Protein Science (1992), *1*, 1173–1184. Cambridge University Press. Printed in the USA.

Molecular dynamics studies of a DNA-binding protein: 1. A comparison of the *trp* repressor and *trp* aporepressor aqueous simulations

ALLISON E. HOWARD AND PETER A. KOLLMAN

Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94143

(RECEIVED October 23, 1991; REVISED MANUSCRIPT RECEIVED March 27, 1992)

Abstract

The results of two 30-ps molecular dynamics simulations of the *trp* repressor and *trp* aporepressor proteins are presented in this paper. The simulations were obtained using the AMBER molecular mechanical force field and in both simulations a 6-Å shell of TIP3P waters surrounded the proteins. The *trp* repressor protein is a DNA-binding regulatory protein and it utilizes a helix–turn–helix (D helix–turn–E helix) motif to interact with DNA. The *trp* aporepressor, lacking two molecules of the L-tryptophan corepressor, cannot bind specifically to DNA. Our simulations show that the N- and C-termini and the residues in and near the helix–turn–helix motifs are the most mobile regions of the proteins, in agreement with the X-ray crystallographic studies. Our simulations also find increased mobility of the residues in the turn–D helix–turn regions of the proteins. We find the average distance separating the DNA-binding motifs to be larger in the repressor as compared to the aporepressor. In addition to examining the protein residue fluctuations and deviations with respect to X-ray structures, we have also focused on backbone dihedral angles and corepressor hydrogen-bonding patterns in this paper.

Keywords: α-helix; AMBER force field; computational chemistry; *Escherichia coli*; helix–turn–helix motif; molecular mechanics; protein conformation; L-tryptophan

The biosynthesis of the dimeric protein *trp* aporepressor is regulated within *Escherichia coli* by the *trp* repressor protein (Rose et al., 1973; Gunsalus & Yanofsky, 1980; Zurawski et al., 1981; Klig et al., 1988). When *trp* repressor binds to the trpR DNA operator, it represses the synthesis of the apoprotein by preventing the binding of RNA polymerase. This action inhibits transcription and the initiation of RNA synthesis. The regulation of two other *E. coli* operons—aroH and trpEDCBA—is also governed by the binding of *trp* repressor to these operator DNA sequences. Repression of the latter two operons inhibits the synthesis of aromatic amino acids and L-tryptophan.

The *trp* aporepressor and repressor protein compositions differ in that the repressor has one molecule of L-tryptophan bound within each of its two symmetry-related monomeric units. The L-tryptophans are considered corepressors in that they activate the repressor so the protein can bind specifically to one of the three operators. Feedback inhibition is involved in the operation of the *trp* repressor: high levels of L-tryptophan cause repression of the operons until there is depletion of the amount of available corepressor and aporepressor.

Recently, Sigler and coworkers have published several papers describing the structures of *trp* aporepressor (Zhang et al., 1987), *trp* repressor (Joachimiak et al., 1983; Schevitz et al., 1985; Lawson et al., 1988), a *trp* repressor/operator complex (Joachimiak et al., 1987; Otwinowski et al., 1988), and a *trp* pseudorepressor protein (Lawson & Sigler, 1988; Marmorstein & Sigler, 1989) in which indole 3-propionate replaces the L-tryptophan ligands in the protein. Although indole 3-propionate binds more strongly to the apoprotein than does L-tryptophan, the protein/ligand complex is unable to bind to DNA. The group has also deposited the X-ray structure coordinates of the trigonal (1WRP) and orthorhombic (2WRP)

Reprint requests to: Peter A. Kollman, Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94143.

crystal forms of the *trp* repressor and the aporepressor (3WRP) coordinates in the Brookhaven Protein Data Bank (PDB) (Bernstein et al., 1977; Abola et al., 1987).

Our group has undertaken several computational studies (Guenot & Kollman, 1992) of the *trp* (apo)repressor system. In this exploratory paper, we examine the molecular dynamics (MD) of the *trp* repressor and *trp* aporepressor proteins in aqueous medium. We report and discuss the MD results, emphasizing the dynamic behavior of the protein backbone, DNA-binding sites, and corepressor binding sites.

We had two objectives in undertaking this study. The first was to examine our molecular mechanical force field, AMBER, and our simulation methodology and software in light of the X-ray results. In this regard, we have found that despite the use of a limited amount of solvent (a 6-Å shell of water) and short simulation times (35-ps MD trajectories) we were able to adequately simulate the charged proteins. Our second objective was to compare and contrast the dynamic behavior of the repressor and aporepressor. Several interesting results have emerged. In particular, we have investigated the motion of the helix-turn-helix DNA-binding motifs and found that the presence of the corepressor influences both the intermonomeric helix-turn-helix distance and the flexibility of this structural unit. Such differences in protein dynamics could help to explain the enhanced repressor specificity for the DNA operator sites.

Results

Overall protein structure

Seventy-two percent of the *trp* repressor residues are in an α-helical secondary structure, and the remaining residues form turns or undefined secondary structure (Schevitz et al., 1985). Each monomer contains six helices, and these are labeled A–F (the residues designating helices A–F in the *trp* (apo)repressor are A, 12–31; B, 35–42; C, 46–63; D, 68–74; E, 79–91; F, 94–105) starting with the NH_3^+-terminus. The DNA-binding motif is formed by the D helix-turn-E helix regions. The two monomers are intertwined and interact extensively through all helices except D. Several scientists (see, for example, Lawson et al., 1988; Perutz, 1989) have defined a region, known as the central core, in which the allosteric conformational changes associated with the binding of the corepressor are absent. This central core contains helices A–C and F. The L-tryptophan-binding sites lie between the central core and helix-turn-helix regions. Approximately 27% of the protein residues are charged; there are 13 basic and 16 acidic residues in each monomer. The modeled proteins each contain six counterions, and the ensembles of atoms have a net charge of zero.

The first 5 ps of MD for the 1WRP and 3WRP models showed considerable variation in the potential energies (Fig. 1). Hence, the MD analysis was based upon the subsequent 30 ps. The average MD structures were generated for both models. During the calculations of the average structures, the MD trajectory coordinate sets were translated and rotated in order to obtain the best least-squares fit to the backbone atoms (N, Cα, C, O) of the initial coordinate set. These average structures are referred to as 1WRP:MD and 3WRP:MD in this paper. (Descriptions of the *trp* repressor and aporepressor computational models, as well as their reference names, are given in the Materials and methods section of this paper.)

Table 1 lists the calculated radii of gyration and solvent-accessible surface areas for several models of the *trp* repressor and *trp* aporepressor. The radius of gyration can provide information regarding the overall shape of a molecule, and Table 1 shows the similarity between the crystal, minimized, and average MD structures. In both the *trp* repressor and aporepressor, the minimized structures appear to have compacted slightly. This is consistent with previous investigations (Weiner et al., 1984) of protein minimization using the Weiner et al. force field. The calculated accessible surface areas also suggest compaction of the minimized structures. Both of the average

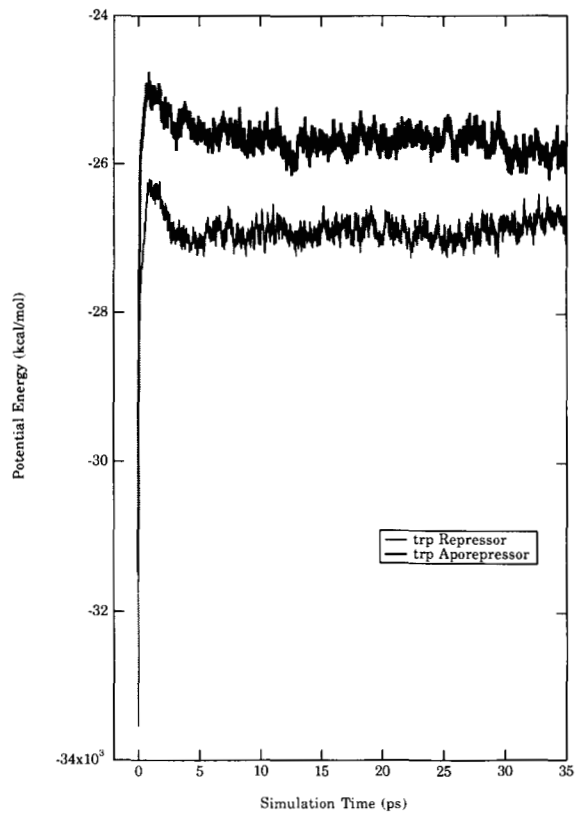

Fig. 1. Plot of potential energy versus simulation time for the molecular dynamics of the *trp* repressor and aporepressor. During the 5–35-ps interval, the mean potential energy of 1WRP:MD was −26,888.3 ± 129.0 kcal·mol^{-1} and the corresponding energy for 3WRP:MD was −25,686.5 ± 146.4 kcal·mol^{-1}.

Table 1. *Radii of gyration and accessible surface areas*[a]

	R_G (Å)	A_S (Å^2)
trp repressor		
1WRP:XRAY	18.68	12,760
1WRP:MIN	18.44	12,300
1WRP:MD	18.69	13,650
trp aporepressor		
3WRP:XRAY	18.33	12,320
3WRP:MIN	18.25	12,290
3WRP:MD	18.33	13,630

[a] Calculated radii of gyration (R_G) and accessible surface areas (A_S) for models of the *trp* repressor and *trp* aporepressor. Counterions and water molecules were excluded during the calculations. During the calculations of accessible surface areas, the radius of the solvent probe was 1.4 Å. The molecular dynamics results were generated using the average dynamical structures. The naming convention for the models is explained in the text.

MD structures show an increased surface area when compared to the crystal structures. Although one might expect the thermal motion during MD to increase the protein surface area, computer graphics inspection of the MD structures suggests that the increased surface area might also be due to burying of water molecules within the protein. The simulations were begun with a solvent shell surrounding the proteins. During the simulations, many water molecules migrated into voids of the protein and some surface residues established new solvent/protein interactions.

In Figure 2, a least-squares superposition of 1WRP:XRAY and 1WRP:MD is shown. Perhaps the most striking difference between the two structures is the compaction of the MD structure. Interestingly, the smallest deviations between the two structures occur in the central core region. We have illustrated the MD root mean square (RMS) fluctuations for the C_α atoms of 1WRP and 3WRP, and the RMS deviations between the C_α atoms of 1WRP:MD and 3WRP:MD and the X-ray structures, in Figure 3A,B. Figure 4 illustrates the RMS deviations between the C_α atoms of the various X-ray structures. The fluctuations and deviations were calculated for the entire protein, and the averages for the two individual monomers are reported in the figures.

Three regions of enhanced RMS fluctuation can be seen for the two protein models: the NH_3^+-termini regions, the COO^--termini regions, and especially the helix–turn–helix motifs. The fact that the NH_3^+- and COO^--termini regions show considerable flexibility is not surprising. As mentioned earlier, the electron density of these regions shows substantial disorder in the crystals (Schevitz et al., 1985; Zhang et al., 1987). Furthermore, the NH_3^+-terminus does not appear to participate, to any great extent, in intramolecular interactions with the rest of the protein, and these NH_3^+-termini residues do not bind with unique DNA sites (Carey, 1989). There is some evidence (Arrowsmith et al., 1989) that the NH_3^+-termini regions stabilize the A helix.

An examination of the RMS deviations for the C_α atoms of 1WRP:MD and 3WRP:MD and their corresponding X-ray structures (Fig. 3A,B) again demonstrates the enhanced mobility of not only the D helix–turn–E helix motif, but also the turn–D helix–turn region. If one further examines the RMS deviations between the X-ray structures (Fig. 4), it can be seen that the origin of the conformational differences between the three crystal structures includes both the DNA-binding motif and the turn–D helix–turn region.

The sequence-specific binding of *trp* repressor with DNA occurs when the helix–turn–helix regions interact with successive DNA major grooves—a distance of approximately 35 Å in canonical B-DNA. The major RMS differences between the 1WRP, 2WRP, and 3WRP crystal structures are also found in the helix–turn–helix regions (Lawson et al., 1988). In the 1WRP and 3WRP MD structures, the fluctuations in the helix–turn–helix regions start to become prominent in the turn between the C and D helices. The fluctuations are minimal in the E helix region of the *trp* repressor but are still substantial in the aporepressor. The magnitude of the fluctuations in the helix–turn–helix regions is also consistently larger in 3WRP.

The crystal structures, minimized coordinates, and MD

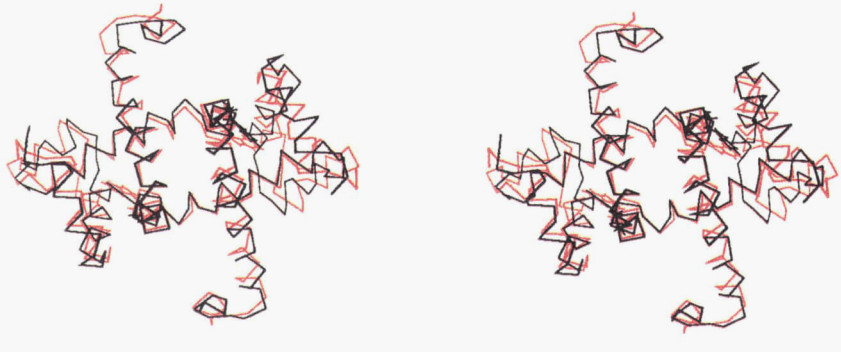

Fig. 2. Stereo view of the RMS least-squares fit between 1WRP:XRAY (red) and 1WRP:MD (black). All atoms of the L-tryptophan corepressors are shown but only the C_α atoms are depicted for the *trp* repressor protein.

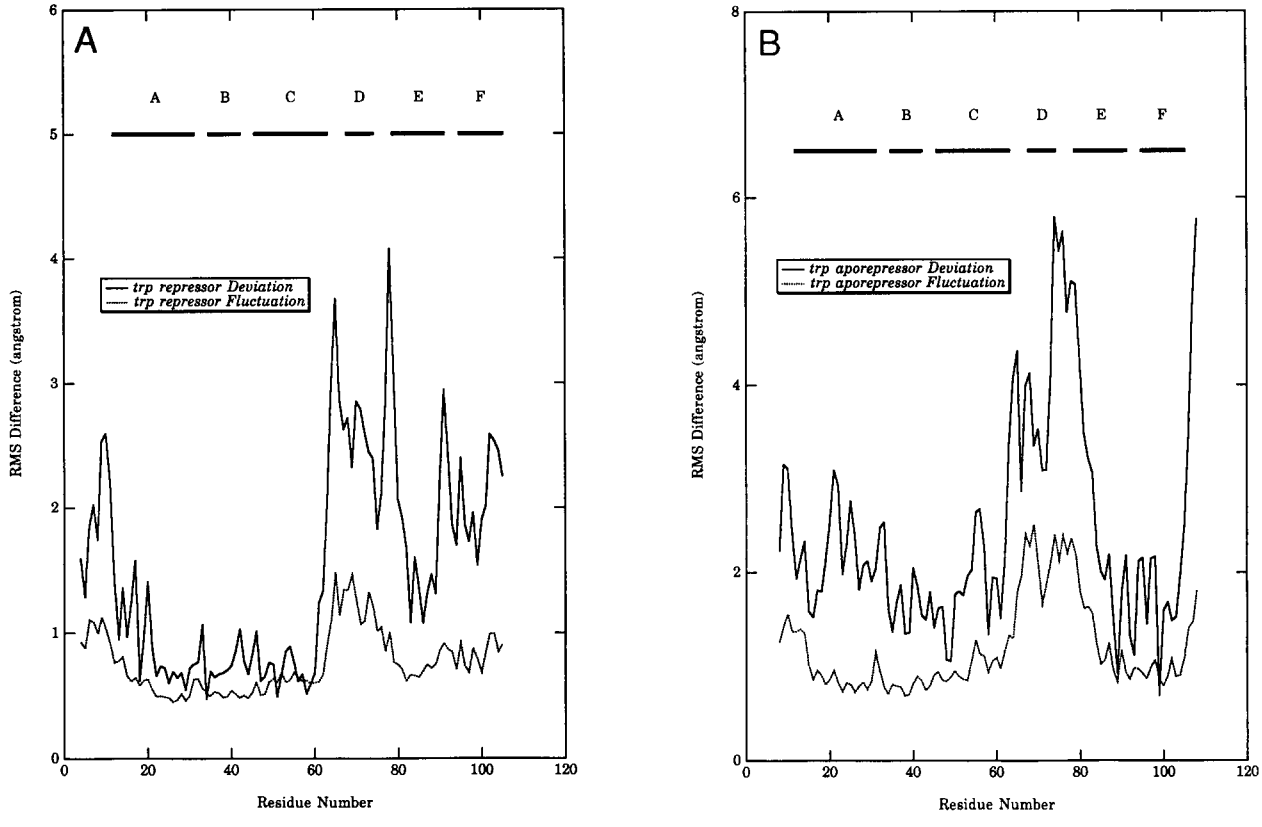

Fig. 3. For *trp* repressor **(A)** and aporepressor **(B)**, the root mean square (RMS) fluctuations of the C_α atoms during molecular dynamics (MD) have been plotted as a function of residue number. The RMS deviations between the average MD and X-ray structures (A: 1WRP:MD and 1WRP:XRAY; B: 3WRP:MD and 3WRP:XRAY) are also shown. A representation of the protein helical secondary structural units has been included in the figures.

coordinates have been compared using an RMS analysis and the results are given in Table 2. The upper right elements of the matrix were calculated by comparing main chain backbone atoms (N, Cα, C, O) in the central core region (residues 12–67, 92–105 of both monomers). The lower left triangle was formed by comparing backbone atoms (N, Cα, C, O) for all residues. In both 1WRP and 3WRP, the RMS differences between the minimized and crystal structures are small. When 1WRP:MD was compared to the 1WRP:XRAY crystal structure, a moderate difference of about 1.5 Å was seen for the central core and a larger difference of approximately 2.0 Å was found when all backbone atoms were compared. Still larger differences were found for the comparison of 3WRP:MD and its X-ray structure. Although the central core comparison gave an RMS difference of about 2.2 Å, the difference increased to almost 3.0 Å when all backbone atoms were compared.

The (ϕ, ψ) backbone dihedral angles (IUPAC-IUB, 1970) were calculated for residues 2–105 (2–104 in the case of 1WRP ψ) from the Cartesian coordinates saved

Table 2. *Root mean square (RMS) structural comparisons*[a]

	1WRP:XRAY	1WRP:MIN	1WRP:MD
trp repressor			
1WRP:XRAY	–	0.39	1.47
1WRP:MIN	0.49	–	1.47
1WRP:MD	1.95	1.95	–
	3WRP:XRAY	3WRP:MIN	3WRP:MD
trp aporepressor			
3WRP:XRAY	–	0.35	2.20
3WRP:MIN	0.41	–	2.20
3WRP:MD	2.91	2.90	–

[a] Comparison of the RMS structural differences (Å) between models of *trp* repressor and *trp* aporepressor. In each of the matrices, the elements in the upper right triangle were calculated by comparing main chain backbone atoms (N, Cα, C, O) in the CORE region (residues 12–67, 92–105 of both monomers). The lower left triangle was formed by comparing backbone atoms (N, Cα, C, O) for all residues. The molecular dynamics results were calculated using the average dynamical structures. The naming convention for the models is explained in the text.

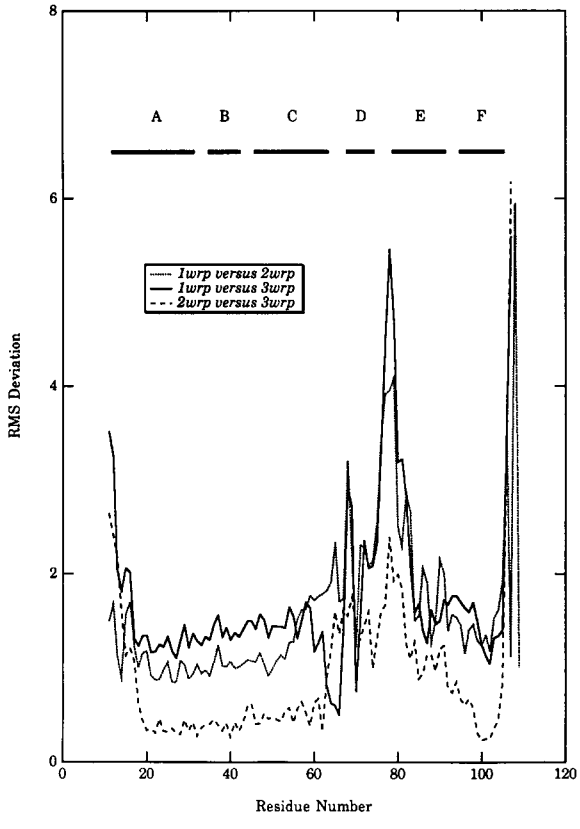

Fig. 4. Root mean square (RMS) deviations, as a function of residue number, between the C_α atoms in several X-ray structures of *trp* (apo)repressor. The figure shows the RMS deviations between 1WRP:XRAY-2WRP:XRAY, 2WRP:XRAY-3WRP:XRAY, and 1WRP:XRAY-3WRP:XRAY. A representation of the protein helical secondary structural units has been included in the figure.

during the course of the 1WRP and 3WRP MD trajectories. The mean dihedral angles and the mean dihedral angular deviations were then calculated (Batschelet, 1981; Spellmeyer & Howard, unpubl.). The mean angle defines an average dihedral angle and the mean dihedral angular deviation is a measure of the width of a sample distribution (see Materials and methods).

The (ϕ, ψ) mean dihedral angles during MD for 1WRP and 3WRP helical secondary structural units are listed in Table 3, along with the corresponding values for the X-ray structures. The agreement between the MD and X-ray values is reasonable. With respect to the mean (ϕ, ψ) dihedral angular deviations, Figure 5A,B displays plots of the angular deviation as a function of residue number. The trend for these data is similar for what was seen in the RMS fluctuations: that is, there is large variability for the NH_3^+- and COO^--termini and the helix-turn-helix motifs. One additionally finds moderately large angular deviations associated with residues having undefined secondary structure. However, unlike the RMS fluctuations of the C_α atoms, the mean dihedral angular deviations in and near the helix-turn-helix motifs are not consistently large. Rather, specific residues appear to act as *pivots* (Brooks et al., 1988; Lolis et al., 1990; Lolis & Petsko, 1990) for the structural motion. In particular, one sees very large angular deviations for some of the (ϕ, ψ) dihedral angles of the turn residues between helices C and D (Gly^{64}-Ser^{67}). Significantly large angular deviations are also found for some residues in and near the D-E (Leu^{75}-Gly^{78}) and E-F turns (Ala^{92}-Pro^{93}).

Helix-turn-helix motif

We were interested in monitoring the movements of the helix-turn-helix regions during the course of the 1WRP and 3WRP MD trajectories. In order to do this, we calculated the center of mass for each monomer's DNA-binding motif and the distance between these two regions for the coordinates of 1WRP and 3WRP during the MD simulation. We applied a similar calculational strategy to the structural elements of the motifs: helix D-helix D, turn-turn, and helix E-helix E. The results of these calculations are reported in Table 4.

For almost all of the measurements, the distance between the two monomeric structural units in the crystal structures was greater than the average found during the course of the MD trajectories. If one plots these distances as a function of trajectory length, it is seen that the period of D helix-turn-E helix motion is on the order of 20–30 ps and it appears to be greater in the aporepressor simulation. Therefore, we certainly have not sampled this motion sufficiently, and it may be that a trajectory length of one to two orders of magnitude greater than these simulations would give results closer to the crystal structures.

A proposed function of the L-tryptophan corepressor is to increase the distance between the helix-turn-helix regions so that the distance becomes appropriate for binding in successive major grooves of DNA. In both the crystal structures and MD simulations, this distance is larger for the repressor than the aporepressor. If one compares the mean distances and standard deviations for the individual structural elements during MD, a difference in the flexibility of the 1WRP and 3WRP structures can be found. For both the D and E helices, the motion between the helix centers of mass is greater in the aporepressor. The range for the helix D-helix D centers of mass motion was 3.6 Å in the repressor simulation and 7.7 Å in the simulation of the aporepressor. The range of motion was 1.7 and 4.8 Å for the helix E-helix E in repressor and aporepressor simulations, respectively.

Tryptophan-binding site

The L-tryptophan-binding site is formed by the B and E helices of one monomer and the B-C turn of the second monomer (Schevitz et al., 1985). We have analyzed the hydrogen-bonding pattern of this site during the course

Table 3. *Mean (ϕ,ψ) dihedral angles*[a]

Model	Mean (ϕ,ψ) dihedral angles (degrees)					
	Helix A	Helix B	Helix C	Helix D	Helix E	Helix F
1WRP:MD ϕ dihedrals	−58.78 ± 2.98	−60.40 ± 6.74	−59.14 ± 5.88	−63.32 ± 9.88	−58.11 ± 6.50	−64.02 ± 11.50
	(−59.86 ± 6.54)	(−62.40 ± 5.83)	(−63.44 ± 6.90)	(−67.32 ± 16.09)	(−63.72 ± 11.73)	(−65.28 ± 9.05)
1WRP:MD ψ dihedrals	−45.66 ± 4.49	−49.60 ± 6.23	−45.50 ± 7.70	−50.26 ± 6.99	−46.22 ± 15.14	−53.73 ± 26.34
	(−43.96 ± 8.20)	(−43.87 ± 11.67)	(−42.20 ± 7.08)	(−43.27 ± 9.35)	(−40.39 ± 12.38)	(−44.94 ± 8.54)‡
3WRP:MD ϕ dihedrals	−60.47 ± 6.02	−60.65 ± 6.21	−60.30 ± 8.24	−62.81 ± 6.44	−58.69 ± 6.55	−66.49 ± 13.06
	(−65.02 ± 8.47)	(−68.45 ± 5.21)	(−63.53 ± 6.67)	(−66.72 ± 9.92)	(−66.73 ± 6.42)	(−67.89 ± 23.51)
3WRP:MD ψ dihedrals	−45.47 ± 5.47	−51.90 ± 7.34	−44.12 ± 9.40	−48.81 ± 9.29	−45.42 ± 15.45	−42.13 ± 11.81
	(−38.88 ± 10.36)	(−39.84 ± 7.16)	(−40.52 ± 7.15)	(−38.47 ± 8.68)	(−36.66 ± 9.66)	(−38.21 ± 27.25)

[a] The mean (ϕ, ψ) dihedral angles (see Materials and methods) for the helical secondary structural units of *trp* repressor and aporepressor are given in this table. The data were generated by calculating the mean (ϕ, ψ) dihedral angles for each residue from Cartesian coordinate sets saved during the course of the trajectory. These angles were then used to calculate the mean (ϕ, ψ) dihedral angle for each structural unit. The error limits associated with each dihedral angle correspond to the mean angular deviation for the mean dihedral angle of each helix. The values in parentheses are those corresponding to the X-ray structure. The symbol ‡ has been used to indicate that the mean dihedral angle for this helix was determined from residues 94–104. This was due to the fact that the Cartesian coordinates were not determined for residue 106 in the X-ray structure and thus the ψ angle could not be calculated for residue 105. The naming convention for the models is explained in the text. The structural features are defined as in the Protein Data Bank files—helix A: monomer residues 12–31; helix B: 35–42; helix C: 46–63; helix D: 68–74; helix E: 79–91; helix F: 94–105.

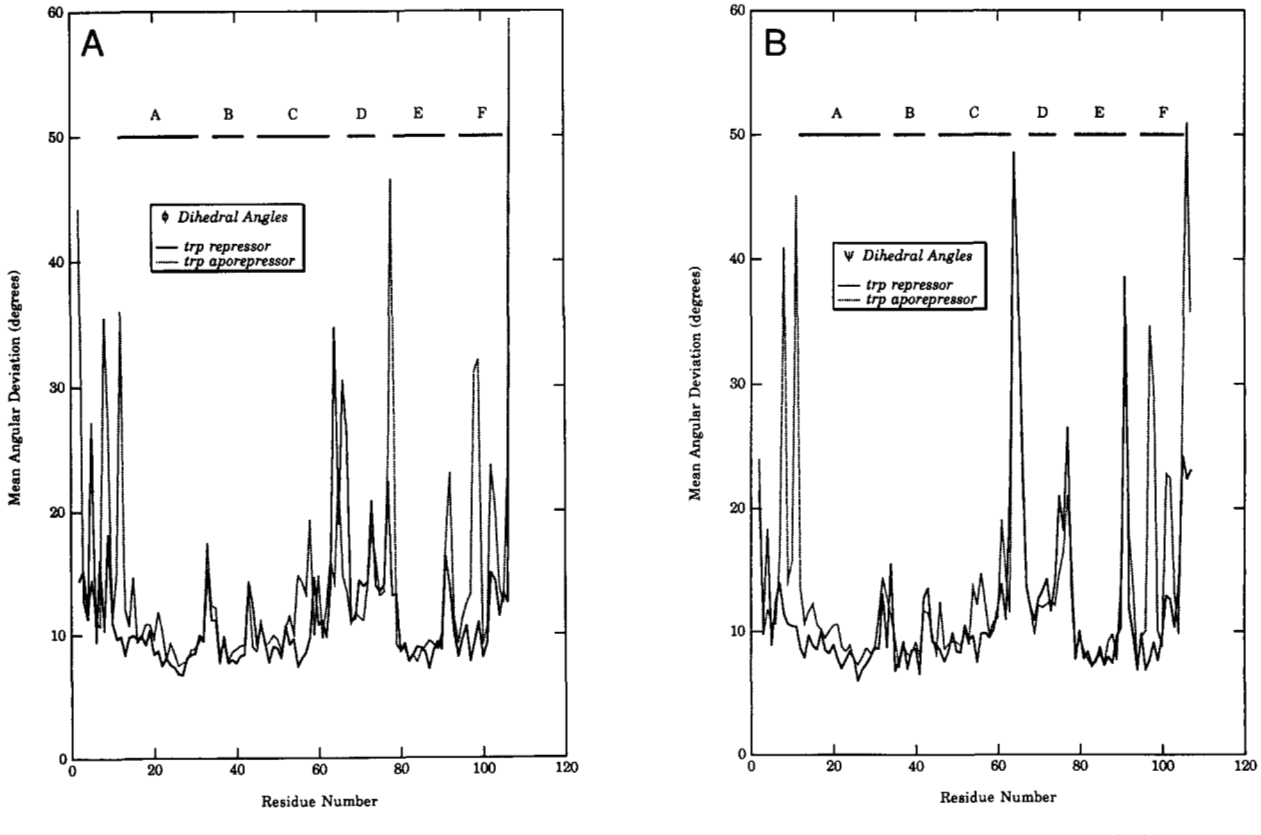

Fig. 5. Mean ϕ **(A)** and ψ **(B)** angular deviations as a function of residue number for *trp* repressor and aporepressor during molecular dynamics. A representation of the protein helical secondary structural units has been included in the figures.

Table 4. *Center of mass distances*[a]

Models	Center of mass distances (Å) for structural units			
	HTH–HTH	Helix D–helix D	Turn–turn	Helix E–helix E
1WRP:XRAY	35.42	45.38	32.96	33.16
2WRP:XRAY	35.32	45.11	33.47	32.46
3WRP:XRAY	33.29	42.89	30.80	31.11
1WRP:MD	33.44 ± 0.63	42.36 ± 0.86	28.96 ± 0.75	33.14 ± 0.40
3WRP:MD	31.91 ± 0.71	38.17 ± 2.58	25.89 ± 0.78	33.88 ± 1.32

[a] Distances (Å) between monomers for several *trp* repressor and *trp* aporepressor models. Each distance refers to the center of mass distance between symmetrically related units in the dimers. The results for the molecular dynamics models are the statistical means—shown with standard deviations—for distances calculated using the individual configurations of the trajectories. The naming convention for the models is explained in the text. The structural features are defined as in the Protein Data Bank files—HTH: monomer residues 68–91; helix D: monomer residues 68–74; turn: monomer residues 75–78; helix E: monomer residues 79–91.

of the 1WRP and 3WRP MD trajectories. In the analysis, we define a hydrogen bond if two criteria are met: (1) the donor–acceptor interatomic distance is less than 3.5 Å, and (2) the donor–hydrogen–acceptor angle is greater than or equal to 120°.

Table 5 lists the identity of hydrogen bonds found in the L-tryptophan-binding sites during the MD simulation of the *trp* repressor. We have not included the hydrogen bonds associated with the α-helical secondary structure. Six major hydrogen bonds are found between the corepressor, L-tryptophan, and the protein residues. The carboxyl terminal of L-tryptophan interacts with the NE and NH nitrogens on Arg^{84} of the second monomer, whereas the ammonium group acts as a hydrogen bond donor to residues Leu^{41} and Leu^{43} of the same subunit and Ser^{88} of the second monomer. Schevitz et al. (1985) have reported a hydrogen bond, which we do not observe in the MD simulation, between Thr^{44} and the L-tryptophan carboxyl group. However, in subsequent refinements of the X-ray structure the hydrogen bond is not found (Sigler, pers. comm.). We do observe, however, a hydrogen bond between Thr^{81} and NE1 of L-tryptophan. In both monomers, we find the Arg^{54} (NH1(2)) of one monomer forming a salt bridge to the Glu^{47} (OE1) of the second monomer.

We did not include X-ray waters in our initial models of 1WRP and 3WRP. Consequently, no water molecules were present in the L-tryptophan-binding site prior to MD. In analyzing the MD trajectory of 1WRP, the L-tryptophan carboxyl and ammonium groups were both found to hydrogen bond with two water molecules. A number of other residues in the binding site also hydrogen bond with water molecules, including Arg^{84} and Ser^{88} in both the *trp* repressor and aporepressor.

Figure 6A and B both show a "snapshot" of the L-tryptophan-binding site. Figure 6B shows the binding site with respect to other structural elements of the repressor. Figure 6A is a close-up of the binding site, and it depicts all residues and waters within 6 Å of the L-tryptophan corepressor.

Discussion

We have compared two 30-ps MD trajectories of the *trp* repressor DNA-binding protein and the aporepressor dimer surrounded by shells of water. The relatively large size of the protein has caused us to limit both the trajectory lengths and the number of waters of solvation. In particular, it is clear that the equilibration time is inadequate for complete relaxation of the protein (Levitt & Sharon, 1988). Future calculations are planned in which the equilibration and sampling times will be increased, and the proteins will be surrounded by boxes of waters in periodic boundary conditions. Another limitation of the calculations is the fact that the average temperature of the protein is significantly cooler than that of the water (see Materials and methods). Subsequent MD trajectories by Daggett and Kollman (unpubl.), using the same protocol on an α-helix in water, have found similar results. Further simulations by Daggett and Kollman on the α-helix in which water and helix were independently coupled to the temperature bath led to both being at ≈300 K and a similar trajectory to that with independent coupling. Thus, we expect that independently coupling both protein and water to a 300 K temperature bath would not change the qualitative features of our trajectory.

Despite the limitations of the simulations, the analysis of the trajectories has provided insight into the structure and dynamics of the two proteins. The most exciting result from this study is the excellent correspondence between the mobile parts of the structure found both by independent crystal structures and in our MD trajectories. In particular, a significantly greater than average movement of the helix D–turn–helix E DNA-binding motif is found in both the *trp* repressor and aporepressor trajectories and this is the part of the crystal structure

Table 5. *Hydrogen-bonding pattern of the L-tryptophan-binding site*[a]

Monomer A(B) donor				Monomer A(B) acceptor			% Occupancy
40A	Asn	ND2	···	40A	Asn	O	18, 0
44	Thr	N	···	47	Glu	OE2	100, 100
44	Thr	OG1	···	47	Glu	OE2	100, 98
47A	Glu	N	···	44A	Thr	OG1	90, 0
48	Arg	NH1	···	43	Leu	O	40, 33 (11)
53A	Thr	OG1	···	49A	Glu	O	23, 0
63	Arg	NE	···	60	Glu	OE2	10, 70 (35)
83	Thr	OG1	···	79	Ile	O	85, 60
86	Ser	OG	···	82	Ile	O	83, 85
88	Ser	OG	···	84	Arg	O	100, 94
48B	Arg	NH1	···	39B	Leu	O	0, 14
48B	Arg	NH1	···	40B	Asn	O	0, 10
63B	Arg	NH1	···	60B	Glu	OE1	85 (53)
88B	Ser	OG	···	85	Gly	O	0, 20
Monomer A(B) donor				**Monomer B(A) acceptor**			**% Occupancy**
54A	Arg	NH1	···	47B	Glu	OE1	91, 0
54B	Arg	NH1	···	47A	Glu	OE1	0, 83
54B	Arg	NH2	···	47A	Glu	OE1	0, 14
Monomer A(B) donor				**Corepressor (L-Trp) acceptor**			**% Occupancy**
84	Arg	NE	···	109	Trp	OB	93, 94
84	Arg	NH2	···	109	Trp	OB	37, 90
84B	Arg	NH2	···	109B	Trp	OA	0, 54
109A	Trp	N	···	41B	Leu	O	73 (17), 86
109A	Trp	N	···	43B	Leu	O	63 (11), 48
109A	Trp	N	···	88A	Ser	OG	17 (12), 0
109	Trp	NE1	···	81	Thr	O	75, 80
Corepressor (WAT) donor				**WAT (corepressor) acceptor**			**% Occupancy**
109	Trp	N	···	–	WAT	O	–
–	WAT	H	···	109	Trp	OA	–
–	WAT	H	···	109	Trp	OB	–

[a] Above is a list of the nonhelical hydrogen bonds found during molecular dynamics simulation of the *trp* repressor between donor-acceptor: A: within a protein monomer (monomer A(B)-monomer A(B)); B: between different monomers (monomer A(B)-monomer B(A)); C: between a monomer and corepressor (monomer A(B)-corepressor (L-Trp)); and D: between a corepressor and water molecule. The occupancies for each monomer are separated by a comma and multiple hydrogen bonds to an atom are indicated with parentheses.

most different between different crystal forms. Because the helix-turn-helix motifs are the structural units that interact with DNA, their flexibility would be expected to enhance the binding of monomers in successive DNA major grooves. The RMS fluctuations are larger in the aporepressor than repressor in the E helix region, and this is most probably a consequence of L-tryptophan interactions decreasing the *trp* repressor mobility. The helix-turn-helix COM distances are somewhat greater in the repressor as compared to the aporepressor. This observation is in agreement with the proposal that the function of the ligand is to increase the distance between helix-turn-helix regions in order to interact with the major grooves in DNA.

The helix-turn-helix residues show a fairly uniform RMS fluctuation. However, it appears that the origins of the motion may be traced to large angular deviations of specific dihedral angles flanking those secondary structural units. This suggestion has been proposed previously (Brooks et al., 1988; Lawson et al., 1988; Lolis et al., 1990; Lolis & Petsko, 1990) after the analysis of *trp* repressor and aporepressor X-ray structures.

A major difference between the MD simulation results and X-ray crystal structures can be found in and near the helix-turn-helix regions of *trp* repressor and aporepressor. Both techniques find enhanced mobility in the D helix-turn-E helix motifs, as compared to the nonterminal regions of the proteins. However, the MD simulations also find large mobility in the turn-D helix-turn protein regions. This type of behavior has also recently been reported during the analysis of the *trp* repressor ^{1}H-NMR solution spectra (Arrowsmith et al., 1990). Arrowsmith et al. found increased disorder in the D helix as compared to the X-ray results. They also found the N-terminal A helix to be more disordered in solution.

Materials and methods

The starting structure for the minimization and MD calculations of *trp* repressor was the 1WRP PDB structure—referred to in this text as 1WRP:XRAY. This model included the dimeric protein with two L-tryptophan ligands (218 residues, 3,564 atoms) and six modeled cations, surrounded by a 6-Å shell of TIP3P (Jorgensen et al., 1983) waters (1,713 residues, 5,139 atoms). The aporepressor model was based upon the 3WRP PDB structure (3WRP:XRAY) and contained a 6-Å shell of TIP3P (Jorgensen et al., 1983) waters (1,584 residues, 4,752 atoms) surrounding six modeled cations and the dimeric *trp* aporepressor protein (216 residues, 3,510 atoms). The X-ray crystallographic waters were removed from both 1WRP:XRAY and 3WRP:XRAY prior to adding the counterions and water shells. The 1WRP:XRAY and 3WRP:XRAY proteins were modeled as zwitterions, as were the L-tryptophan ligands. All atoms were explicitly represented in both the 1WRP:XRAY and 3WRP:XRAY models.

The nucleotide sequence of *trp* aporepressor (Gunsalus & Yanofsky, 1980) shows each monomer to contain 108 amino acids. However the first residue, methionine, is found to be cleaved in approximately 90% of the experimental preparations. We modeled 1WRP:XRAY and 3WRP:XRAY with the NH_3^+-terminal amino acids as methionine and not alanine, the second residue of the nu-

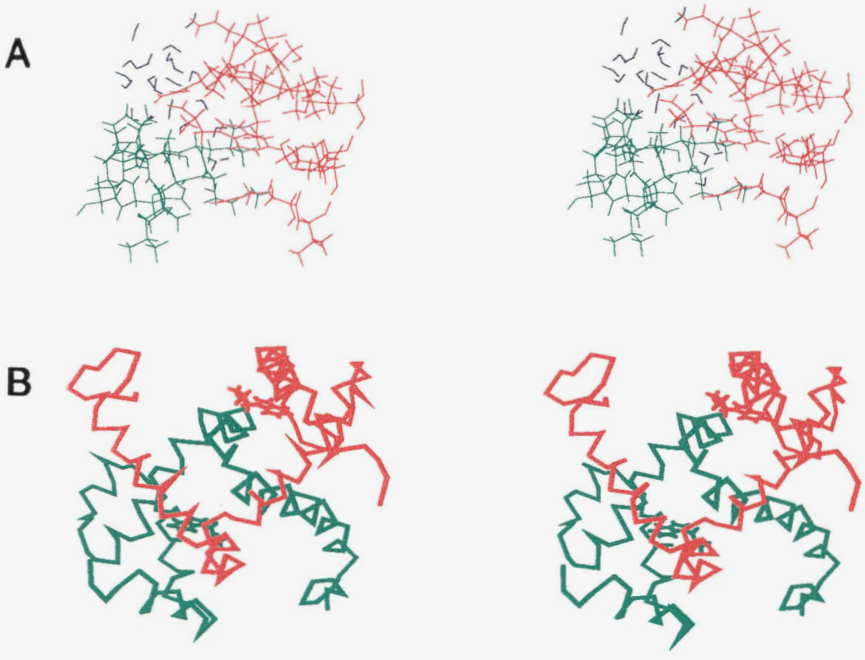

Fig. 6. A: Stereo view of one of the two corepressor (L-tryptophan)-binding sites. The figure is a snapshot from the MD trajectory of 1WRP after 5 ps of equilibration and an additional 30 ps of MD. The figure shows all residues within 6.0 Å of any corepressor atom. The color scheme shows water residues in blue. The protein residues of one monomer, including L-tryptophan, are shown in red, and those associated with the second monomer are colored green. **B:** Stereo view of *trp* repressor. This figure was drawn from the same coordinate set, and using the same coloring scheme and orientation as part A and it depicts the L-tryptophan-binding site of part A in relation to the rest of the *trp* repressor molecule. Water molecules have not been represented in this figure. Only the C_α atoms are depicted for the *trp* repressor protein monomers.

cleotide sequence. A high degree of disorder prevented the determination of the Cartesian coordinates for residues 2–3, 106–108 and 2–7 in the X-ray crystallographic structures of 1WRP (Joachimiak et al., 1983; Schevitz et al., 1985; Lawson et al., 1988) and 3WRP (Zhang et al., 1987), respectively. These residues were modeled in extended chain conformations.

Molecular mechanical calculations were conducted using the AMBER force field (Weiner et al., 1984, 1986) as implemented in the AMBER 3.0 software (Singh et al., 1986). The form of the AMBER force field potential was as follows:

$$V_{\text{total}} = \sum_{\text{bonds}} K_r(r - r_{eq})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_{eq})^2$$
$$+ \sum_{\text{dihedrals}} \sum_{\eta} \frac{V_\eta}{2}[1 + \cos(\eta\phi - \gamma)]$$
$$+ \frac{1}{VDW_{scale}} \sum_{j=1}^{\text{atoms}} \sum_{i>j}^{\text{atoms}} \varepsilon_{ij}^* \left[\left(\frac{R_{ij}^*}{r_{ij}} \right)^{12} - \left(\frac{R_{ij}^*}{r_{ij}} \right)^{6} \right]$$
$$+ \sum_{j=1}^{\text{Hbonds}} \sum_{i>j}^{\text{Hbonds}} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right)$$
$$+ \frac{1}{EEL_{scale}} \sum_{j=i}^{\text{atoms}} \sum_{i>j}^{\text{atoms}} \frac{q_i q_j}{\varepsilon r_{ij}},$$

where $\varepsilon_{ij}^* = \sqrt{\varepsilon_i^* \varepsilon_j^*}$ and $R_{ij}^* = R_i^* + R_j^*$. In this equation, the bond and angle force constants are represented by K_r and K_θ, respectively. The equilibrium bond length parameters are described by r_{eq} in the above equation and the equilibrium angles by θ_{eq}. The dihedral angles, ϕ, are represented by a truncated Fourier series potential in which V is the barrier height, η are the Fourier terms (i.e., the periodicities of the torsion), and γ is the dihedral phase shift angle. For the nonbonded terms, the distance between two atoms, i and j, is given by r_{ij}. The van der Waals radius and well depth are described by the terms R_{ij}^* and e_{ij}^*, respectively. Hydrogen-bonding nonbonded interactions are evaluated with a 10–12 potential, in which the C_{ij} and D_{ij} parameters are coefficients specific for that atom pair. It should be noted that *either* the 6–12 or the 10–12 potential component was used for any atom pair—for atom pairs designated as hydrogen-bonding acceptor/donor pairs, the 10–12 potential was used and otherwise the 6–12 term was substituted. The electrostatic term is proportional to the product of the ab initio electrostatic potential derived point charges (Singh & Kollman, 1984), q, on atoms i and j and it is inversely proportional to the dielectric constant, ε.

The dihedral term was used for both proper and improper dihedrals, where an improper dihedral is defined as one in which the designated atoms are not sequentially bonded to each other. Only pairwise nonbonded interactions were calculated during the simulations, and these were calculated for atom pairs separated by at least three bonds. For 1,4 nonbonded interactions, both VDW_{scale} and EEL_{scale} were set equal to 2 and for all other nonbonded interactions, the values of these scale factors were unity. A nonbonded cutoff of 8 Å was used for all cal-

culations. As indicated in the force field equation, a constant dielectric was used and $\varepsilon = 1$.

There is some ambiguity in the AMBER force field with respect to how hydrogen bonding should be handled between TIP3P water (Jorgensen et al., 1983) and protein hydrogen-bonding atoms. We chose to use 10–12 parameters, consistent with the Weiner et al. (1984) protein-protein hydrogen-bonding parameters, for the $H_{TIP3P} \cdots Donor_{PROTEIN}$ and $O_{TIP3P} \cdots Acceptor_{PROTEIN}$ atoms. The H_{TIP3P} and O_{TIP3P} parameters were taken from protein backbone amide-hydrogen and carbonyl-oxygen atoms, respectively. Also, the C and D values of the 10–12 potential were set equal to zero ($C = 0.0$ kcal/mol·Å^{12}, $D = 0.0$ kcal/mol·Å^{10}) for the $O_{TIP3P} \cdots H_{TIP3P}$ parameter in order to ensure that water hydrogen bonding was handled as in Jorgensen et al. (1983). Some other nonstandard force field parameters were used in this study and they are listed in Table 6.

The following strategy was used to prepare each model (1WRP:XRAY and 3WRP:XRAY) for the MD simulation: the waters and ions were first minimized and subjected to 10 ps of MD in the presence of rigid protein. This procedure allows for the reorientation of the electric vectors in the solvent/ion system. The protein/water/ion Cartesian coordinates were then minimized for several more sequences. In each sequence, harmonic positional constraints were applied to the backbone atoms (N, Cα, C, O) of the protein and the RMS gradient was reduced to 0.1 kcal/Å with conjugate gradient minimization. Positional constraints of 100, 50, 15, and 2 kcal/Å^2 were used during the minimizations. Finally, all constraints were removed from the system, and the RMS gradient was reduced to 0.05 kcal/Å using conjugate gradient minimization. In this paper, the minimized structures are referred to as 1WRP:MIN and 3WRP:MIN. The 1WRP:MIN structure was calculated to have an energy of $-37{,}017.2$ kcal·mol^{-1}, whereas the molecular mechanics energy of 3WRP:MIN was found to be $-34{,}682.9$ kcal·mol^{-1}. These energies cannot be compared with each other due to the different number of waters in each complex. The MD simulation was then started.

For both models, 1WRP:MIN and 3WRP:MIN, a 35-ps MD trajectory was calculated (1 ps MD ≈ 1 Cray X-MP/48 h) at constant temperature (Berendsen et al., 1984) using the "leapfrog" algorithm (Hockney, 1970) to integrate the equations of motion. The average temperature during the simulations was 300 K for the configuration of atoms. However, local heating effects were evident, and the average temperature of the protein atoms was 235 K, whereas the water atoms had a temperature of approximately 345 K. A time step of 0.0015 ps was used during the MD simulations. A nonbonded cutoff of 8 Å was applied, and the nonbonded list was updated every 10 steps. The SHAKE algorithm (Ryckaert et al., 1977; van Gunsteren & Berendsen, 1977) was used to constrain all bonds. During the course of the trajectories, the Cartesian coordinates were saved to disk every 200 steps.

In our MD analysis, we have calculated the RMS deviations between various structures using the equation:

$$\text{RMS deviation} = \left[\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{(d_{ij}^{ref} - d_{ij})^2}{n^2} \right]^{1/2},$$

where

$$d_{ij} = [(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2]^{1/2}.$$

Although we report the overall RMS deviations between structures, it is often illustrative to compare the deviations between smaller structural units. Thus, most of our analyses report the deviations between C_α atoms for each protein residue. The MD analysis in this paper also reports the RMS atomic fluctuations during the simulations, where the average MD structure is represented by the subscript j.

$$\text{RMS fluctuation} = \langle d_{ij}^2 \rangle^{1/2},$$

where

$$d_{ij} = [(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2]^{1/2}.$$

Circular statistics (Batschelet, 1981) have been used in this paper to analyze protein dihedral angles from the MD trajectories. Using this methodology, a mean angle, $\bar{\phi}$, is defined by

Table 6. *AMBER force field parameters*[a]

Atom	Mass (g/mol)	Description
CT	12.01	All-atom aliphatic carbon
LP	3.00 (12.00)	Sulfur lone pair
S	32.02	Sulfur in MET
SO	22.99	Sodium cation
Angle	**K_θ (kcal/radian)**	**θ_{eq} (degrees)**
CT-S-LP	150.00 (600.00)	96.70
LP-S-LP	0.00 (600.00)	160.00
VDW	***R**** (Å)	***e**** (kcal/mol)
SO	1.60	0.10

[a] This table contains a description of the nonstandard AMBER force field parameters used during this study. The sulfur parameters were revised (G. Seibel, unpubl.) in order to eliminate the erratic behavior of the S-LP atoms during molecular mechanical minimizations and molecular dynamics calculations. The original sulfur parameters, enclosed in parentheses, are included in the table.

$$\bar{\phi} = \begin{cases} \arctan(\bar{y}/\bar{x}) & \text{if } \bar{x} > 0 \\ 180° + \arctan(\bar{y}/\bar{x}) & \text{if } \bar{x} < 0 \\ 90° & \text{if } \bar{x} = 0 \text{ and } \bar{y} > 0 \\ 270° & \text{if } \bar{x} = 0 \text{ and } \bar{y} < 0 \\ \text{undetermined} & \text{if } \bar{x} = 0 \text{ and } \bar{y} = 0, \end{cases}$$

where

$$\bar{x} = \frac{1}{n}(\cos\phi_1 + \cos\phi_2 + \ldots + \cos\phi_n),$$

and

$$\bar{y} = \frac{1}{n}(\sin\phi_1 + \sin\phi_2 + \ldots + \sin\phi_n)$$

for n dihedral angles. The mean angular deviation, s, is a measure of dispersion and this quantity is analogous to the standard deviation in linear statistics. It is defined with respect to r, the mean vector length, and r is a measure of concentration.

$$s(\text{degrees}) = \frac{180°}{\pi}[2(1-r)]^{1/2},$$

where

$$r = [(\bar{x})^2 + (\bar{y})^2]^{1/2}.$$

Note added in proof

After this paper was submitted for publication, the results of an MD simulation of the *trp* aporepressor appeared in the literature (Komeiji et al., 1991). These researchers similarly find the protein to have a relatively rigid core and flexible DNA-binding motif. One interesting difference should be noted between the calculations: the fluctuations of the helix-turn-helix motif were found to be larger during our calculations. The origin of this difference is not clear but it may arise from the larger solvent environment of the Komeiji et al. (1991) calculation.

Acknowledgments

We gratefully acknowledge useful discussions with Paul Sigler (Yale), and we thank both him and his research group for providing us with several *trp* (apo)repressor X-ray structures prior to publication. The authors acknowledge Jeanmarie Guenot (UCSF) for useful discussions. We thank Jeanmarie Guenot (UCSF) for providing the modeled proteins *trp* (apo)repressor and George Seibel (UCSF) for providing revised sulfur lone pair force field parameters. We also thank D. Eisenberg (UCLA) and David Spellmeyer (DuPont) for providing programs to calculate solvent-accessible surface areas (DE) and circular statistical quantities (DS). The molecular dynamic trajectories discussed in this paper were calculated on the Cray X-MP/48 at the San Diego Supercomputer Center (SDSC) and we are grateful to the NSF for supporting the computer usage and to NIH GM-29072 for research support. We are also pleased to acknowledge the use of the UCSF Computer Graphics facilities through grant RR-1014 to R. Langridge.

References

Abola, E.E., Bernstein, F.C., Bryant, S.H., Koetzle, T.F., & Weng, J. (1987). In *Crystallographic Databases—Information Content, Software Systems, Scientific Applications* (Allen, F.H., Bergerhoff, G., & Seivers, R., Eds.), pp. 107–132. Data Commission of the International Union of Crystallography, Bonn, Cambridge, Chester.

Arrowsmith, C.H., Carey, J., Treat-Clemons, L., & Jardetzky, O. (1989). NMR assignments for the amino-terminal residues of *trp* repressor and their role in DNA binding. *Biochemistry 28*, 3875–3879.

Arrowsmith, C.H., Pachter, R., Altman, R.B., Iyer, S.B., & Jardetzky, O. (1990). Sequence-specific 1H NMR assignments and secondary structure in solution of *Escherichia coli trp* repressor. *Biochemistry 29*, 6332–6341.

Batschelet, E. (1981). *Circular Statistics in Biology*. Academic Press, New York.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., & Haak, J.R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys. 81*, 3684–3690.

Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977). The Protein Data Bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol. 112*, 535–542.

Brooks, C.L., III, Karplus, M., & Pettitt, B.M. (1988). *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*. John Wiley & Sons, New York.

Carey, J. (1989). *trp* repressor arms contribute binding energy without occupying unique locations on DNA. *J. Biol. Chem. 264*, 1941–1945.

Guenot, J. & Kollman, P.A. (1991). Molecular dynamics studies of a DNA-binding protein: 2. An evaluation of implicit and explicit solvent models for the molecular dynamics simulation of the *Escherichia coli trp* repressor. *Protein Sci. 1*, 1185–1205.

Gunsalus, R.P. & Yanofsky, C. (1980). Nucleotide sequence and expression of *Escherichia coli trp*R, the structural gene for the *trp* aporepressor. *Proc. Natl. Acad. Sci. USA 77*, 7117–7121.

Hockney, R.W. (1970). *Methods Comput. Phys. 9*, 136–211.

IUPAC-IUB Commission on Biochemical Nomenclature 1969. IUPAC-IUB Commission on Biochemical Nomenclature. (1970). Abbreviations and symbols for the description of the conformation of polypeptide chains. Tentative rules (1969). *Biochemistry 9*, 3471–3479.

Joachimiak, A., Marmorstein, R.Q., Schevitz, R.W., Mandecki, W., Fox, J.L., & Sigler, P.B. (1987). Crystals of the *trp* repressor–operator complex suitable for X-ray diffraction analysis. *J. Biol. Chem. 262*, 4917–4921.

Joachimiak, A., Schevitz, R.W., Kelly, R.L., Yanofsky, C., & Sigler, P.B. (1983). Functional inferences from crystals of *Escherichia coli trp* repressor. *J. Biol. Chem. 258*, 12641–12643.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., & Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926–935.

Klig, L.S., Carey, J., & Yanofsky, C. (1988). *trp* repressor interactions with the *aro*H and *trp*R operators. Comparison of repressor binding in vitro and repression in vivo. *J. Mol. Biol. 202*, 769–777.

Komeiji, Y., Uebayasi, M., Someya, J., & Yamato, I. (1991). Molecular dynamics simulation of trp-aporepressor in a solvent. *Protein Eng. 4*, 871–875.

Lawson, C.L. & Sigler, P.B. (1988). The structure of *trp* pseudorepressor at 1.65 Å shows why indole propionate acts as a *trp* 'inducer.' *Nature 333*, 869–871.

Lawson, C.L., Zhang, R.G., Schevitz, R.W., Otwinowski, Z., Joachimiak, A., & Sigler, P.B. (1988). Flexibility of the DNA-binding domains of *trp* repressor. *Proteins 3*, 18–31.

Levitt, M. & Sharon, R. (1988). Accurate simulation of protein dynamics in solution. *Proc. Natl. Acad. Sci. USA 85*, 7557–7561.

Lolis, E., Alber, T., Davenport, R.C., Rose, D., Hartman, F.C., & Petsko, G.A. (1990). Structure of yeast triosephosphate isomerase at 1.9-Å resolution. *Biochemistry 29*, 6609–6618.

Lolis, E. & Petsko, G.A. (1990). Crystallographic analysis of the complex between triosephosphate isomerase and 2-phosphoglycolate at 2.5-Å resolution: Implications for catalysis. *Biochemistry 29*, 6619–6625.

Marmorstein, R.Q. & Sigler, P.B. (1989). Stereochemical effects of L-tryptophan and its analogues on *trp* repressor's affinity for operator-DNA. *J. Biol. Chem. 264*, 9149–9154.

Otwinowski, Z., Schevitz, R.W., Zhang, R.-G., Lawson, C.L., Joachimiak, A., Marmorstein, R.Q., Luisi, B.F., & Sigler, P.B. (1988). Crystal structure of *trp* repressor/operator complex at atomic resolution. *Nature 335*, 321–329 (published erratum: *Nature 335*, 8371).

Perutz, M.F. (1989). Mechanisms of cooperativity and allosteric regulation in proteins. *Q. Rev. Biophys. 22*, 139–237.

Rose, J.K., Squires, C.L., Yanofsky, C., Yang, H.L., & Zubay, G. (1973). Regulation of *in vitro* transcription of the tryptophan operon by purified RNA polymerase in the presence of partially purified repressor and tryptophan. *Nature 245*, 133–137.

Ryckaert, J.P., Ciccotti, G., & Berendsen, H.J.C. (1977). Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. *J. Comput. Phys. 23*, 327–341.

Schevitz, R.W., Otwinowski, Z., Joachimiak, A., Lawson, C.L., & Sigler, P.B. (1985). The three-dimensional structure of *trp* repressor. *Nature 317*, 782–786.

Singh, U.C. & Kollman, P.A. (1984). An approach to computing electrostatic charges for molecules. *J. Comput. Chem. 5*, 129–145.

Singh, U.C., Weiner, P.K., Caldwell, J.W., & Kollman, P.A. (1986). *AMBER (UCSF Version 3.0)*, Department of Pharmaceutical Chemistry, University of California, San Francisco, California.

van Gunsteren, W.F. & Berendsen, H.J.C. (1977). Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys. 34*, 1311–1327.

Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S., Jr., & Weiner, P. (1984). A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc. 106*, 765–783.

Weiner, S.J., Kollman, P.A., Nguyen, D.T., & Case, D.A. (1986). An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem. 7*, 230–252.

Zhang, R.-G., Joachimiak, A., Lawson, C.L., Schevitz, R.W., Otwinowski, Z., & Sigler, P.B. (1987). The crystal structure of *trp* aporepressor at 1.8 Å shows how binding tryptophan enhances DNA affinity. *Nature 327*, 591–597.

Zurawski, G., Gunsalus, R.P., Brown, K.D., & Yanofsky, C. (1981). Structure and regulation of *aro*H, the structural gene for the tryptophan-repressible 3-deoxy-D-arabino-heptulosonic acid-7-phosphate synthetase of *Escherichia coli*. *J. Mol. Biol. 145*, 47–73.