as DNA), and a single glucocortcoid receptor DBD encapsulated in water(referred to as GR DBD). We will also briefly characterize the molecular dynamics simulations conducted.

Construction of Initial Structures

Our simulations were based on the crystal structure of the DNA-binding domain(DBD) of the glucocorticoid receptor dimer complex with a non-consensus segment of DNA (GRE4, see figure 2) [1]. The GR–DNA complex with consensus spacing of three base pairs corresponding to the strand GRE3 in figure 2 was created from the original crystal structure by separating the DNA from the rest of the structure and deleting one of the four base pairs that served as spacing, thus dividing the strand of DNA into two segments.

One of the resulting segments of DNA was a consensus half-site already aligned in the crystal structure to make specific contacts with a monomer, while the other segment was a consensus half-site displaced one base pair along the helical axis. Since the DNA was observed to be in the B–form, rotating the second segment by 36deg and translating it by 3.4A (A will denote a length measured in angstrom.) along the helical axis properly connected the two segments of DNA. This was accomplished using the coordinate transformation routines available in the structure refinement package X-PLOR [15].

In order to relax any local defects in the combined structure 500 steps of energy minimization were conducted on the DNA segment using X-PLOR. During the minimization only the base pairs adjacent to the deletion were allowed to relax. At the end of the minimization helical parameters for the patched region were found to be in agreement with B–form DNA.

The DNA was then recombined with the dimer such that the relationship, as observed in the crystal structure, to the monomer making contacts with the consensus DNA half–site was preserved. This procedure brought the second monomer in alignment with a consensus DNA half–site. Energetic analysis revealed that the alignment was favorable, so no further manipulation of the system in order to induce specific contacts between the second monomer and the consensus sequence was attempted.

The monomer that realized specific contacts with the consensus half–site in the crystal structure, henceforth, will be referred to as monomer 1, while the the other monomer that was aligned with a consensus half–site after altering the crystal structure will be referred to as monomer 2. The resulting complex of the DNA strand, monomer 1, and monomer 2 is presented in figure 3.

Figure 3: GR DBD dimer complex with DNA. In order to create this structure an extra nucleotide pair, has been removed from the crystallographic structure. The protein on the right is monomer 1 and the one on the left is monomer 2.

Previous molecular dynamics simulations of DNA have demonstrated that measures must be taken to stabilize the helical conformation of DNA. Early simulations modified the charges on the phosphate groups [16, 17], later simulations obtained stability by including counter ions, explicit water, and/or base pair constraint functions [18–22]. The protein dimer as described above containes four zinc ions and several positively charged residues which partially compensate the charge of the DNA; the resulting net charge of the system, however, still being rather large, -22e .

A short simulation of the GR-DBD complex, as described above without any water molecules, revealed that the DNA structure was not stable and that the DNA rapidly lost B-form helical parameters. Therefore, water was added to the system using the algorithm WATER [23]. WATER employs a search and replace algorithm in which a box of equilibrated water, consisting of 125 waters and measuring 15.55A X 15.55A X 15.55A is replicated on a grid in order to fill the simulated volume. A search of the space 1.6A from each water's oxygen atom is conducted; if van der Walls contacts with the solvated DNA-protein complex are favorable, the respective water molecule is included, otherwise it is not included.

The procedure outlined placed several waters inbetween monomer 2 and the DNA. These waters were removed from the system in order to obtain an overall arrangement of water similar to the water conformation observed in the cystal structure between monomer 1 and DNA. No waters were added through the procedure described to the region between the DNA and monomer 1, only the waters which were observed in the crystal structure remained included. Thus the placement of waters inbetween protein and DNA was similiar for monomer 1 and monomer 2. Obviously, attention to the

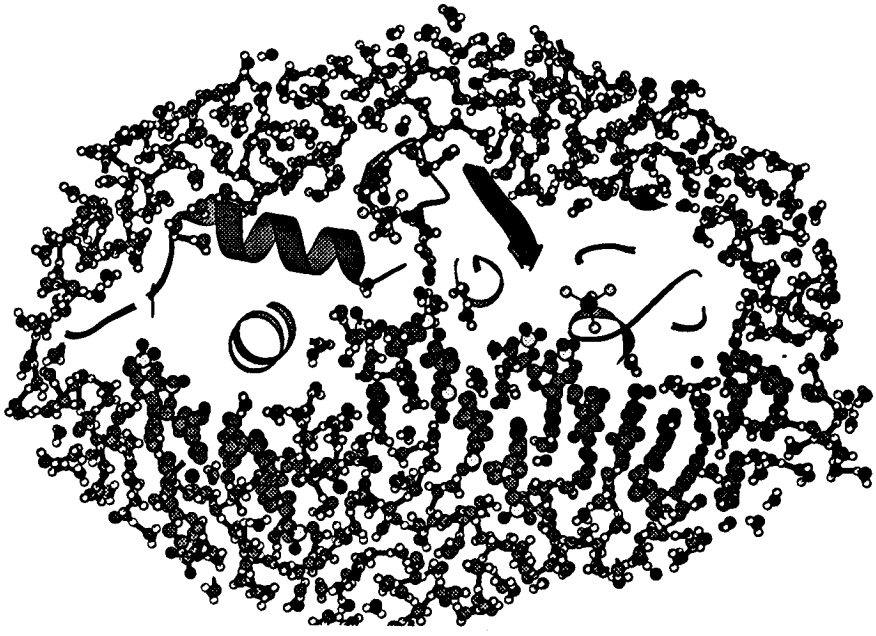

Figure 4: Vertical slice through the DNA–GR–DBD dimer system that was simulated. In order to increase stability 3,359 water molecules were added to the protein-DNA complex. The system measured approximately 60 A X 60 A X 80 A.

placement of waters inbetween the protein and DNA is required since the time scale of this molecular dynamics simulation is too brief to account for diffusion effecting water in the interior of the system.

The simulated volume for the complete GR-DNA system was chosen as an ellipsoid enclosing the complex with major axes measuring approximately 60 A X 60 A X 80 A. WATER added 3359 waters to the system which subsequent dynamics proved is sufficient to stabilize the B-form of DNA while keeping the total number of additional atoms to a minimum. Therefore, no additional steps were taken to reduce the -22e charge. A slice through the resulting water-DNA-GR system which was used for the molecular dynamics simulation is presented in figure 4.

Two other systems were also constructed using the same procedures. A second system, referred to as the GR DBD system, consisted of only one GR DBD monomer and 1052 water molecules. The computational domain for this system was defined as an ellipsoid enclosing the monomer with major axes measuring 60 A X 40 A X 35 A. This system served to demonstrate how binding to DNA and dimerization affects the properties of the GR DBD.

A third system, referred to as the DNA system, consisted of the sequence of DNA used for the GR-DNA system enclosed in an ellipsoid of 2112 water molecules and measured 86 A X 43 A X 43 A. The DNA strand was offset 4.00A from the center along the minor axis of the ellipse such that the location of the DNA in the water environment was similar to the location of the DNA in the GR-DNA system. Simulations of this system were used as a control to demonstrate that the boundary

conditions and local water environment did not induce bending of the DNA.

Molecular Dynamics Simulations

Parameters for the all atom representation of protein and DNA were taken from release 3.1 of X-PLOR [15]. All atom types for the DNA were renamed in order to avoid overlap with the protein parameterization, and specific protein-DNA interaction parameters from X-PLOR were employed. For water the TIP3P parameterization was used [24].

All simulations, except for the initial minimization of the constructed DNA sequence, were conducted on a Parsytec GCel-64, a 64 processor transputer system, using the molecular dynamics program EGO [25,26]. EGO has been developed specifically for message passing parallel computers. The program is completely compatible with the CHARMm energy routines [27], and uses standard molecular dynamics input/output formats. EGO does not incorporate cut–off approximations for the non–bonded interactions, instead all interactions are calculated by means of a distance class algorithm combined with a generalized Verlet algorithm [28].

To avoid waters escaping from the system, a harmonic well was implemented with boundaries slightly larger than the measurements used in constructing the system. For this implementation atoms were not constrained to a specific location in space, rather any atom crossing the elliptical boundary experienced a restoring force directed normal to the surface of the ellipse circumscribing the volume of interest.

The simulations for each system (DNA-GR-DBD dimer complex, GR DBD, and DNA) were carried out in three phases. All phases of the study used a time per integration step of 0.75 fs, SHAKE distance constraints [29] for only the hydrogen atoms, and harmonic well constraints for all atoms attempting to cross the elliptic boundary. Previous test simulations with the program EGO and the above parameters demonstrated that 0.75 fs per integration step yields the most stable energy conservation. In phase one of our simulations, zero velocities were assigned to the system and 3072 integration steps of minimization dynamics were conducted. The minimization technique that EGO employs removes energy from the system at every time step by a friction coefficient, chosen as 0.7, which rescales the velocities that result from the molecular dynamics conversion of unfavorable potential energies into kinetic energy. In addition, the maximum move of any atom for a single integration step was limited to 8.0 X 10E−3 A.

In the second phase of the simulations, 3072 steps of equilibration dynamics were run in which the velocities were gradually rescaled, according to the following scheme in order to achieve a stable target temperature of 300 K.

$$V_{\text{new}} = V_{\text{old}} \times \sqrt{1 + (\frac{T_{\text{target}}}{T} - 1) \times \frac{\Delta t}{\tau}} \tag{1}$$

T–target is the target temperature; T is the instantaneous temperature of the system;

Delta t is the time per integration step, and tau is the coupling constant. The time constant,tau, was 0.75 X 10E−13.

In the third phase of the simulations, molecular dynamics calculations were conducted with stochastic forces applied only to the added waters and the Verlet integration scheme was employed for all other particles. This phase of the simulation covered 85120 integration steps for the DNA-GR-DBD dimer complex, or approximately 64 ps.

Results

In this section a preliminary analysis of the structural properties of the system and most of the significant interactions between protein and DNA are discussed; the interactions between the monomer units 1 and 2,although significant, will not be considered in this paper. For the purpose of the analysis, snapshots of the molecular dynamics trajectory have been taken approximately every 1.0ps. Particular attention is focused on the overall structure of the DNA and protein subunits.

STRUCTURAL ANALYSIS

The structural details of the control of transcription by nuclear receptors, or other regulatory proteins, is at present still unclear, though it is apparent that the interaction of regulatory proteins with DNA often induces a bend in the DNA [11,30–38,44,45],but this is not a necessity for transcriptional control since there also exist DNA binding proteins controlling transcription that do not induce a bend [39,30,31]. For the present molecular dynamics study a relatively short segment of DNA is being simulated so that only local effects can be observed. However, there are sufficient interactions in this segment to confer sequence specificity and to observe a bend in the DNA. Rms deviations from the crystal structure have been calculated in order to assess the overall stability of the complex and individual structural elements. For the rms calculations each particular subset of atoms was first least square fit to the crystal structure in order to remove any rotation or translation of the whole system.

Each monomer has been further broken down into three segments: the back-bone of the recognition helix (residues 457 to 469), the back-bone of the parallel helix (residues 492 to 503), and the dimer region (residues 470 to 491). These segments are represented schematically in figure 5. The back-bone of the individual monomers has been fit independently, and the backbone of the whole dimer has also been fit as a single unit to the original configuration. The phosphate back-bone of the DNA has also been independently fit to the original configuration. The rms deviations resulting from the simulations of all three systems,DNA-GR-DBD dimer complex, GR DBD, and DNA, are presented in Table 1 and in figure 6.

The rms deviations of the complete back-bone of the individual monomers in water and of the monomers engaged in a complex with DNA are similar, indicating that

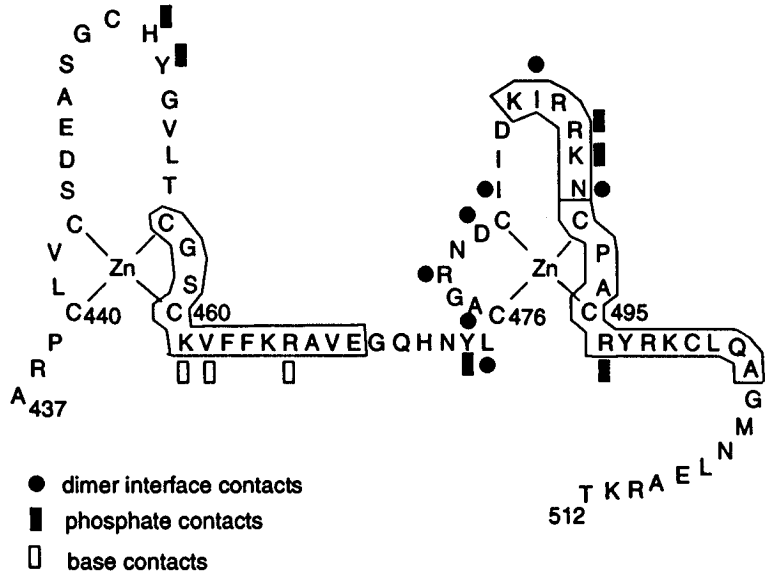

Figure 5: One letter protein sequence of each monomer. The boxes indicate sequence specific and non-sequence specific contacts, and dots indicate dimer interface contacts observed in the crystal structure. Outlined regions are the alpha helical units. The first alpha helix is the recognition helix.

each monomer is stable in the initial configuration. Graphics animation indicates little change in the spatial relationship between monomer units, and the fact that the rms deviation of the complete dimer back-bone is not significantly higher than the observed deviation for either of the individual monomers indicates that the spatial relationship of the monomers is stable in the initial conformation.

The rms values presented in figure 7 indicate that the major alpha helical units (residues 457 to 469 and residues 492 to 503) remain quite stable and contribute little to the overall deviation, while the dimer region (residues 470 to 491) accounts for most of the deviation from the crystal structure. Comparison of the rms values for monomer 1 and monomer 2 from the DNA-GR-DBD dimer complex to the values from the simulation of the GR DBD system indicate a similar stability. A much greater stability was observed for the DNA in the simulation of the DNA-GR-DBD dimer complex, average rms deviation for the phosphate backbone being 1.9A , than the for the DNA in the simulation of the DNA system, average rms deviation for the phosphate backbone being 2.7A.

The instabilities of the DNA in the simulation of the DNA system can be partially attributed to a widening of the major groove and loss of helical parameters that other simulations of such system have also observed [20,19].

The simulation of the DNA system was conducted as a control to demonstrate that the bending of the DNA strand engaged in the complex with GR is not an artifact of the simulation. In fact, the uncomplexed DNA strand does not exhibit

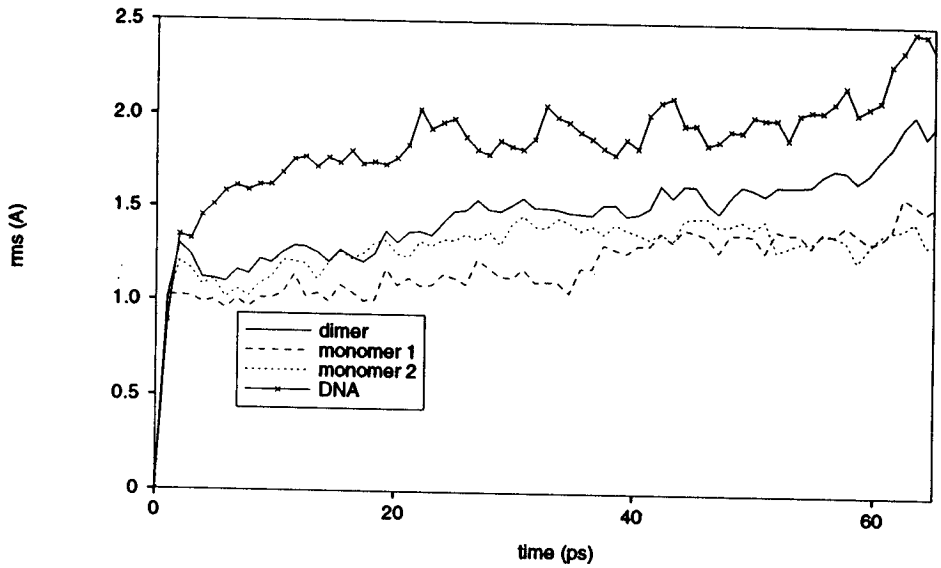

Figure 6: Rms deviation from crystal structure. These values are taken from the simulation of the complete DNA-GR-DBD system. The alpha carbon atoms were used for the protein segments and the phosphate back-bone atoms were used for the DNA segments. Separate lines indicate the rms values calculated using the following: only monomer 1 back-bone atoms, only monomer 2 back-bon atoms, the complete dimer back-bone, and the DNA back-bone.

any bending in our simulation of the DNA system.

Major and minor groove widths, which measure interstrand phosphate–phosphate distances, and twists which measure the turn from one base pair to the next along the helical axis are indicators of the stability of the DNA strand and are used as indicators of the specific form of DNA (A,B,Z,etc). Values of these quantities resulting from our simulation of the DNA system indicate large standard deviations, with the minor groove remaining more stable in the simulation. The relevant quantities are listed in Tables 2 and 3.

The average groove width values obtained from the simulation of the complete protein-DNA system indicate that the complexed DNA remains in the B-form with slightly wider than average, but very stable, major and minor grooves. The slight widening of the major and minor groove in the simulation of the complex may be due to the nature of molecular dynamics simulations but a wider than average major groove, resulting from the intrustion of the recognition helix, was also observed in the crystal structure.

The major groove distance vectors A–7 to T–2, G–6 to A–1, T+2 to A+7, and A+1 to G+6 pass through the N-terminal and beneath the C-terminal regions of the back-bone of the recognition helices of monomer 1 and monomer 2, respectively. The two sets of measurements show the same trend of being tighter at the N-terminal re-

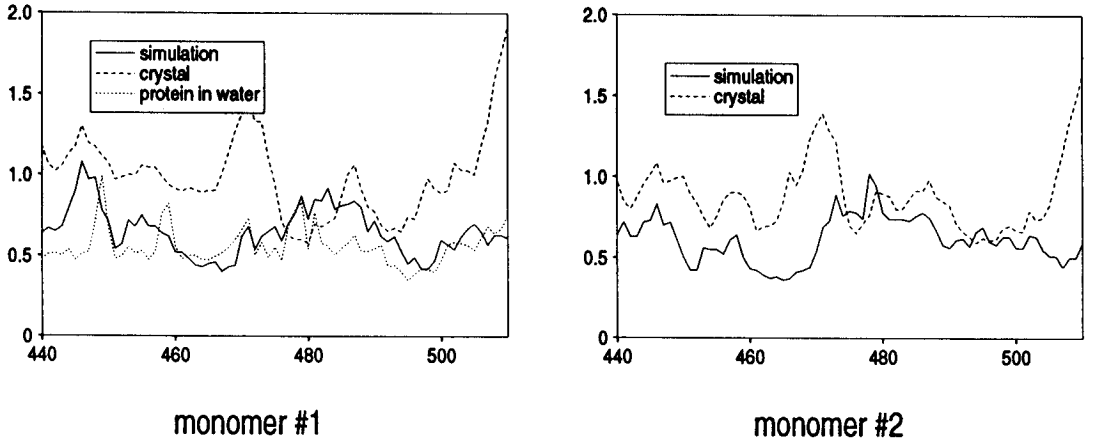

Figure 7: Graph of average rms values in angstrom versus residue number for the last 30 ps of simulation. Left monomer 1, right monomer 2. These values compare experimental rms deviations obtained from the crystal structure, labeled crystal, to theoretical values obtained by the simulations of the molecular assembly in water, labeled simulation. The values from the simulation of only monomer 1 in water have also been included in the diagram on the left.

gion, where the alpha helix makes specific base pair contacts, than at the C-terminal region where the alpha helix is farther away from the bases. Animation of the molecular dynamics trajectory reveals that the sets of lines joining these particular phosphate atoms initially have different intersections with the back-bone of the recognition helices but very quickly develop similar spatial relations. This is a geometrical indication of the molecular recognition which takes place and is described in the following section.

The most obvious structural change developing in our simulation of the complexed DNA strand is a bend which is initiated during the minimization and equilibration

complex		(A)	
	monomer 1	1.2	0.2
	monmer 2	1.3	0.1
	DNA	1.9	0.3
water			
	monmer 1	1.1	0.2
	DNA	2.7	1.0

Table 1: Table of average and standard deviation of rms values. These values compare rms deviations from the crystal structure for the simulation of the DNA–GR–DBD dimer complex labeled complex, to the simulations of the DNA and GR DBD systems, labeled water.

groove	avg	std	avg	std
C-8 to G-3	19.1	0.3	23.2	3.8
A-7 to T-2	18.5	0.9	25.0	5.6
G-6 to A-1	21.0	0.2	26.5	5.7
A-5 to C 0	20.7	1.3	24.9	5.5
A-4 to T+1	20.0	0.6	25.5	4.9
C-3 to A+2	18.2	0.5	23.5	4.8
A-2 to C+3	19.9	0.8	22.4	3.3
T-1 to A+4	20.8	1.4	22.8	3.3
C 0 to A +5	16.3	0.7	20.9	4.0
A+1 to G+6	20.5	0.0	20.4	0.0
T+2 to A+7	18.0	0.5	18.5	1.0
G+3 to C+8	15.7	0.6	18.5	2.1
B-form	17.5		17.5	

MAJOR GROOVE

Table 2: Major groove widths as measured by interstrand phosphate–phosphate distances. The table displays measurements from the simulation of the DNA-GR-DBD dimer system (left column) and DNA system (right column) The labeling for the groove column refers to figure(2)

phases and develops fully within 20ps. The bend remains stable in this configuration for the final 40ps of the simulation as evidenced by the curves algorithm of the MD-toolchest [40–42] which fits an axis to the DNA. The stabilization of the bend is also exhibited by a decreased slope in the `rms` deviation of the DNA back-bone in figure 6. The bend, presented in figure 8, is approximately 35deg from the linear strand. This angle agrees well with results from gel–assay studies. The DNA strand exhibits only a minor cork screw deviation from planarity, and the base pairs producing the bend are localized to the three central base pairs of the DNA strand, while the remaining segments of the strand remain relatively linear as can readily be seen in figure 8. The bend is driven by several favorable non-bonded interactions between residues located in the dimer region, which is rich in side chains capable of making phosphate contacts, and the DNA back-bone. In order to attain these favorable interactions the DNA back-bone is pulled into a hollow region beneath the dimer interface.

DNA–Protein Interactions

Interaction energies were determined from snapshots of the dynamics trajectory by reading the coordinates into X-PLOR and using the constraints interaction statements to isolate specific interactions. Results from these calculations are presented in figure 9. In this preliminary paper we consider only DNA–protein interactions. The DNA interaction energy values for monomer 1 and for monomer 2, which was not aligned with a consensus sequence in the crystal structure, become nearly iden-

groove	avg	std	avg	std
A-4 to T-7	11.1	0.4	12.7	1.2
C-3 to C-6	12.8	0.8	14.3	0.7
A-2 to T-5	13.1	0.1	13.4	0.2
T-1 to T-4	12.5	0.2	11.4	1.3
C 0 to G-3	12.7	0.1	11.5	1.2
A+1 to T-2	12.2	0.4	11.6	1.0
T+2 to A-1	13.6	0.8	13.4	1.0
G+3 to C 0	13.5	2.2	15.6	0.1
T+4 to T+1	13.9	0.3	15.0	0.8
T+5 to A+2	16.2	1.0	14.3	0.9
C+6 to C+3	14.9	0.2	13.9	0.8
T+7 to A+4	14.5	0.3	13.7	1.1
B-form	11.5		11.5	

MINOR GROOVE

Table 3: Minor groove widths as measure by interstrand phosphate–phosphate distances. The table displays measurements from the simulation of the DNA-GR-DBD dimer system (left column) and DNA system (right column) The labeling for the groove column refers to figure(2)

tical within 20ps indicating comparable levels of recognition. The major component of this total energy results from the electrostatic energy function.

The time scale for the drop in interaction energy and subsequent stabilization correlates very well with the bending motion observed for the DNA, which suggests that the bent conformation is necessary for molecular recognition. The favorable contacts cannot be attributed to any one region, rather it is distributed over several domains with key residues in each domain being responsible for the major contributions: residues 451 and 452 of the N-terminal region; residues 459, 461, 465 and 466 of the recognition helix; residues 486, 489 and 490 in the dimer region; and residue 496 of the parallel helix.

N-Terminal Region, Residues 437 to 456. Residues His451 and Tyr452 in this region are responsible for non-sequence specific contacts and are the primary source of interaction for this region. His451 of monomer 1 makes a stable contribution of approximately -80kcal through contacts with the phosphate group of Cyt+8. His451 of monomer 2, while making the corresponding contact with DNA is only able to achieve a stable interaction of -30kcal. The tyrosine residues 452 of each monomer make phosphate interactions with Ade±7 of -30 and -40kcal for monomers 1 and 2, respectively. The preference for this configuration is evidenced by monomer 1 which initially has this contact, loses it, then regains it as the terminal hydrogen, which makes the contact, rotates about the proximate oxygen atom.

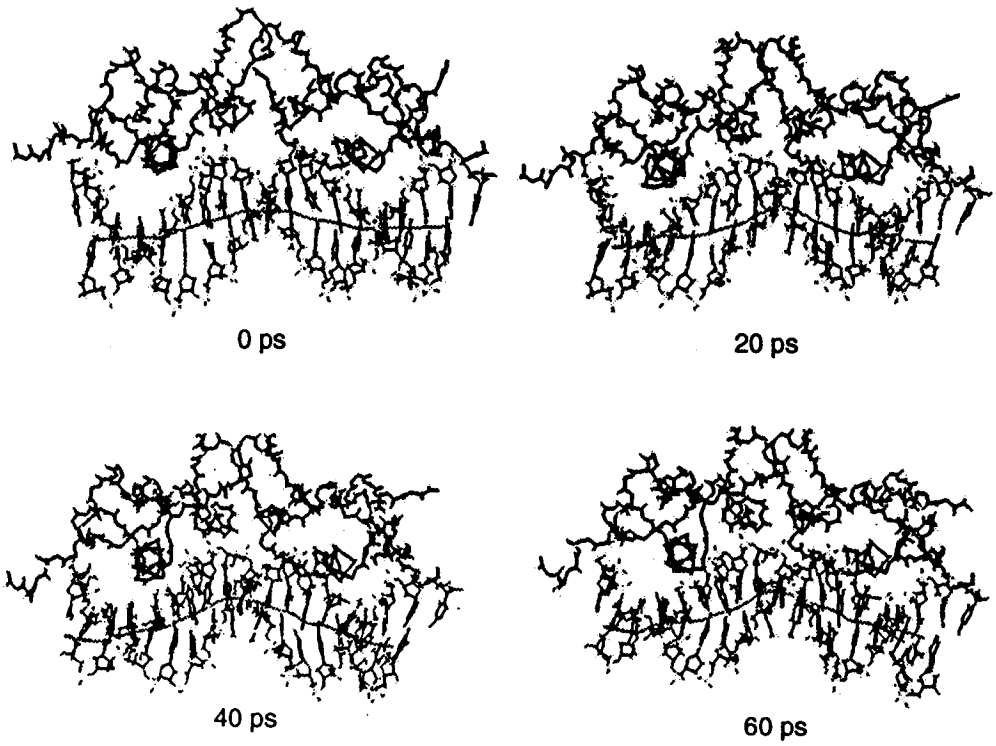

Figure 8: Bending of the DNA axis. An axis drawn down the DNA helix indicates how the bend develops.

Recognition Helix, Residues 457 to 469. Residues Lys461, Val462, and Arg466 have been identified as making specific contacts with base pairs through mutational and methylation protection experiments. In addition, residues Gly458, Ser459 and Lys461 are responsible for discrimination between the estrogen and glucocorticoid response sequences [2]. Residues Lys461 and Arg466 of each monomer, in fact, are making significant contributions to the total interaction energy; however, Val462 of each monomer never achieves a total interaction energy lower than -7 kcal. The limited interaction energy for Val462 is the result of interactions with the methyl groups of thymines at positions four and five (Thy±4, Thy±5).

Ser459 of each monomer makes important contributions to the total energy through identical contacts with phosphate groups of the thymines at postition four(Thy±4); neither residue was reported as making such contact in [1]. As noted above Ser459 is one of the three residues which is required to differentiate the glucocorticoid response element from the estrogen response element.

Lys461 of each monomer makes strong contacts with N7 of the adenines at position seven (Ade±7), but Lys461 of monomer 1 is coordinated between the N7 of Gua-6 and the phosphate of Ade-7 while Lys461 of monomer 2 is coordinated with only the phosphate of Cyt+8. The structural difference can be traced to rotations about the χ_3 angle of Lys461. Both achieve about the same interaction energies.

Lys465 of each monomer makes phosphate back-bone contacts; however, monomer

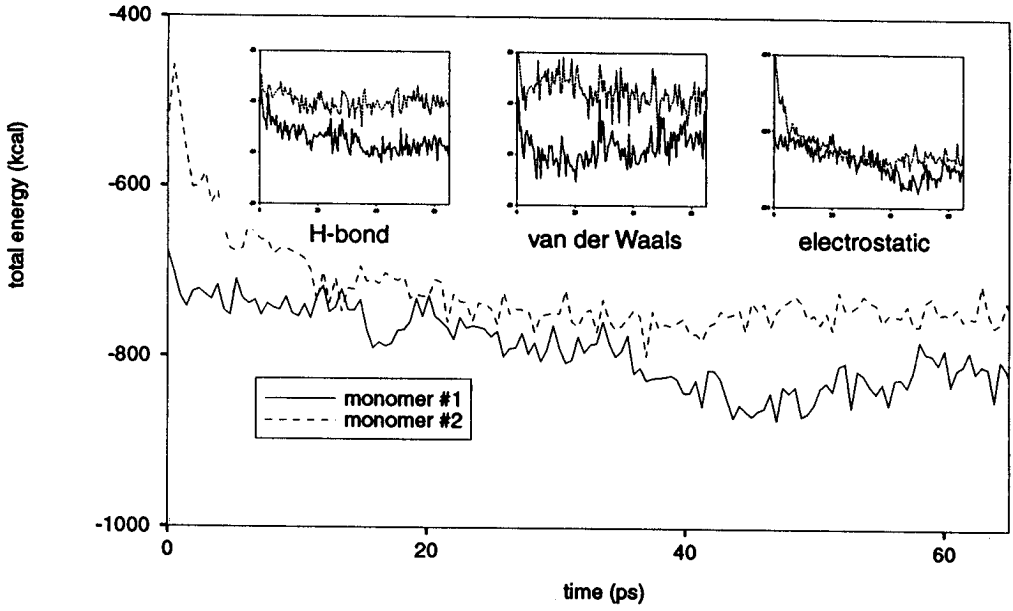

Figure 9: Protein-DNA interaction energies. Monomer 2 achieves nearly the same molecular recognition as monomer 1 within 20ps.

2 is able to achieve a lower energy for this residue by nearly a factor of two, through contacts with the phosphates of Ade+7 and Gua+6 while monomer 1 only contacts the phosphate of Gua-6. These phosphate contacts were not reported in [1] either.

The interaction energies of Arg466 for each monomer show markedly different energies because they are making different contacts with the DNA. The terminal hydrogens on the side chain of Arg466 of monomer 2 make contact with the phosphate group of Gua+3 while the the corresponding hydrogens of monomer 1 make contacts with Gua-3 and Cyt-3. The interaction of monomer 1 is the one observed in the crystal structure and corresponds well to the role which Arg466 plays in determining sequence specificity.

Extended Region, Residues 470 to 474. Tyr474 has previously been recognized as making a phosphate back-bone contact. Our simulations showed that monomer 1 achieves stable contacts through this residue with a corresponding interaction energy of -25kcal/mol. Additional interactions measuring -40kcal/mol and -15kcalmol are due to His472 and Asn473, respectively. Monomer 2 never developed such contacts as the side chains had an orientation which directed them away from the DNA and placed the phosphate groups out of reach.

Dimer Region, Residues 475 to 491. The contacts between Lys486 and the DNA back-bone were not present for either monomer in the crystal structure. However,

each of the monomers was able to achieve stable interactions of -60kcal through phosphate contacts with Thy±4. These contacts developed quickly and remained stable.

Arg489 of each monomer have initially very similar energies and phosphate contacts, but the side chain of the arginine on monomer 2 flips into a conformation, that separates it from the corresponding phosphate group thus destabilizing the interaction energy by some 60kcal. The residues of monomer 1 and 2 maintain similar spatial relationships to DNA so that the lower interaction energy of -110kcal could also be achieved by monomer 2 flipping back into the original conformation. It should be noted that both residues maintain favorable contacts that were not identified in the crystal structure. Lysine residues 490 have been identified as making phosphate contacts and each monomer is able to achieve these contacts, but with varing degress of stability.

Parallel Helix, Residues 492 to 503. There is only one residue in this region which accounts for virtually all of its interactions with DNA, namely, Arg496, which makes a phosphate contact. Monomer 2 is able to achieve a stable interaction which is about 40kcal lower than that of monomer 1 which measures -80kcal.

Discussion

In this preliminary report several significant results have been achieved, which are in agreement with experimental results. The results in this study agree very favorably with published results for the estrogen receptor DBD binding to DNA [13], a very similiar system in structure,sequence and function [3,7,43], and since the conclusion of the conference it has been brought to our attention that more recent results from gel assay studies of a GR-DBD dimer bound to the same consensus response element as used in this molecular dynamics study will bend DNA to a degree which agrees with our results (personal communication with Dr. David Shapiro, University of Illinois Urbana–Champaign, Oct. 1993). This beautiful agreement between vastly different experimental and theoretical results provides mutal support to the validity of the results. The gel-assay studies being a macroscopic measurement of properties however do not provide the details of the interaction that we have acheived. In fact protein–DNA interactions which were not previously described in the structural analysis have been identified: namely interactions between the Ser459, Lys465, Lys486 and Arg489 residues of each monomer and DNA. From our results it is also clear that the DNA bends away from the dimer, in other words, the protein is on the outside of the curve. The author does not know of any studies to date for the GR-DBD dimer that have predicted the direction of the bend although unpublished results from phasing studies of the intact ER dimer bound to DNA indicate a bending direction,toward the major groove (personal communication with Dr. Ann Nardulli, University of Illinois Urbana–Champaign, Oct. 1993) which is in agreement with our results for the GR-DBD dimer bound to DNA. The direction of bending is again a significant result since

there is evidence that for other protein systems such as the the Phage 434 CRO–OR1 complex [44] and the catabolic gene activator protein complexed with DNA [45] that the protein is on the inside of the DNA curve. Even more interesting, is the fact that the hetero dimer fos–jun and homodimer jun–jun will actually bend DNA in different orientations [37]. This molecular dynamics simulation clearly demonstrates that as the DNA bends away from the protein a more favorable interaction is acheived. Further analysis of the data will attempt to identify the specific structural elements and structural changes that facilitate the bending. For this preliminary report the interactions between the protein monomers have not been analysized but it is clear, as evidenced in the crystal structure which favors the dimer contacts over specific DNA interactions, that the strength of the dimer interactions plays a significant role in determining the relative orientation of the individual molecules in the complex.

The failure of the test simulation of DNA-GR-DBD dimer without explicit water molecules demonstrates the profound effect that the local environment has on molecular dynamics simulations, and the need for molecular dynamics simulations to account for such effects. Further comparisons of the simulation of the DNA system to the DNA-GR-DBD dimer system demonstrates the structural effects that molecular systems can induce upon one another. The fact that within 20ps of simulation time monomer 2 developed the same level of interaction energy with DNA as did monomer 1, which was presumed to be making specific contacts with DNA in the crystal structure, is a very strong indicator of the ability of molecular dynamics simulations to accurately represent interactions for supramolecular systems. Additionally this simulation demonstrates the ability of molecular dynamics to predict macroscopic observables that bridge the gap between the limitations of crystallographic techniques which can obtain detailed atomic structure but not necessarily for the desired structure or under the desired physiologic conditions and the limitations of macroscopic techniques, such as gel-assay studies, which can specify overall conformational properties but not reveal detailed atomic information.

Finally, it should be noted that in complementary studies of the GR-DBD-DNA system with the four base pair spacing element, conducted by Lennart Nilsson and also presented in these conference proceedings, similar protein stabilities were identified; however, a bent DNA conformation was not observed in their simulation. We presume that the dimer was not able to develop specific interactions with a non–consensus DNA sequence and therefore unable to induce a bend. Currently, we are conducting a molecular dynamics study of the DNA-GR-DBD with the four base pair spacing as in the Nilsson study in order to further verify our results and to further understand the nature of the bending conformation.

References

[1] B. F. Luisi, W. X. Xu, L. P. Freedman, K. R. Yamamoto, and P. B. Sigler. Crystallographic analysis of the interaction for the glucocorticoid receptor with

DNA. *Nature*, 352:497–505, 1991.

[2] M. G. Parker, editor. *Nuclear Hormone Receptors.* Academic Press, San Diego, CA, 1991.

[3] R. M. Evans. The steroid and thyroid hormone receptor superfamily. *Science*, 240:889–895, 1988.

[4] Vincent Laudet, Catherine Hänni, Jean Coll, Francois Catzeflis, and Dominique Stéhelin. Evolution of the nuclear receptor gene superfamily. *EMBO J.*, pages 1003–1013, 1992.

[5] Ernest Martinez, Francoise Givel, and Walter Wahli. A common ancestor DNA motif for invertebrate and vertebrate hormone response elements. *EMBO J.*, 10:263–268, 1991.

[6] Leonard P. Freedman, Ben F. Luisi, Z. Richard Korszun, Ravi Basavappa, Paul B. Sigler, and Keith R. Yamamoto. The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain. *Nature*, 334:543–546, 1988.

[7] Torleif Härd, Edwin Kellenbach, Rolf Boelens, Bonnie A. Maler, Karin Dahlman, Leonard P. Freedman, Jan Carlstedt-Duke, Keith R. Yamamoto, Jan ke Gustafsson, and Robert Kaptein. Solution structure of the glucocorticoid receptor DNA-binding domain. *Science*, 249:157–160, 1990.

[8] Min S. Lee, Garry P. Gippert, Kizhake V. Soman, David A. Case, and Peter E. Wright. Three-dimensional solution structure of a single zinc finger DNA-binding doamin. *Science*, 245:635–637, 1989.

[9] Per Eriksson and Öjan Wrange. Protein–protein contacts in the glucocorticoid receptor homodimer influence its DNA binding properties. *J. Biol. Chem.*, 265:3535–3542, 1990.

[10] Karin Dahlman-Wright, Helmi Siltala-Roos, Jan Carlstedt-Duke, and Jan ke Gustafsson. Protein-protein interactions facilitate DNA binding by the glucocorticoid receptor DNA-binding domain. *J. Biol. Chem.*, 265:14030–14035, 1990.

[11] D. M. Crothers and H. M. Wu. The locus of sequence–directed and protein–induced DNA bending. *Nature*, 308:509–513, 1984.

[12] J. F. Thompson and A. Landy. Empirical estimate of protein-induced DNA bending angles: applications to lambda site-specific recombinaton complexes. *Nucleic Acids Research*, 16:9687–9705, 1988.

[13] Ann M. Nardulli and David J. Shapiro. Binding of the estrogen receptor DNA-binding domain to the estrogen response element induces DNA bending. *Molecular and Cellular Biology*, 12:2037–2042, 1992.

[14] G. Redeuith M. Sabbah, S. Le Ricousse and E. Baulieu. Estrogen receptor–induced bending of hte xenopus vitellogenin a2 gene hormone response element. *Biochemical and Biophysical Research Communications*, 185(3):944–952, 1992.

[15] Axel T. Brünger. *X-PLOR, Version 3.1, A System for X-ray Crystallography and NMR.* The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, 1992.

[16] M. Levitt. *Cold Spring Harbor Symp. Quant. Biol.*, 47:251–261, 1983.

[17] B. Tidor, K. K. Irikura, B. R. Brooks, and M. Karplus. *J. Biomol. Struct. Dyn.*, 1:231–252, 1983.

[18] U. C. Singh, S. J. Weiner, and P. Lollman. *Proc. Natl. Acad. Sci. USA*, 82:755–759, 1985.

[19] Klaus Boehncke, Marco Nonella, and Klaus Schulten. Molecular dynamics investigation of the interaction between DNA and dystamycin. *Biochemistry*, 30:5465–5475, April 1991.

[20] S. Swaminathan, G. Ravishanker, and D. L. Beveridge. Molecular dynamics of B–DNA including water and counterions: A 140–ps trajectory for d(CGCGAATTCGCG) based on the GROMOS force field. *J. Am. Chem. Soc.*, 113:5027–5040, 1992.

[21] Mats A. L. Eriksson and Aattoo Laaksonen. A molecular dynamics study of conformational changes and hydration of left-handed d(CGCGCGCGCGCG) in a nonsalt solution. *Biopolymers*, 32:1035–1059, 1992.

[22] Richard M. Venable, Göran Widmalm, Bernard R. Brooks, William Egan, and Richard W. Pastor. Conformational states of a TT mismatch from molecular dynamics simulation of duplex d(CGCGATTCGCG). *Biopolymers*, 32:783–794, 1992.

[23] Bernard Walberer and Klaus Schulten. WATER — A program for the implicit inclusion of water in molecular dynamics simulations of proteins. Beckman Institute Technical Report TB-91-17, University of Illinois, 1991.

[24] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparision of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79:926–935, 1983.

[25] Brad Banko and Helmut Heller. *User Manual for EGO — Release 1.1.* Theoretical Biophysics Group at the University of Illinois at Urbana-Champaign, Beckman Institute, 405 N. Mathews Ave., Urbana, IL 61801, U.S.A., 1991. [Beckman Institute Technical Report TB-92-07].

[26] Helmut Heller, Helmut Grubmüller, and Klaus Schulten. Molecular simulation on a parallel computer. *Molecular Simulation*, 5:133–165, 1990.

[27] Bernard R. Brooks, Robert E. Bruccoleri, Barry D. Olafson, David J. States, S. Swaminathan, and Martin Karplus. CHARMm: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.*, 4(2):187–217, 1983.

[28] Helmut Grubmüller, Helmut Heller, Andreas Windemuth, and Klaus Schulten. Generalized Verlet algorithm for efficient molecular dynamics simulations with long-range interactions. *Molecular Simulation*, 6:121–142, 1991.

[29] H. J. C. Berendsen and W. F. van Gunsteren. Practical algorithms for dynamic simulations. In G. C. Ciccotti and W. G. Hoover, editors, *Proceedings of the International School of Physics, "Enrico Fermi"*, volume on course 97, pages 43–65, Amsterdam, 1986. North-Holland.

[30] C. P. Verrijzer, J. A. W. M. van Oosterhout, W. W. van Weperen, and P. C. van der Vleit. Pou proteins bend DNA via the pou-specific domain. *EMBO J.*, 10:3007–3014, 1991.

[31] C. Zwieb, J. Kim, and S. Adhya. DNA bending by negative regulatory proteins: gal and lac repressors. *Genes Dev.*, 3:606–611, 1989.

[32] M. L. Vignais and A. Sentenac. Asymmetric DNA bending induced by the yeast multifunctional factor TUF. *J. Biol. Chem.*, 264:8463–8625, 1989.

[33] D. J. Shuey and C. S. Parker. Bending of promoter DNA on binding of heat shock transcription factor. *Nature*, 323:459–461, 1986.

[34] R. Schreck, H. Zorbas, E. L. Winnacker, and P. A. Baeuerle. The nf-kb transcription factor induces DNA bending which is modulated by its 65-kD subunit. *Nucleic Acids Research*, 18:6497–6502, 1990.

[35] C. Moskaluk and D. Bastia. DNA bending is induced in an enhancer by the DNA-binding domain of the bovine papillomavirus sc E2. *Proc. Natl. Acad. Sci. USA*, 85:1826–1830, 1988.

[36] G. C. Kuhnke, H. J. Theres, and R. Ehring. RNA polymerase and gal repressor bind simultaneously and with DNA bending to the control region of *escherichia coli* galactose operon. *EMBO J.*, 8:1247–1255, 1989.

[37] T. K. Kerppola and T. Curran. Fos-jun heterodimers and jun homodimers bend DNA in opposite orientations: implications for transcription factor cooperativity. *Cell*, 66:317–326, 1991.

[38] T. A. Gustafson, A. Taylor, and L. Kedes. DNA bending is induced by a transcription factor that interacts with the human c-fos and α–actin promoters. *Proc. Natl. Acad. Sci. USA*, 86:2162–2166, 1989.

[39] M. R. Gartenberg, C. Ampe, T. A. Steitz, and D. M. Crothers. Molecular characterization of the GCN4-DNA complex. *Proc. Natl. Acad. Sci. USA*, 87:6034–6038, 1990.

[40] G. Ravishanker, S. Swaminathan, D. L. Beveridge, R. Lavery, and H. Sklenar. *J. Biomol. Str. Dyn.*, 6:669, 1989.

[41] R. Lavery and H. Sklenar. *J. Biomol. Str. Dyn.*, 6:63, 1988.

[42] H. Sklenar, C. Etchebest, and R. Lavery. *Proteins: Structure, Function Genet.*, 6:46–60, 1989.

[43] John W. R. Schwabe, David Neuhaus, and Daniela Rhodes. Solution structure of the DNA–binding domain of the oestrogen receptor. *Nature*, 348:458–461, 1990.

[44] A. Mondragon and S. C. Harrison. The phage 434 CRO–OR1 complex at 2.5 å resolution. *J. Mol. Biol.*, 219:321–, 1991.

[45] I. T. Weber and T. A. Steitz. Model of specific complex between catabolite gene activator protein and B–DNA suggested by electrostatic complementarity. *Proc. Natl. Acad. Sci. USA*, 81:3973–, 1984.