Effect of Zn^{2+} on DNA Recognition and Stability of the p53 DNA-Binding Domain†

Jianxin Duan and Lennart Nilsson*

Department of Biosciences and Nutrition, Karolinska Institutet SE-141 57 Huddinge, Sweden

Received February 15, 2006; Revised Manuscript Received April 7, 2006

ABSTRACT: Zinc ions are frequently found in DNA-binding proteins. p53 is a cancer-related transcriptional factor, and its DNA-binding domain (DBD) contains a Zn^{2+}, which has been shown to be important for aggregation and sequence-specific DNA binding. We have carried out molecular dynamics simulations to investigate the influence of Zn^{2+} on the p53 DNA recognition and the stability of the DBD. In the simulation with Zn^{2+} present, the protein attracted to the DNA phosphate backbone, allowing for Arg248 on loop L3 to be inserted into the minor groove for specific contact with the DNA base. The insertion of Arg248 between the backbone phosphate groups in the minor groove caused a narrowing of the minor groove, which is not seen in the simulation without Zn^{2+}. Structurally, the zinc ion coordinated the motions among the different protein structural elements, which could also be important for optimal binding and core packing. The influence of Zn^{2+} on protein stability was mainly localized to the L2 loop. Our results suggest that L2 may be a frustrated and highly flexible element and play an important role in aggregation of Zn-free p53. Zn^{2+} keeps the L2 structured and probably prevents aggregation.

The tumor-suppressor protein p53 is expressed in its latent form in all cells. It responds to DNA damage, transactivates its target genes, and initiates cell-cycle arrest and/or apoptosis. When p53 interacts with other proteins, p53 is also capable of influencing replication and transcription. Further, p53 has been demonstrated to directly bind to damaged DNA sites and exhibit 3′ → 5′ exonuclease activity. When p53 enhances DNA repair and prevents proliferation of cells with damaged genomes, it contributes to DNA integrity and chromosomal stability. Mutations in p53 are associated with over 50% of human mutations found in tumors (*1*−*7*), and rescuing p53 is an important cancer therapeutic strategy (*8*).

The p53 protein is a flexible multidomain protein containing 393 residues, and it forms a tetramer when active. The N terminus of the protein is the transcriptional activation domain capable of interacting with a number of cellular and viral proteins. The tetramerization domain, indispensable for tetramerization, is between residues 324 and 355. The C terminus recognizes damaged DNA and has some DNA repair activities, and it also has been shown to have negative regulatory functions. The central part of the protein is the sequence-specific DNA-binding domain (DBD)[1] that binds to the consensus 5′-PuPuPuC(A/T) sequence (*4*, *6*, *7*). The p53-binding sites contain four consensus sequences, oriented in → ←-X-→ ← fashion, where the arrows represent the consensus sequences and X represents variable spacing. Each → ← (Figure 1C) is a half-site and binds a p53 dimer. Most of the cancer-related mutations are mapped to the DBD.

The crystal structure of p53 DBD revealed a large β-sandwich scaffold, positioning the DNA-binding loop−sheet−helix motif in the major groove and loop L3 in the minor groove (*9*) (Figure 1). A single Zn^{2+} is coordinated by Cys176, Cys238, Cys242, and His179 connecting the L2 and L3 loops. Removal of Zn^{2+} reduces the DNA-binding specificity (*10*−*13*). It has also been shown that Zn^{2+}-free DBD (apoDBD) is structurally different from DBD (*10*, *12*−*14*) and is aggregation-prone (*13*). Butler and Loh studied the structure and function of apoDBD using nuclear magnetic resonance (NMR) (*13*). Although they were not able to acquire a three-dimensional structure of the apoDBD, they have shown that most of the differences are located in the L2 and L3 loops and concluded that the contacts made by L3 in the minor groove are crucial for specific binding because Zn^{2+} containing DBD has an 11-fold higher binding affinity for the specific sequence, even though nonspecific binding to DNA is retained. They have also shown that Zn^{2+} dissociates at physiological temperature and the most common mutations do not appreciably affect the stability of apoDBD (*13*).

In this paper, molecular dynamics simulations were performed to investigate the detailed structure of the apoDBD and the Zn^{2+}-free DBD−DNA complex (apocomplex) in comparison to the DBD and DBD−DNA complex (complex) structure. We examined the influence of Zn^{2+} on specific sequence recognition, DNA binding, and protein dynamics. We found that Zn^{2+} not only participates actively in DNA binding but also coordinates movements of different structural elements in the protein.

† This work is supported by grants from The Swedish Research Council.

* To whom correspondence should be addressed: Department of Biosciences and Nutrition, Center of Structural Biochemistry, SE-141 57 Huddinge, Sweden. Telephone: +46-8-608-9228. Fax: +46-8-608-9290. E-mail: lennart.nilsson@biosci.ki.se.

[1] Abbreviations: DBD, DNA-binding domain; rmsd, root-mean-square deviation; rmsf, root-mean-square fluctuation.

10.1021/bi0603165 CCC: $33.50
Published on Web 05/23/2006

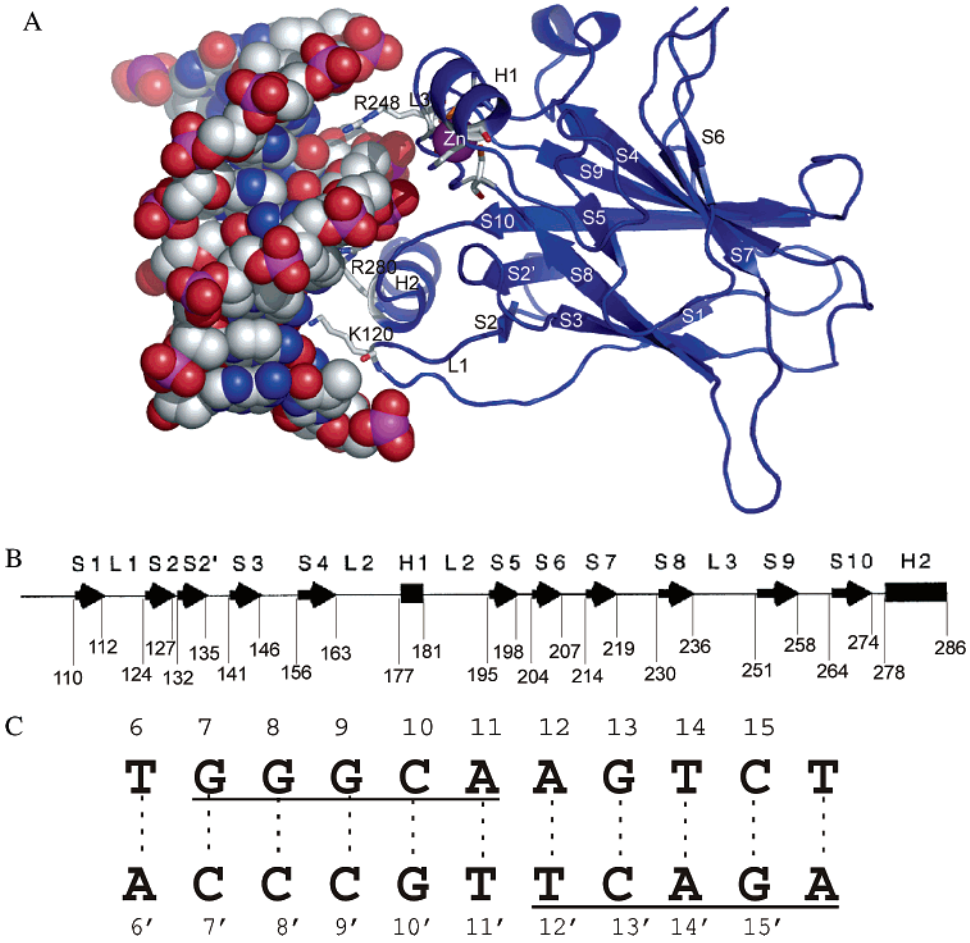

FIGURE 1: (A) Cartoon of the p53 DBD−DNA complex. The DNA molecule is shown as spheres; DBD is in a cartoon representation; and the Zn^{2+}-coordinating residues are shown as sticks. Zn^{2+} and all structural elements are marked, except L2, the loop between S4 and S5 and with H1 in the middle. (B) Secondary-structure line-up. The numbers are the residue number at which the secondary structure starts or ends. (C) p53-binding half-site. The consensus sequences are underlined, and the sequence used in this paper is numbered.

MATERIALS AND METHODS

System Setup. The monomeric p53 DBD bound to its consensus DNA sequence was extracted from the 2.2 Å crystal structure solved by Cho et al. (PDB entry 1TSR) (*9*). The DNA sequence is TGGGCAAGTC, with the consensus half-site sequence underlined (*9*). Four systems were set up: the DBD−DNA complex with and without Zn^{2+} (complex and apocomplex) and the DBD itself with and without Zn^{2+} (DBD and apoDBD) using CHARMM (*15*). The all-atom force-field parameters used were CHARMM22 for proteins (*16*) with modified parameters for the zinc−methiolate interaction potential (*17*) using the formal Zn^{2+}-ion model (*18*) and CHARMM27 for nucleic acids (*19*). To reproduce full flexibility, Zn^{2+} is not covalently bound to the protein. The hydrogen atoms were added with HBUILD (*20*). Crystallographic water was added to the system prior to immersing in a TIP3P (*21*) water sphere with a 41 Å radius with a minimum of 10 Å from the water boundary to any protein or DNA atoms. All water molecules with the oxygen atom closer than 2.8 Å to the solutes were deleted. The systems were made overall neutral by adding either sodium or chloride ions.

The solvent was first minimized with 500 steps steepest descent (SD) with a fixed solute, followed by another 500 steps SD minimization, with 20.0 kcal mol^{-1} $Å^{-2}$ harmonic restraint on the backbone atoms of the protein and DNA. The restraint was then decreased to 10.0 kcal mol^{-1} $Å^{-2}$ for another 500 steps adopted basis Newton−Rapson (ABNR) minimization. In the final minimization with 500 steps ABNR, the restraint was removed. Finally, the water molecules were allowed to equilibrate around the fixed solutes in a 10 ps molecular dynamics simulation followed by a 500 steps SD minimization.

Simulation Protocol. The simulations were performed using CHARMM (*15*) with stochastic boundary conditions (*22*) at 300 K. Water molecules in the outer 4 Å solvation shell were treated as Langevin particles with a friction constant β of 50 ps^{-1} on all water oxygen atoms. The temperature was maintained at 300 K with a stochastic heat bath. All covalent bonds to hydrogen atoms are constrained with the SHAKE algorithm (*20*), permitting a time step of 2 fs. To prevent the solutes from migrating toward the boundary sphere, a harmonic restraint with a force constant of 10 kcal mol^{-1} $Å^{-2}$ was applied on one solute atom close to the center of the sphere. The Watson−Crick hydrogen bonds between the outer DNA base pairs were harmonically restrained to their initial length with a force constant of 1.0 kcal mol^{-1} $Å^{-2}$. All electrostatic interaction forces and energies were smoothly shifted to 0 at 12 Å by an atom-based cutoff (CHARMM FSHIFT option). The nonbonded lists were built with a 14 Å cutoff and updated whenever an atom moved more than 1 Å since the last update. It has been

shown that this is a spherical cutoff scheme that performs well and is comparable to PME (*23*−*26*). Coordinates were saved every 0.5 ps, and the simulations lasted 10 ns each.

Analysis. All references to numbering of nucleotides and secondary structures are according to Cho et al. (*9*). Briefly, the DNA starts at T6 (5′) and ends at C16 (3′). The complementary strand is marked with a prime sign, and the numbering still follows the parent strand (Figure 1C). For secondary structures, S stands for strand, H stands for helix, and L stands for loop (Figure 1B).

Hydrogen bonds are defined if a hydrogen-donor atom is <2.4 Å from an acceptor atom, and the lifetime was analyzed with a 2 ps time window. When analyzing structures with explicit hydrogen atoms, a distance cutoff has been shown to be sufficient (*27*).

To analyze the number of water molecules in the space between Zn^{2+} and the phosphate group of T14, a dummy atom was placed every 1 Å along the vector between the ion and phosphorus atom for snapshots taken every 2 ps. The number of water molecules within a radius of 1 Å from the dummy atoms were counted and averaged over all snapshots. Essentially, a corridor of 2 Å width was created, and water molecules that resided in this corridor were counted. It should be noted that the result is an overestimate because water molecules may be counted twice if they are in the overlapping volumes; however, not all volumes were overlapping.

The width of the DNA minor groove was calculated by Curves5.1 (*28*) for snapshots every 10 ps for the last 8 ns.

The Poisson−Boltzmann equation was solved at 300 K with the software package MOE (Chemical Computing Group, Inc.). The partial charges were according to CHARMM22 all-atom force field for proteins (*16*). The Gaussian charge densities are derived from charge deviation by setting the standard deviation of the Gaussian to be R/D, where R is the van der Waals radius and D is the charge deviation, which is set to 2.5 in the calculations. If the D is 0, then the charges are modeled with point charges. The protein interior dielectric constant was set to 4, and the exterior dielectric constant was set to 80. The dielectric offset, defined as the distance away from the solute surface at which the dielectric becomes that of the solvent, was set to 1.4 Å, as was the solvent radius. The solute concentration was chosen to be 1 mM, and no ion was present in the calculation. The grid spacing was set to 0.925 Å. The molecular surfaces were calculated with a probe of 1.4 Å radius for the protein, and the electrostatic potentials were mapped on the surfaces.

The spatial atom displacement covariance was calculated for Cα atoms and phosphorus atoms along the simulation trajectories as follows:

$$\sigma_{ij} = \langle (x_i - \langle x_j \rangle)(x_j - \langle x_j \rangle) \rangle$$

and the normalized covariance matrix is given by

$$\mathrm{CO}_{ij} = \sigma_{ij}/\sqrt{\sigma_{ii}\sigma_{jj}}$$

The secondary-structure definition was calculated using STRIDE (*29*), implemented in VMD1.8.2 (*30*). The averaged structures were calculated on the basis of the last 8 ns of the trajectory, allowing 2 ns for equilibration.

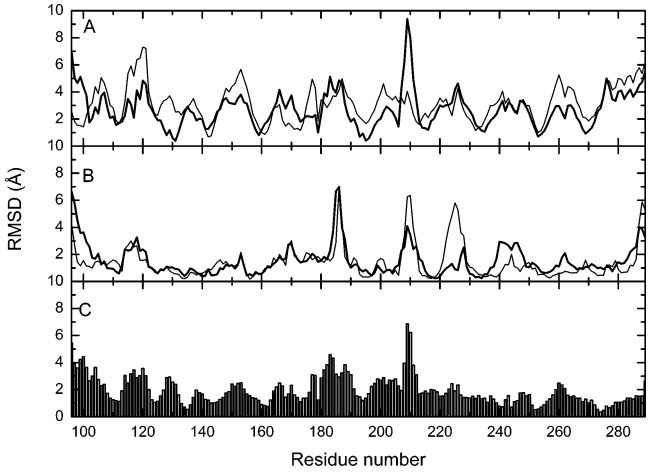

FIGURE 2: Protein backbone rmsd per residue of the averaged structures of (A) the DBD (thick line) and apoDBD (thin line) simulations compared to the starting structure, (B) the complex (thick line) and apocomplex (thin line) simulations compared to the starting structure, and (C) the averaged structure of DBD compared to apoDBD.

Table 1: Pairwise Comparison of the Backbone rmsd of the Averaged Structures (Å)

	start	DBD	apoDBD	complex	apocomplex
start	0				
DBD	3.04	0			
apoDBD	3.44	2.23	0		
complex	1.8	3.19	3.76	0	
apocomplex	1.91	2.98	3.52	1.48	0

RESULTS

Protein Structure and Dynamics. The pairwise backbone root-mean-square deviation (rmsd) between structures averaged over the last 8 ns of the trajectories and the starting structure (Table 1) shows that complex and apocomplex are structurally close, while the difference between DBD and apoDBD is much greater. From the rmsd per residue compared to the starting structure (Figure 2), a correlation between increased rmsd and removal of either Zn^{2+} or DNA can be noticed. The root-mean-square fluctuation (rmsf) per residue values around the averaged structures (Figure 3) were of similar magnitude in all cases but with clear differences; in particular, the rmsf of L2 loop in apoDBD is much higher than in the apocomplex.

The covariance plots between the different structural elements are presented in Figure 4, and a list of the correlating secondary structures are shown in Table 2. In the protein DNA interface, L1 is correlated with H2 but not with L3, because L1 and H2 interact with the major groove, while L3 interacts with the minor groove. On the other hand, L3 is correlated with H1 through Zn^{2+}. Aside from the correlations in the DNA- and Zn^{2+}-binding regions, there were also correlations between sequentially distant but spatially close structural elements. For example, motions in S4 correlated to S7, S9, and S10 because they belong to the same β sheet, and S1 correlated to S3−S4, S9−S10, and S10 because they are located in the far end of the β sandwich. Curiously, there is a weak correlation between S1 and H1, which are on opposite ends of the sandwich.

Strikingly, the covariance decreased significantly when DNA, Zn^{2+}, or both are removed from the system. Along

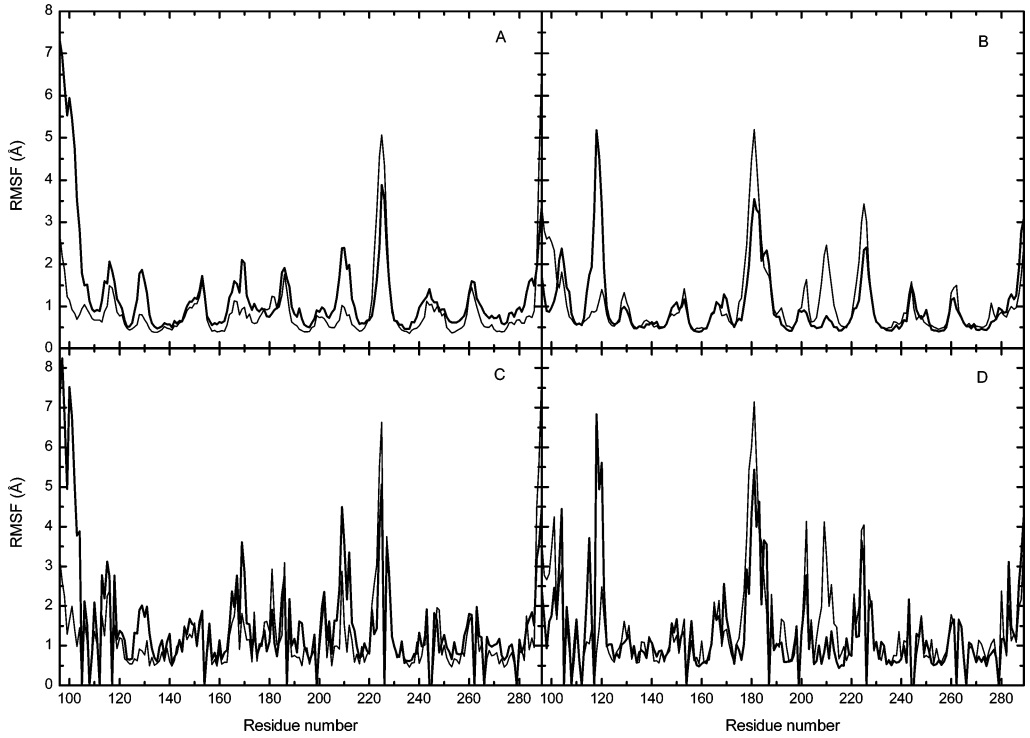

FIGURE 3: Backbone (A and B) and side-chain (C and D) rmsf per residue with the averaged structures as a reference. The data for the complex (thick lines) and apocomplex (thin lines) simulations are collected in A and C, while the data for DBD (thick lines) and apoDBD (thin lines) are in B and D.

Table 2: Structural Elements Showing High Covariance with Each Other (>0.6)[a]

	complex	apocomplex	DBD	apoDBD
S1	S3−S4, H1, S9-**S10**, **S10**	S3−S4, **S10**		
L1	**H2**	**H2**		
S2−S2′	S2′−S3, **H2**	S2′−S3		
S2′	**S10**−**H2**		**S10**−**H2**	**S10**−**H2**
S3	S8	S8		S8
S4	S7, S9, **S10**	S7	S7, S9	S9
H1	S5, **L3**			
S5	**L3**, S8, S9			S8
S5−S6	S7		S7	
S6	S6−S7			
S9	S9-**S10**, **S10**	**S10**	**S10**	S9-**S10**, **S10**

[a] The secondary-structure elements involved in DNA binding are highlighted in bold.

Table 3: (A) Occupancy of Specific Protein DNA Hydrogen Bonds and (B) Sum of the Occupancies for the Hydrogen Bonds between the Protein and DNA

A			occupancy	
protein		DNA	complex	apocomplex
K120	Hζ	G8 O6	0.28	0.12
		G8 N7	0.86	0.86
		G9 O6	0.25	0.40
		G9 N7	0.08	0.22
R248	Hη12	T12′ O2	0.34	0
R280	Hη	G10′ O6	0.51	0.52

B	to base	to phosphate	sum
complex	2.5	6.4	8.9
apocomplex	2.3	6.6	8.9

with many correlations involving the DNA- and Zn^{2+}-binding elements, much of the covariance in the β sandwich also vanished. It may appear that apoDBD retained more covariance patterns than DBD; however, the peaks in both plots are weak, and the difference may be insignificant.

Protein DNA Interactions. Further, we examined the contact pattern between the DBD and DNA in the complexes. The binding surfaces of the DBD and DNA complement each other very well, leaving basically no space for interfacial water molecules. Indeed, the analysis for water-mediated hydrogen bonds between DBD and DNA revealed that no such bridges were present during more than 6% of the simulation time. The sequence-specific contacts in the major grove were hydrogen bonds between Lys120 H_ζ atoms and O6 and N7 atoms on bases G8 and G9 and Arg280 $H_\eta 11$ and $H_\eta 12$ atoms and G10′ O6 atoms in both the complex and apocomplex (Table 3A). The hydrogen bond between Cys277 and C9′ described in the crystal structure (*9*) was however not observed in our simulations. In the minor groove, the specific contact between Arg248 $H_\eta 12$ on L3 and T12′ O2 was only found in the complex system and not in the apocomplex system (Table 3A). Interestingly T12′ does not belong to the 5′-PuPuPuC(A/T) sequence but to the next reversed consensus sequence (Figure 1C). The nonspecific hydrogen bonds were made by residues in loop 3, strand S10, and helix H2. Summing up the hydrogen-bond occupancies (Table 3B) for both systems reveals that the sum of specific and nonspecific contacts are the same, but the specific contact occupancy is 0.18 higher for the complex system. The behavior of Arg248 in the minor groove in the complex simulation confirms the critical importance of L3 for site-specific DNA binding (*9*).

The covariance of the DNA backbone phosphorus atoms is similar in the complex and apocomplex simulations (parts E and F of Figure 4). Nucleotides 11−13 on one chain and 10′ on the complementary chain are correlated in both simulations, while the unexpected correlation between nucle-

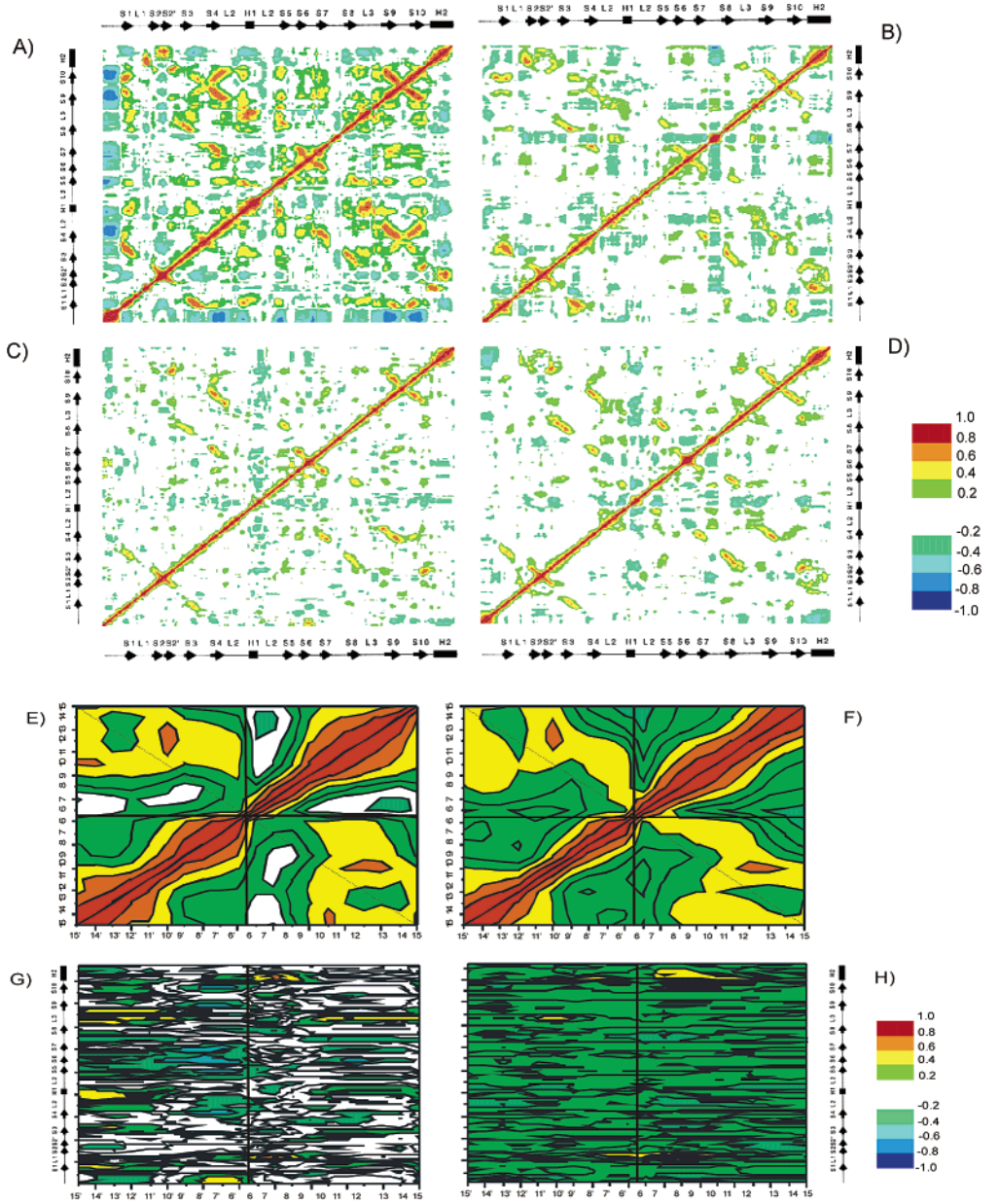

FIGURE 4: Covariances of the spatial atom displacements for Cα atoms of the protein and phosphorus atoms of the DNA backbone along the trajectories. The first four panels show the covariance of the protein Cα atoms: A shows complex simulation; B shows apocomplex; C shows DBD; and finally D shows apoDBD. E and F are the covariance plots of the DNA backbone phosphorus atom for complex and apocomplex simulations. A weak correlation between the paired DNA strands is highlighted by the diagonal thin dashed line. G and H show the covariance between the protein and DNA backbone atoms for complex and apocomplex simulations.

otides 10 and 14′−15′ was only found in the complex simulation. This second correlation is unexpected because of the spatial distance between nucleotide 10 in the middle of the DNA and nucleotides 14′−15′ in the end. However, as described below, the DNA in the complex simulation was wrapped around the L3 loop, which could be related to the observed correlation.

A more striking difference was evident for the covariance between the protein and DNA backbone atoms in the apocomplex. Regions with low correlation ($-0.2 < CO_{ij} < 0.2$) in the complex exhibit a slight increase in correlation in the apocomplex. Generally, the correlation between the protein and DNA backbone atoms in the apocomplex simulation is more uniform and lower than in the complex simulation. In both simulations, high correlations can be found between bases 6−9 and H2/L1, identified as major-groove contacts. One patch of high covariance between bases 9′−10′ and H2 is found in the apocomplex (not seen in the complex simulation) corresponding to the major-groove contact on the complementary strand. In the complex simulation, there is additionally a large region, between L3 and bases 10′−15′, with high covariance and symmetrically, and correlations between bases 10−15 (complementary bases of 10′−15′) and the L3 loop are also fairly high. These are the minor-groove contacts, and to our expectation, these correlations are not found in the apocomplex. In conclusion, the covariance data implies that the presence of Zn^{2+} is crucial for sustaining minor-groove contacts.

Zn^{2+} and DNA Binding. Apart from the hydrogen bond between Arg248 and T12′ (Table 3A), Arg248 formed salt bridges with both T11′ and T14 phosphate backbones in the complex system. Neither this specific hydrogen bond nor

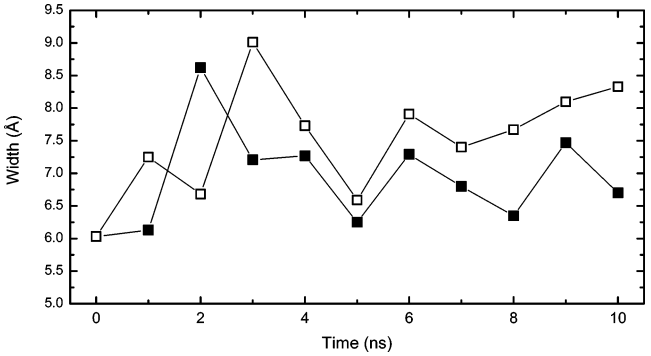

FIGURE 5: DNA minor-groove width at T12′ as calculated using Curves 5.1 (*28*) for the complex (■) and apocomplex (□) simulations.

the bifurcated salt bridges were seen in the apocomplex simulation. Instead, the Arg248 formed salt bridges with T14 and C15 phosphates and no hydrogen-bonding contact with the bases in the apocomplex system. The crystal structure suggests that the insertion of Arg248 into the minor groove leads to salt bridges with the phosphates on both DNA chains, which in turn causes a narrowing of the minor groove (*9*). A statistically significant ($p < 0.001$, independent Student's *t* test) narrowing of the minor groove was observed (Figure 5) in the complex system (6.7 ± 0.6 Å averaged over the last 8 ns) compared to the apocomplex (7.7 ± 0.9 Å averaged over the last 8 ns). The insertion of Arg248 into the minor groove in the complex simulation is unambiguously manifested by the hydrogen bond between Arg248 and T12′ on the minor-groove side. To accommodate such an insertion, the L3 loop must be close to the DNA. Therefore, the distances between the Arg248 Cα atom and T14 and T11′ phosphorus atoms were monitored, and we found that the distances were larger (>2 Å) in the apocomplex in both cases (parts A and B of Figure 6). The interaction energies between Zn^{2+} and phosphate groups above and below Arg248 (i.e., T11′ and T14) were clearly favorable despite the long distance (parts C and D of Figure 6). Figure 7 shows that L3 and Arg248 are in better contact with the DNA during the course of the complex simulation, as compared to the apocomplex case.

To calculate the electrostatic potential experienced by the DNA in close vicinity of the protein, the Poisson−Boltzmann equation was solved for p53 DBD with and without Zn^{2+} and the electrostatic potentials were mapped to the molecular surface of the solutes (Figure 8). The figure shows the DNA-contacting surface, and the electrostatic potential around L2, L3, and H2 is more positive in DBD because of the presence of the zinc ion. At the position of the T14 phosphorus atom, the grid values for the electrostatic potential energy contributions from the p53 DBD with or without Zn^{2+} were 2.9 and 2.2 kcal, respectively.

Zn^{2+}-Binding Site. Zn^{2+} is bound by cysteines 176, 238, and 242 and His179 and it keeps the L2 and the DNA-binding L3 loops together. The distances (Figure 9) from His179 $N_\delta 1$ to Zn^{2+} and S_γ atoms of the cysteines show that the tetrahedral structure around Zn^{2+} was maintained in the complex simulation but not in the DBD simulation, despite the fact that both systems contained Zn^{2+}. The histidine in DBD lost contact with Zn^{2+} and was replaced by a water molecule after almost 5 ns of simulation, but the cysteines remained bound to Zn^{2+}.

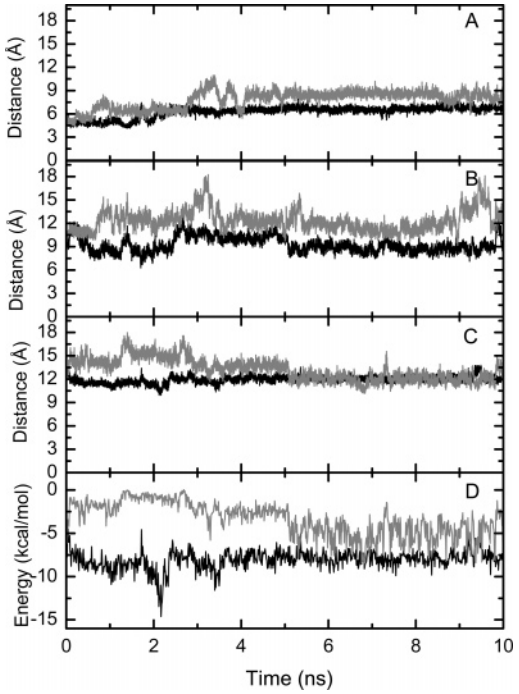

FIGURE 6: Protein−DNA distances and interaction energies. (A) Arg248 Cα−T11′ P distance (black, complex; gray, apocomplex). (B) Arg248 Cα−T14 P distance (black, complex; gray, apocomplex). (C) black, Zn^{2+}−T11′ P distance; gray, Zn^{2+}−T14 P distance. (D) black, Zn^{2+}−T11′ phosphate interaction energy; gray, Zn^{2+}−T14 phosphate interaction energy.

In the apoDBD and apocomplex simulations, the tetrahedral coordination structure was lost already at a very early stage; however, for the apocomplex simulation, the shift in the Zn^{2+}-binding site is less pronounced than in apoDBD, as indicated by the distance mapping between His179 and the cysteines (Figure 9). The rmsd plots for the averaged structures (Figure 2) show that apoDBD has larger shifts in both the L2 and L3 region than the apocomplex. Further, the rmsf plots point out that the fluctuation of the L2 loop in apoDBD is much higher than in the apocomplex, while the L3 loop remained rather stable (Figure 3). While the L2 loop adopted a different structure, in both DBD and apoDBD, in apoDBD, however, L2 has a more open conformation (Figure 10). Helix H1 in the middle of L2 loop spontaneously unfolded in all simulations, except the complex simulation. This suggests that H1 is intrinsically unstable, and indeed, a series of secondary-structure predictions for the L2 region [PHD (*31*), GOR4 (*32*), JUFO (*33*), JPRED (*34*), nnPredict (*35*), SAM-T99 (*36*), SSpro (*37*) and SSpro8 (*37*)], all predicted H1 to be random-coil and in the case of PHD with a reliability score of 8−9 (9 is the maximum). Another curious result from the secondary-structure predictions is that the sequence QHMTEVVRR (residues 167−175) directly upstream from H1 had mixed predictions of both the helix and strand. PHD, nnPredict, SAMT-99, and SSpro all predicted this peptide to be helical, while GOR4, JPRED (a consensus method), and SSpro8 predicted a β structure and JUFO simply predicted a mixture of both.

DISCUSSION

Zinc ions are frequently found in proteins, where they contribute to protein stability and catalytic functions more widely than any other transition metal (*38*, *39*). Zinc ions

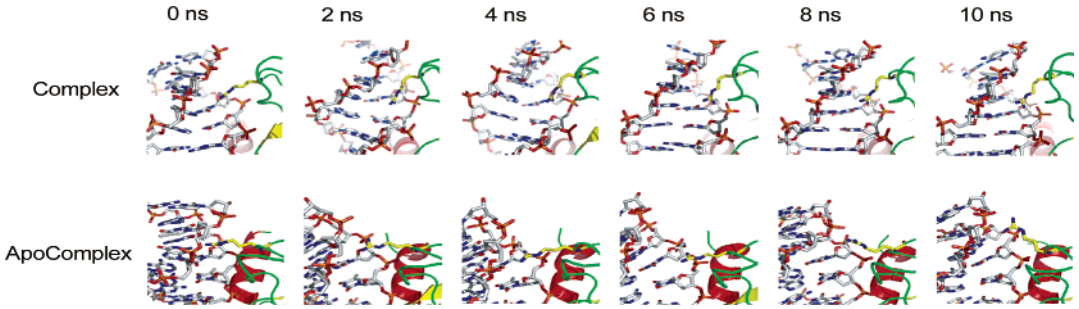

FIGURE 7: Snapshots taken every 2 ns for complex amd apocomplex simulations. The protein is shown in a ribbon representation, and the DNA is shown in sticks. Arg248 is shown in yellow sticks.

are also found in DNA-binding proteins, typically the zinc-finger motif. In all cases, Zn^{2+} is believed to contribute to protein structural integrity without interacting with the DNA (*40*). The binding of Zn^{2+} to p53 has been shown to be crucial for sequence-specific DNA binding (*10*−*12*), providing an 11-fold higher affinity to consensus DNA for DBD compared to apoDBD, while there is no difference when binding to generic DNA sequences (*13*).

Impact of Zn^{2+} on Protein Structure and Dynamics. The simulation of DBD suggested that the Zn^{2+}-binding site is not a stable structure. The His179 in H1 dissociated from Zn^{2+} accompanied by unfolding of H1 at the same time. The opening of L2 and unfolding of H1 leaves a possible pathway for spontaneous Zn^{2+} release in agreement with experiments showing that Zn^{2+} spontaneously dissociates from the DBD at 37 °C (*13*).

In general, apoDBD retained the same structure as DBD, with the L2 loop containing helix H1 as one of the regions with the highest structural shift and fluctuation in apoDBD compared to DBD (Figures 2B, 3, and 10). This observation hints at the importance of L2 in p53. ApoDBD has been found to readily aggregate at 37 °C and also to promote aggregation of DBD (*13*). L2 may play a nonignorable role in the aggregation process. A popular notion concerning aggregation is that regions susceptible to fluctuations may give rise to aggregation-prone structures and that aggregation requires partial unfolding of the native state (*41*). Frustrated secondary-structure elements, defined by the incompatibility of secondary-structure predictions and experimentally determined structures (*41*), have been suggested to be harbingers for polymerization. A score $S_{\alpha/\beta}$ for finding a predicted β strand in a helix determined by experimental methods can be defined as

$$S_{\alpha/\beta} = \frac{1}{L}\sum_{i=1}^{L}(R_i - 5)$$

where R_i is the reliability score predicted by PHD at position i, 5 is the mean score, and L is the sequence length (*41*). A score of 4 indicates maximum frustration (*41*). L2 seems to fit with this concept because (1) H1 located in L2 was predicted to be a coil and a corresponding $S_{\alpha/\text{coil}}$ score is 3.5, thus conforming with the definition of a frustrated secondary-structure element and (2) L2 is highly flexible in apoDBD. Curiously, the peptide QHMTEVVRR in L2 upstream from H1 was predicted to be both a helix and strand by different algorithms indicating the presence of an additional frustrated structural element in L2. Considering partial unfolding and frustration together with the high fluctuation in L2, we conjecture that L2 may be the nucleus for aggregation of apoDBD.

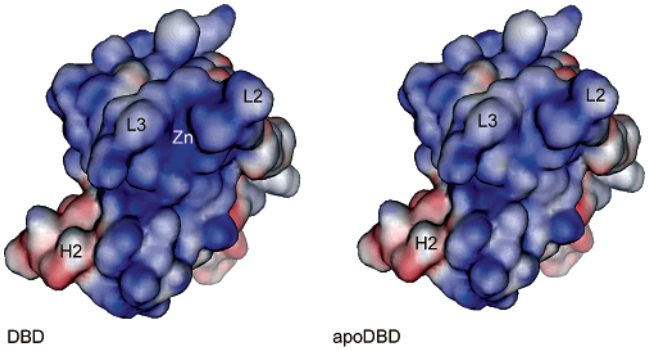

FIGURE 8: Electrostatic potential mapped onto the molecular surface of DBD and apoDBD starting structures. The L2, L3, and H2 are highlighted as well as the zinc. The color code ranges from blue for 30 kcal/mol to red for −30 kcal/mol in electrostatic potential.

Several studies have suggested that there is a core−core interaction between DBDs when bound to DNA and that the interface region was proposed to be around H1 in the L2 loop (*42*−*44*). In a model of full-length p53 dimer based on NMR constraints (*45*), four key residues are identified as being involved in core−core interactions and three of these, Glu171, Val172, and Arg 174, are found in the QHMT**EV**-V**RR** sequence. The role of L2 as a contact interface and its possible propensity to aggregate suggest that the extremely flexible L2 loop in apoDBD may form a dimer with the L2 in DBD and promote heteroaggregation. The H1 structure was preserved during the whole simulation in the complex simulation but not in the DBD simulation. Thus, DNA also has a strong stabilizing effect on H1 and L2, most likely asserted via loop L3 and Zn^{2+}.

It is also interesting to notice the impact of DNA on the protein structure. The difference between the complex and apocomplex is less dramatic than that between the complex and DBD because the backbone rmsd between the complex and apocomplex averaged structures is only 1.5 Å, whereas between the complex and DBD, the rmsd is as high as 3.2 Å. Even the backbone rmsd between the averaged structures of DBD and apoDBD is only 2.2 Å. The covariance plots (Figure 4) of the apocomplex and DBD revealed a similar pattern. Further, the rmsd and rmsf plots (Figures 2 and 3) show that removing DNA from the complex caused a more dramatic change than removing Zn^{2+}. As mentioned earlier, Zn^{2+} has the strongest influence on the stability of the L2 region; thus, DNA seems to have a stronger impact on the protein global stability, whereas the effect of Zn^{2+} is more local, although not less important.

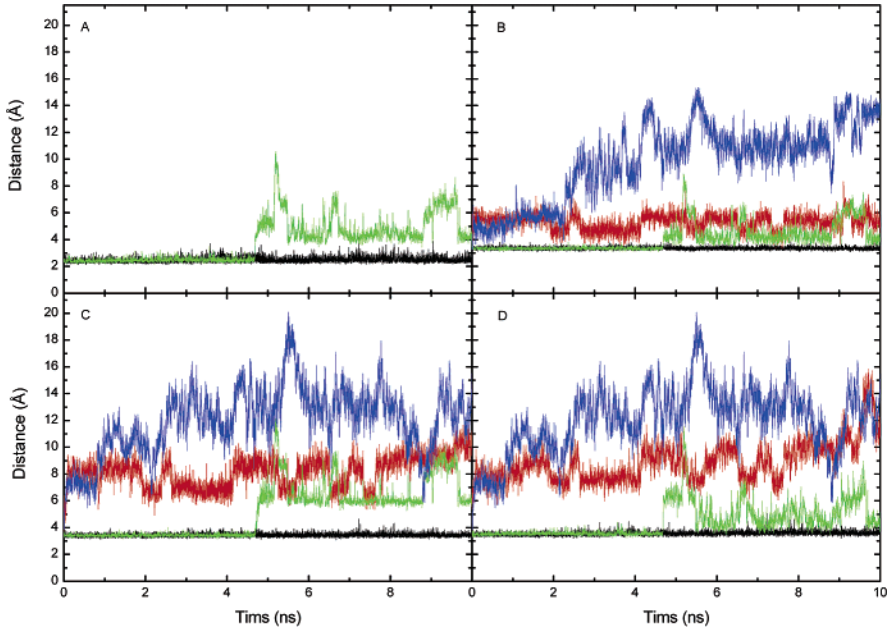

FIGURE 9: (A) Distances between Zn^{2+} and the His179 Nδ1 atom in the complex (black line) and DBD (green line) simulations. Distances between the His179 Nδ1 atom and Sγ atoms in (B) Cys176, (C) Cys238, and (D) Cys242 in the complex (black), apocomplex (red), DBD (green), and apoDBD (blue) simulations.

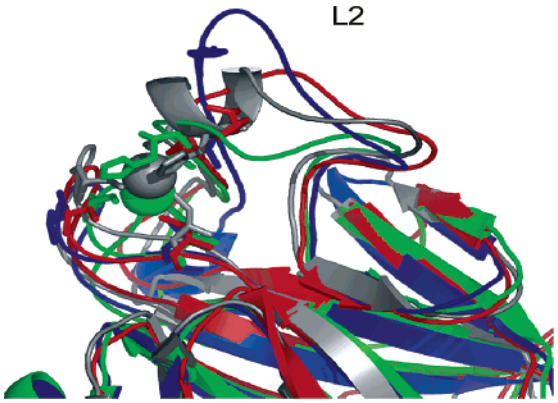

FIGURE 10: Last snapshots of the trajectories from all simulations showing the different conformations of the L2 loop. The colors are the same as in Figure 9 with the complex in black, apocomplex in red, DBD in green, and apoDBD in blue. The Zn^{2+}-coordinating residues are shown in sticks. In DBD, the L2 loop still partially covers Zn^{2+}, while in apoDBD, L2 is unfolded, leaving the Zn^{2+} site fully exposed.

Impact of Zn^{2+} on DNA Binding. The difference in the protein DNA contact pattern between the complex and apocomplex is small and probably cannot alone account for the experimentally observed 11-fold difference in affinity, corresponding to about 1.4 kcal/mol in binding free energy. However, it clearly demonstrates that adding Zn^{2+} increased the sequence specificity because Arg248 is able to form specific interactions only in the complex simulation. In the complex simulation, the DNA was in a much closer proximity to the L3 loop. It is reasonable to assume that Zn^{2+} has a direct influence on the DNA phosphate backbone, because there is a favorable interaction between the ion and phosphate backbone (Figure 6D) and a stronger positive electrostatic potential from DBD according to the Poisson−Boltzmann calculation (Figure 8). The attraction between Zn^{2+} and phosphates facilitated insertion of Arg248 into the minor groove and provided the possibility for sequence-specific interactions with the bases in the minor groove. The salt bridges between Arg248 and the phosphate backbone of the DNA caused the observed contraction of the minor groove. In this picture, the sum of these three weak interactions, the long-range electrostatic interactions between Zn^{2+} and phosphates, the 38% occupancy hydrogen bond between Arg248 and T12′, and the improved van der Waals interaction because of the contraction of the DNA minor groove, accounts for the binding free-energy difference of 1.4 kcal/mol.

One might argue that the contribution from Zn^{2+} is strongly reduced by the screening effects of bulk water and ionic strength such that the described effect may be less pronounced under physiological conditions. Zn^{2+} is however buried, and there is not much water between the DNA and Zn^{2+}. We assessed the amount of water involved by calculating the average hydration number for Zn^{2+} within a radius of 3 Å for the complex simulation and found it to be 0, showing that the ion is entirely shielded by the protein in all directions; we also found that on average there were <0.15 water molecules in a $\sim$2 Å wide corridor from Zn^{2+} to the phosphorus atom of T14 for snapshots taken every 2 ps. Basically, the space between the ion and DNA backbone was occupied by the protein. Because the electrostatic shielding effect of the protein is much weaker than for bulk water, the attractive force between the ion and phosphate backbone is not efficiently screened in this case, as also evidenced by the Poisson−Boltzmann calculation discussed above.

The covariance of the structural elements revealed that surprisingly many of the correlations observed in the complex system were lost in the apocomplex system, many involving elements important for DNA and Zn binding such as the correlation between H1 and the L3 loop. However, the rmsf of the apocomplex was still at the same level as the complex system, indicating that motions in the apocomplex are more random than those in the complex simulation. We propose that not only is the protein−DNA contact pattern important

for DNA binding/recognition but concerted movements between regions of the proteins are also vital, especially among the DNA-binding motifs. Significant covariances found between the DNA-binding structural elements and several β strands in the complex simulation suggest that any structural change upon binding to DNA would have implications in the β-sheet region, which is exactly what has been observed by NMR chemical-shift changes (*42*).

Our data clearly demonstrates that there is an attractive interaction between Zn^{2+} and the DNA phosphate backbone and that Zn^{2+} stabilizes p53 DBD both locally and globally. It is difficult to determine which contribution is dominating in terms of DNA sequence recognition; therefore, we cannot exclude the possibility that the main role of Zn^{2+} is to stabilize the protein.

CONCLUSION

In this molecular dynamics study, not only have we shown the impact that Zn^{2+} has on the stability and dynamics of the p53 DBD, but we also propose a more direct role of Zn^{2+} in DNA binding. Adding the fact that apoDBD is aggregation-prone makes it of great interest to prevent zinc release and/or aggregation of apoDBD. One possible strategy to achieve this is to decrease the intrinsic stress in the L2 loop without affecting the conformation of the Zn^{2+}-binding site. Site-directed mutagenesis in, for example, the H1 helix to lower its propensity to unfold and also in the QHMTEV-VRR peptide to remove any possible frustration would be interesting to perform in the future.

REFERENCES

1. Milner, J. (1997) Structures and functions of the tumor suppressor p53, *Pathol. Biol. 45*, 797−803.
2. Soussi, T., and May, P. (1996) Structural aspects of the p53 protein in relation to gene evolution: A second look, *J. Mol. Biol. 260*, 623−637.
3. Wolkowicz, R., and Rotter, V. (1997) The DNA binding regulatory domain of p53: See the C, *Pathol. Biol. 45*, 785−796.
4. Yonish-Rouach, E. (1997) A question of life or death: The p53 tumor suppressor gene, *Pathol. Biol. 45*, 815−823.
5. Friend, S. (1994) p53: A glimpse at the puppet behind the shadow play, *Science 265*, 334−335.
6. Hainaut, P., and Hollstein, M. (2000) p53 and human cancer: The first ten thousand mutations, *Adv. Cancer Res. 77*, 81−137.
7. Levine, A. J. (1997) p53, the cellular gatekeeper for growth and division, *Cell 88*, 323−331.
8. Bullock, A. N., and Fersht, A. R. (2001) Rescuing the function of mutant p53, *Nat. Rev. Cancer 1*, 68−76.
9. Cho, Y., Gorina, S., Jeffrey, P. D., and Pavletich, N. P. (1994) Crystal structure of a p53 tumor suppressor−DNA complex: Understanding tumorigenic mutations, *Science 265*, 346−355.
10. Verhaegh, G. W., Parat, M. O., Richard, M. J., and Hainaut, P. (1998) Modulation of p53 protein conformation and DNA-binding activity by intracellular chelation of zinc, *Mol. Carcinog. 21*, 205−214.
11. Rainwater, R., Parks, D., Anderson, M. E., Tegtmeyer, P., and Mann, K. (1995) Role of cysteine residues in regulation of p53 function, *Mol. Cell Biol. 15*, 3892−3903.
12. Meplan, C., Richard, M. J., and Hainaut, P. (2000) Metalloregulation of the tumor suppressor protein p53: Zinc mediates the renaturation of p53 after exposure to metal chelators in vitro and in intact cells, *Oncogene 19*, 5227−5236.
13. Butler, J. S., and Loh, S. N. (2003) Structure, function, and aggregation of the zinc-free form of the p53 DNA binding domain, *Biochemistry 42*, 2396−2403.
14. Coffer, A. I., and Knowles, P. P. (1994) Divalent metal ions induce conformational change in pure, human wild-type p53 tumor suppressor protein, *Biochim. Biophys. Acta 1209*, 279−285.
15. Brook, B., Bruccoleri, R., Olafson, B., States, D., Swaminathan, S., and Karplus, M. (1983) Charmm: A program for macromolecular energy, minimization and dynamic calculations, *J. Comput. Chem. 4*, 187−217.
16. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B 102*, 3586−3616.
17. Bredenberg, J., and Nilsson, L. (2001) Modeling zinc sulfhydryl bonds in zinc fingers, *Int. J. Quantum Chem. 83*, 230−244.
18. Stote, R. H., and Karplus, M. (1995) Zinc-binding in proteins and solution−A simple but accurate nonbonded representation, *Proteins: Struct., Funct., Genet. 23*, 12−31.
19. Foloppe, N., and MacKerell, A. D. (2000) All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data, *J. Comput. Chem. 21*, 86−104.
20. van Gunsteren, W. F., and Berendsen, H. J. C. (1977) Algorithms for macromolecular dynamics and constraint dynamics, *Mol. Phys. 34*, 1311−1327.
21. Jorgensen, W. L., Chandrasekar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys. 79*, 926−935.
22. Brooks, C. L., III, Brunger, A., and Karplus, M. (1985) Active site dynamics in protein molecules: A stochastic boundary molecular-dynamics approach, *Biopolymers 24*, 843−865.
23. Norberg, J., and Nilsson, L. (2000) On the truncation of long-range electrostatic interactions in DNA, *Biophys. J. 79*, 1537−1553.
24. Steinbach, P. J., and Brooks, B. R. (1994) New spherical cutoff methods for long-range forces in macromolecular simulation, *J. Comput. Chem. 15*, 667−683.
25. Prevost, M., van Belle, D., Lippens, G., and Wodak, S. (1990) Computer simulations of liquid water: Treatment of long-range interactions, *Mol. Phys. 71*, 587−603.
26. Beck, D. A., Armen, R. S., and Daggett, V. (2005) Cutoff size need not strongly influence molecular dynamics results for solvated polypeptides, *Biochemistry 44*, 609−616.
27. de Loof, H., Nilsson, L., and Rigler, R. (1992) Molecular dynamics simulation of galanin in aqueous and nonaqueous solution, *J. Am. Chem. Soc. 114*, 4028−4035.
28. Lavery, R., and Sklenar, H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids, *J. Biomol. Struct. Dyn. 6*, 63−91.
29. Frishman, D., and Argos, P. (1995) Knowledge-based protein secondary structure assignment, *Proteins 23*, 566−579.
30. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics, *J. Mol. Graphics 14*, 27−38.
31. Rost, B., and Sander, C. (1993) Prediction of protein secondary structure at better than 70% accuracy, *J. Mol. Biol. 232*, 584−599.
32. Garnier, J., Gibrat, J. F., and Robson, B. (1996) GOR method for predicting protein secondary structure from amino acid sequence, *Methods Enzymol. 266*, 540−553.
33. Meiler, J., Mueller, M., Zeidler, A., and Schmaeschke, F. (2001) Generation and evaluation of dimension-reduced amino acid parameter representation by artificial neural networks, *J. Mol. Model. 7*, 360−369.
34. Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M., and Barton, G. J. (1998) JPred: A consensus secondary structure prediction server, *Bioinformatics 14*, 892−893.
35. Kneller, D. G., Cohen, F. E., and Langridge, R. (1990) Improvements in protein secondary structure prediction by an enhanced neural network, *J. Mol. Biol. 214*, 171−182.
36. Karplus, K., Barrett, C., and Hughey, R. (1998) Hidden Markov models for detecting remote protein homologies, *Bioinformatics 14*, 846−856.
37. Pollastri, G., Przybylski, D., Rost, B., and Baldi, P. (2002) Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles, *Proteins 47*, 228−235.
38. Maret, W. (2004) Zinc and sulfur: A critical biological partnership, *Biochemistry 43*, 3301−3309.

39. Coleman, J. E. (1992) Zinc proteins: Enzymes, storage proteins, transcription factors, and replication proteins, *Annu. Rev. Biochem. 61*, 897−946.
40. Garvie, C. W., and Wolberger, C. (2001) Recognition of specific DNA sequences, *Mol. Cell 8*, 937−946.
41. Thirumalai, D., Klimov, D. K., and Dima, R. I. (2003) Emerging ideas on the molecular basis of protein and peptide aggregation, *Curr. Opin. Struct. Biol. 13*, 146−159.
42. Rippin, T. M., Freund, S. M., Veprintsev, D. B., and Fersht, A. R. (2002) Recognition of DNA by p53 core domain and location of intermolecular contacts of cooperative binding, *J. Mol. Biol. 319*, 351−358.
43. Klein, C., Planker, E., Diercks, T., Kessler, H., Kunkele, K. P., Lang, K., Hansen, S., and Schwaiger, M. (2001) NMR spectroscopy reveals the solution dimerization interface of p53 core domains bound to their consensus DNA, *J. Biol. Chem. 276*, 49020−49027.
44. Nagaich, A. K., Zhurkin, V. B., Durell, S. R., Jernigan, R. L., Appella, E., and Harrington, R. E. (1999) p53-induced DNA bending and twisting: p53 tetramer binds on the outer side of a DNA loop and increases DNA twisting, *Proc. Natl. Acad. Sci. U.S.A. 96*, 1875−1880.
45. Veprintsev, D. B., Freund, S. M., Andreeva, A., Rutledge, S. E., Tidow, H., Canadillas, J. M., Blair, C. M., and Fersht, A. R. (2006) Core domain interactions in full-length p53 in solution, *Proc. Natl. Acad. Sci. U.S.A. 103*, 2115−2119.

BI0603165