Research Article

Journal of Molecular Recognition

Received: 18 February 2009, Revised: 29 June 2009, Accepted: 19 August 2009, Published online in Wiley InterScience: 23 October 2009

(www.interscience.wiley.com) DOI:10.1002/jmr.990

Preferred drifting along the DNA major groove and cooperative anchoring of the p53 core domain: mechanisms and scenarios

Yongping Pan[a] and Ruth Nussinov[a,b]*

While the importance of specific p53-DNA binding is broadly accepted, the recognition process is still not fully understood. Figuring out the initial tetrameric p53-DNA association and the swift and *cooperative* search for specific binding sites is crucial for understanding the transactivation mechanism and selectivity. To gain insight into the p53-DNA binding process, here we have carried out explicit solvent molecular dynamic (MD) simulations of several p53 core domain-DNA conformations with the p53 and the DNA separated by varying distances. p53 approached the DNA, bound non-specifically, and quickly drifted along the DNA surface to find the major groove, cooperatively anchoring in a way similar to the specific binding observed in the crystal structure. Electrostatics was the major driving force behind the p53 movement. Mechanistically, this is a cooperative process: key residues, particularly Lys120 and Arg280 acted as sensors; upon finding their hydrogen-bonding partners, they lock in, anchoring p53 into the major groove. Concomitantly, the DNA adopted a conformation that facilitated p53 easy access. The initial non-specific core domain-DNA contacts assist in shifting the DNA and the p53 substrates toward conformations "ready" for specific major groove binding, with subsequent optimization of the interactions. This work is an invited contribution for the special issue of the Journal of Molecular Recognition dedicated to Professor Martin Karplus.

Supporting information may be found in the online version of this paper.

Keywords: protein-DNA interactions; cooperativity; molecular dynamics simulations; DNA sliding; specific versus non-specific binding; binding mechanism

INTRODUCTION

In response to cellular stress such as DNA damage, tumor suppressor p53 is activated and functions as a transactivator to regulate several genes, preventing the development of cancer (Kastan *et al*., 1991; el-Deiry, 1998; Vousden, 2002; Vousden and Lu, 2002). A critical biochemical event for the p53 transactivation activity is its sequence-specific binding to DNA (Bargonetti *et al*., 1991). More than 500 different p53-response elements in the human genome have been characterized (el-Deiry *et al*., 1992; Wei *et al*., 2006; Horvath *et al*., 2007; Smeenk *et al*., 2008; Zeng *et al*., 2008), and their physical and functional properties have been analyzed and categorized (Riley *et al*., 2008). These DNA binding sites are usually composed of two 10-base pair (bp) repeats 5′-PuPuPuC(A/T)|(A/T)GPyPyPy-3′, separated by 0–13 base pairs (el-Deiry *et al*., 1992; Funk *et al*., 1992), where Pu and Py stand for purine and pyrimidine bases, respectively (el-Deiry *et al*., 1992; Balagurumoorthy *et al*., 1995; Balagurumoorthy *et al*., 2002). While most p53 response elements contain no or smaller spacers, those involved in negative regulation were found to have larger spacer sizes relative to those playing a role in positive regulation (Riley *et al*., 2008). The spacer size has been suggested to be important in binding cooperativity and selective transactivation (Pan and Nussinov, 2009). These p53 binding sites are embedded in a very large sized genome, making an efficient search of the binding site critical.

Structural information relating to the p53-DNA complex can greatly promote the elucidation of the binding process. The p53 protein is a tetramer of four homologous peptide chains. Each chain is composed of multiple functional domains, including the core domain (CD) and the C-terminal basic domain (BD) that are involved in DNA binding (Fields and Jang, 1990; Cho *et al*., 1994; Clore *et al*., 1994). While the BD is generally believed to bind to DNA non-specifically, the CD binds to specific DNA sequences in a specific manner (Cho *et al*., 1994; Wang *et al*., 1995; Waterman *et al*., 1995; McLure and Lee, 1998), and each pair of core domains associate with each other through two salt bridges in the H1 helix

* *Correspondence to: R. Nussinov, NCI-Frederick, Bldg 469, Rm 151, Frederick, MD 21702, USA.*
E-mail: ruthnu@helix.nih.gov

a *Y. Pan, R. Nussinov*
Basic Research Program, SAIC-Frederick, Inc. Center for Cancer Research Nanobiology Program, NCI-Frederick, Frederick, MD 21702, USA

b *R. Nussinov*
Sackler Institute of Molecular Medicine, Department of Human molecular Genetics and Biochemistry, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

region (Klein *et al*., 2001b; Rippin *et al*., 2002; Sun *et al*., 2003; Dehner *et al*., 2005). High resolution chemical foot-printing and cross-linking experiments have confirmed the specific binding of each of the four p53 core domains and the C_2 symmetry of the tetrameric p53 in the DNA-bound state (Nagaich *et al*., 1997b). The crystal structures of the DNA-p53 complex also reveal the important interactions and key residues involved in the specific binding. Among the six residues that are in direct contact with DNA, Arg280 and Lys120 stretch into the major groove while Arg248 is buried in the minor groove of the DNA. Relative to the rich experimental data for the p53 core domain-DNA complex structure, the available high-resolution oligomer structures of p53 in the DNA-unbound state is less consistent, with different interacting interfaces between the core domains detected by different groups, illustrating the variability in the CD–CD association (Zhao *et al*., 2001; Veprintsev *et al*., 2006; Wang *et al*., 2007). The NMR-based CD dimer model in the full length p53 suggests conformational changes for the dimer from the unbound to the DNA-bound states (Veprintsev *et al*., 2006). A p53 tetramer in the unbound state has also been solved based on electron microscopy (EM) (Okorokov *et al*., 2006). However, the D_2-symmetry organization of the tetramer and the structural details relating to the CD orientation make it difficult to deduce the conformational transition during the process of binding to a full 20-base pair site that has no base pair insertions. In addition, experiments further suggest that in solution p53 may exist mainly in the dimer form (Veprintsev *et al*., 2006; Poon *et al*., 2007) and possibly in a different organization between the core domains (Cho *et al*., 1994; Zhao *et al*., 2001; Veprintsev *et al*., 2006; Wang *et al*., 2007). It thus appears that current structural data are insufficient to evolve a clear picture regarding the p53-DNA binding process.

Experimental results regarding the *dynamic* p53-DNA binding process show that p53 can directly bind to a specific DNA sequence or initially bind to DNA non-specifically with subsequent translocation by one-dimensional diffusion (Jiao *et al*., 2001). The C-terminal BD has been shown to be important for the initial non-specific binding to DNA and linear diffusion along the DNA (Jayaraman and Prives, 1995; McKinney and Prives, 2002; McKinney *et al*., 2004; Liu and Kulesz-Martin, 2006). The difference in the C-terminal domain is also partially responsible for differences in DNA binding (Sauer *et al*., 2008). As a result, the complex experiences conformational changes such as axial DNA bending (Nagaich *et al*., 1997a; Cherny *et al*., 1999; Nagaich *et al*., 1999), and exposure of p53 motifs for acetylation or other post-transcriptional modification (Ceskova *et al*., 2006). The DNA topology itself and its sequence can also influence the p53 specific binding (Jagelska *et al*., 2008) and binding cooperativity (Pan and Nussinov, 2008), respectively. While the search can be enhanced by intra-molecular motifs such as the C-terminal domain, or residue modification such as acetylation (Luo *et al*., 2004), it is unclear *how* p53 searches the DNA sequence. More importantly, such a dynamic binding process model does not account for the role of p53 CD that is deemed to be of importance for the specific recognition process. Particularly interesting issues include whether p53 finds its binding site one core domain at a time or a dimer of p53 core domain acts in a concerted mode. Current experimental evidence indicates that p53 exists mainly in a dimer form before it binds to the DNA (Veprintsev *et al*., 2006), which suggests that p53 is more likely to bind the DNA as a dimer rather than as monomers randomly binding to each quarter site; however, even within this framework, the first step is likely to involve the cooperative monomer-DNA interactions.

The study of the process leading to the initial interactions and recognition between a single p53 CD and DNA is essential to the understanding of the binding mechanism. Because the p53 CD and its DNA-bound structures are available, such studies can be reasonably addressed through molecular modeling and molecular dynamic (MD) simulations. Here, we performed MD simulations on a series of conformations in which the DNA and p53 were placed at different distance intervals with respect to each other. We show that when p53 was placed near the minor groove, it quickly drifted to the major groove, guided by electrostatic interactions and geometry match. Our data reveal the details of the p53-DNA specific binding process, highlighting the roles of factors such as the presence of basic residues and the overall favorable interaction due to shape complementarity between the two molecules.

MATERIAL AND METHODS

Model construction

First, p53 core domain dimer was constructed based on the crystal structure (PDB code 1TSR) (Cho *et al*., 1994), as described previously (Pan and Nussinov, 2007). This DNA-unbound p53 dimer was relaxed through MD simulations, resulting in a series of conformations in which the DNA binding interfaces of the two core domains were further apart due to the charge-charge repulsions but were still associated with each other through the salt bridges. Two snapshots at 5 and 3 ns were selected from the simulation and the DNA was placed back into the system by matching the p53 CD in the snapshots with the crystal structure, resulting in a conformation in which one of the core domains was bound to DNA specifically while the other had little contact with the DNA. These two snapshots were chosen based on the distance between the H1 helices from the two CD, with the 5 ns the furthest and 3 ns at the middle. Removal of the p53 monomer that was specifically bound to DNA resulted in p53 monomer-DNA dimer entity, referred to as I and II (Figure 1). These two systems were the starting points of this study and were subjected to MD simulations to observe the dynamic process of p53 monomer-DNA interaction from different starting conformations of the p53 with respect to the DNA. A third starting conformation was obtained by moving the p53 further away from the DNA by

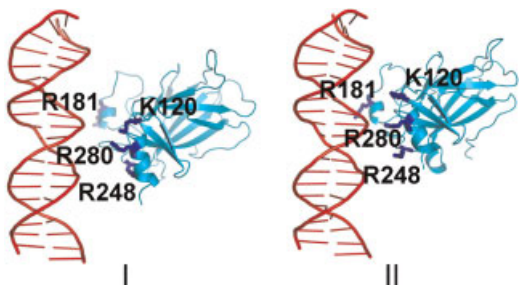

Figure 1. Cartoon representation of the starting conformations of the p53 core domain-DNA systems. Conformations (I) and (II) were extracted from the p53 core domain dimer simulation at 5 and 3 ns in the trajectory. For clarity, only the backbones and the key residues were shown. DNA, p53, and key residues were colored red, cyan and blue, respectively.

2 Å (not shown). The DNA segment with the sequence of 5′CTAGACTTGCCCAAT3′ was extracted from the E chain of the crystal structure (Cho *et al.*, 1994).

MD simulation protocol

Each system was first solvated with a TIP3P water box (Jorgensen *et al.*, 1983) with a margin of at least 10 Å from any edge of the water box to any protein or DNA atom. Solvent molecules within 1.6 Å of the DNA or within 2.5 Å of the protein were removed. The systems were then neutralized by adding sodium ions. The solvated systems were subjected to a series of minimizations and equilibrations before the production MD simulations with the CHARMM program (Brooks *et al.*, 1983) and the CHARMM 27 force field (MacKerell *et al.*, 1998). Minimizations were first performed for 500 steps with the steepest decent algorithm with the backbone of the p53 and DNA constrained and additional 500 steps for the whole system to eliminate residual strains in the system. Periodic boundary conditions were applied and the non-bonded lists were updated every 20 steps. NPT ensemble was used with the pressure kept at 1 atom and temperature at 300 K using Langevin–Nose-Hoover coupling. SHAKE constraints on all hydrogen atoms to allow the time step of 2 fs in production simulations. Electrostatic energies were calculated with the PME algorithm. The systems were first equilibrated for 20 ps and the production simulations lasted for 27 and 21 ns, respectively. The simulations were terminated when no further conformational changes toward the crystal structure were observed. Structures were saved every 2 ps for analysis.

RESULTS

The process of DNA recognition by p53 core domain

In the initial structure of the complex I (Figure 1), the p53 core domain was placed in a position such that Lys120 and Arg280 faced the DNA minor groove but had little contact with it. The snapshots in Figure 2 (taken from the 20 ns trajectory) present the progression of the conformational changes in the p53 and the DNA with respect to these specific interaction residues. At the beginning of the trajectory, p53 quickly approached the DNA gaining significant interactions with the DNA backbone through the protruding positively charged Arg280. At 0.3 ns, p53 was already in close contact with the DNA, with Arg280 making interactions with the DNA backbone and Lys120 pointing to the major groove (Figure 2B). Thus, within such a short period, p53 not only moved closer to the DNA, but also traveled upward, along the DNA (Figure 2B). At 0.75 ns, p53 moved further along the DNA axis, with Lys120 penetrating deeper into the major groove while still maintaining the Arg280 interaction with the DNA backbone, albeit with a slightly changed orientation (Figure 2C). However, the interaction between Lys120 and the DNA was still limited (Figures 2C, 3B). At 1.6 ns, Lys120 made contact with the DNA backbone while Arg248 moved toward the minor groove (Figure 3D). At the same time, between 0.75 and 1.6 ns the DNA conformation also changed, with the DNA bending toward p53 facilitating its interactions with Lys120, Arg280 and Arg248. At 2.63 ns, Arg248 moved even closer toward the minor groove while Arg280 and Lys120 maintained the interactions with the DNA backbone (Figure 2E). Nevertheless, from this point to 5 ns, the Lys120 and Arg248 interactions were still unstable, and Arg280 continued to adjust itself and by 5 ns into the trajectory was partially inserted into the major groove (Figure 2F). After that, the protein and DNA further optimized their interactions with both Lys120 and Arg280 frequently getting into the major groove (Figure 2G). By 10 ns these two residues obtained better interactions, pushing Arg248 away from the minor groove. Further dynamic interactions allowed the molecule to adjust at 19.8 ns, now with all three key residues interacting with the DNA simultaneously (Figure 2H). Figure 3B shows that Arg248 contributed to the overall interaction only at the beginning and after 19 ns into the trajectory. Comparing the snapshots at 2.63 ns and 19.8 ns reveals the similarity, indicating that it is the preferred conformation for the two partners to interact. Further, at 19.8 ns, both Lys120 and Arg280 were more buried in the major groove, a conformation similar to the specific binding mode observed in the crystal structures (Cho *et al.*, 1994; Ho *et al.*, 2006). These results show that p53 adjusted its orientation, reaching its preferred organization. At this point, with Lys120 and Arg280 nestling in the major groove, the relative movement of p53 and DNA with respect to each other became much more confined. Detailed analysis of the specific interactions between p53 and DNA indicates that the Lys120-Gua2 and Arg280-Gua4 hydrogen bonds distances decreased rapidly at the first half ns and then changed gradually. However, in the 20+ ns simulation reported here, the distance did not reach the hydrogen bonding range, with the shortest at about 5 Å (Figure 3C).

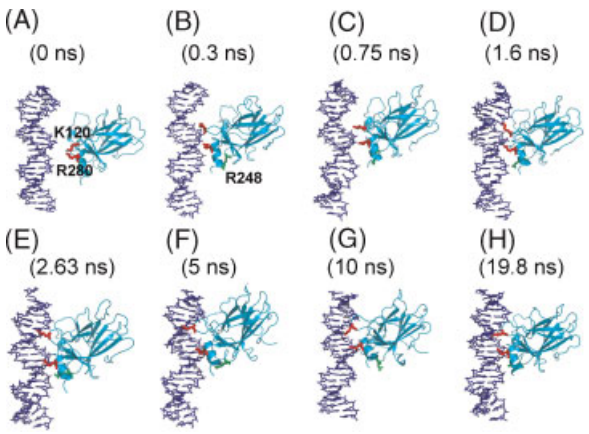

Figure 2. Snapshots from the simulation trajectory from starting structure I. Only the backbone and residues Lys120, Arg280 and Arg248 are shown for clarity. In the starting structure (A), both Lys120 and Arg280 faced the minor groove of the DNA. At the end of the simulation, Lys120, Arg280 and Arg248 were anchored in the major groove and minor groove respectively. The overall organization of the complex is very similar to that observed in crystal structures.

For the simulation of complex II, the starting structure was chosen such that Lys120 faced the major groove but was not in contact with any of the DNA bases and Arg280 pointed toward the minor groove (Figure 1). This conformation was also slightly closer to the crystal structure than conformation I. Figure 4 illustrates the progression of the binding process. Within a short time (1.04 ns) p53 quickly moved toward the DNA and made contact through Arg280. In the meantime, Lys120 and Arg280 adjusted their positions, with Lys120 hovering around the major groove (Figure 4B). At 2.19 ns, Arg280 crossed the backbone and interacted with the DNA backbone from the major groove side

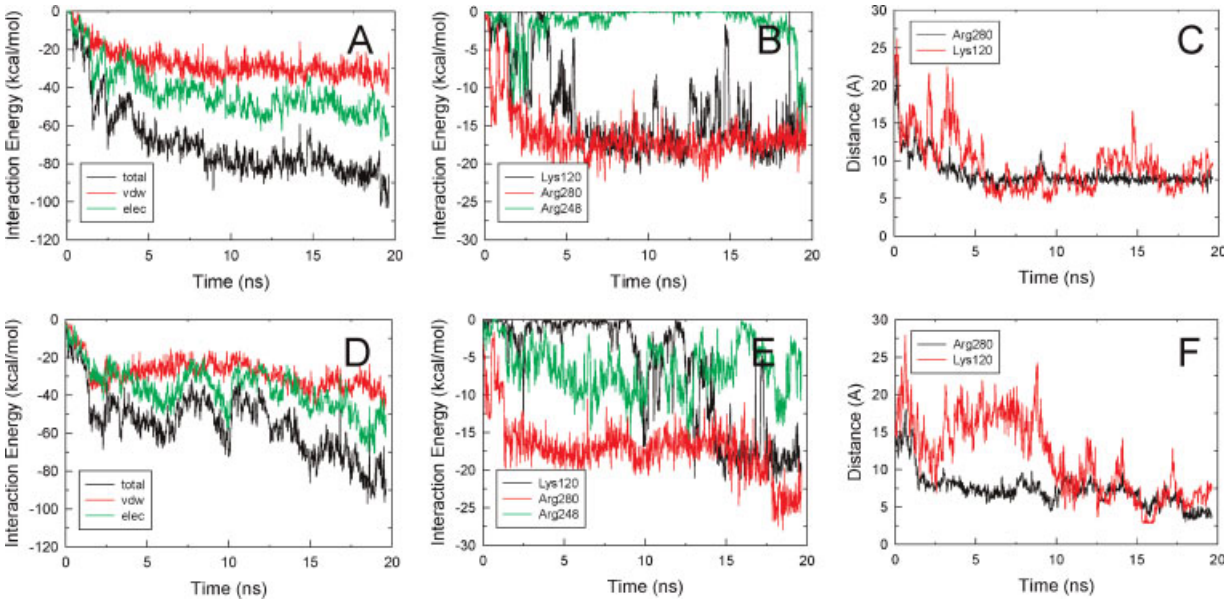

Figure 3. Quantitative description of the interactions between p53 and DNA. The top panel is from simulation I and the bottom panel for simulation II. (A and D) total interaction energies (black), van der Waal (red) and electrostatic interaction energies (green) between p53 and DNA; (B and E) interaction energies for each of the three basic residues Lys120 (black), Arg280 (red) and Arg248 (green); (C and F) the distances of Lys120 (black) and Arg280 (red) to their respective Guanidine bases with which they formed hydrogen bonds in the crystal structure.

(Figure 4C). Importantly, Arg248 was also able to find its position toward the minor groove (Figure 4C). However, Lys120 was unable to quickly establish stable interactions with the DNA and settle down snugly into the major groove. Instead, it floated around much of the time and made significant contact with the DNA only at around 10 ns into the simulation (Figure 4D, E). During the 5 to 10 ns period, the major conformational change was in the orientation of Lys120; this resulted in a conformation in which both Arg280 and Lys120 were inserted into the major groove (Figure 4E). However, this organization fell apart again; until at about 14 ns Lys120 re-built its major groove interactions and the continued dynamics further strengthened the Lys120 and Arg280 interactions with the DNA. By 15.99 ns, both residues penetrated deeper into the major groove; the residues pointed toward the Gua bases, resulting in a conformation similar to that observed in the crystal structure, where the bases interacted with Lys120 and Arg280 through hydrogen bonding. The distances corresponding to the hydrogen bonds between Lys120 (side-chain Nitrogen atom) and Gua2 (Nitrogen atom at the major groove side) and between Arg280 (one of the terminal Nitrogen atoms) with Gua4 were 2.75 and 4.43 Å, respectively. In addition, the amide hydrogen of Lys120 made close contact/hydrogen bond with the DNA backbone oxygen at the same position as in the crystal structure, although the Arg248 interaction was very dynamic. After 18 ns into the trajectory, all three residues made significant interactions with the DNA, with the Arg280 interactions the most favorable (Figure 3E). Geometrical analysis shows that the distance between Arg280 and the Gua base was within the hydrogen bonding range. This was not observed in the first simulation. The overall conformation at this point was very similar to the crystal structure (Figure 4H).

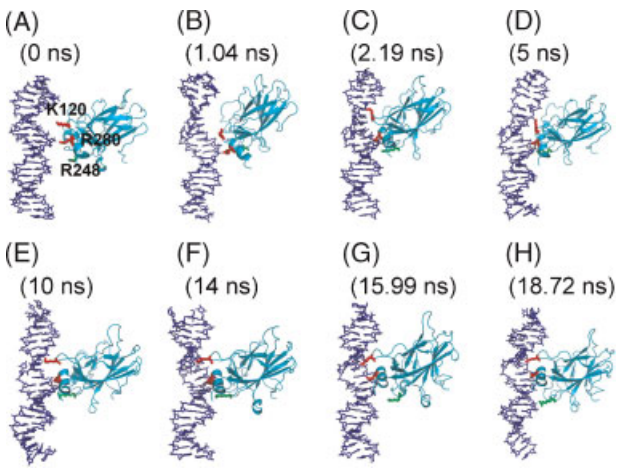

Figure 4. Snapshots from the simulation trajectory from starting structure II. Only the backbone and residues Lys120, Arg280 and Arg248 are shown for clarity. In the starting structure (A), Lys120 faced the major groove while Arg280 faced the minor groove of the DNA. At the end of the simulation, Lys120, Arg280 and Arg248 were anchored in the major groove and minor groove respectively. The overall organization of the complex is very similar to that observed in crystal structures.

Thus, in two simulations, with different starting structures, both away from the native binding site and binding orientation, the interactions of the three charged residues with the DNA ended in a complex conformation similar to the p53 specific binding. However, this was not the case in all simulations. Due to the positive charges, the surface of the p53 core domain possesses different regions that can potentially interact with the DNA. In one simulation where the p53 and DNA were further apart in the starting structure, in the initial stage of the Lys120 and Arg280 interaction with DNA, Arg181 also approached the DNA at the same time (data not shown). As a result, all three residues were in contact with the DNA, leading to a conformation in which Arg280

and Lys120 interacted with the minor groove while Arg181 attached on the opposite side of the DNA (data not shown). Although such a p53-DNA complex organization was not observed in crystal structures, this result still suggests that the interaction of Arg181 with DNA may interfere with the specific binding. Coincidentally, Arg181 was involved in the p53 dimerization interface in the DNA-bound conformation. The fact that the co-crystallization of p53-DNA encounters great difficulties suggests that other non-specific interactions between the p53 core domain and DNA may compete and interfere with the cooperative association among the species of the complex. A comparative study of the DNA binding with a family of p53 proteins, including p53, p63, and p73 suggested that cooperative binding is due to the p53 dimerization at the H1 helix interface (Klein *et al*., 2001a). The study of the binding of p53 modified at the H1 helix supported the importance of the dimerization of p53 in achieving the native p53-DNA interactions (Sun *et al*., 2003). The possible dimerization of the core domain prior to the binding to DNA could prevent non-specific interactions and accelerate the binding process.

Overall conformational change and comparison with the crystal structures

As described above, after p53 docked into the major groove, further movement of the p53 with respect to the DNA became limited. In both simulations the final conformations were very different from the starting structures, with p53 observed to be significantly rotated. The colored structural motifs in Figure 5 highlight the changes in the orientation. From the comparison between A and B and between C and D in Figure 5, it can be seen that the helix H2 orientation changed to become more perpendicular toward the viewer, and the loop (highlighted in cyan) became invisible. The final position of p53 was very similar to what was observed in the crystal structures (Figure 5E). Comparison of the orientations of the same motifs in the structure shows that the final structure in simulation II was even closer to the crystal structure conformation. Such a similar organization further suggests that the binding/recognition process observed here may be a reasonable representation of the natural process.

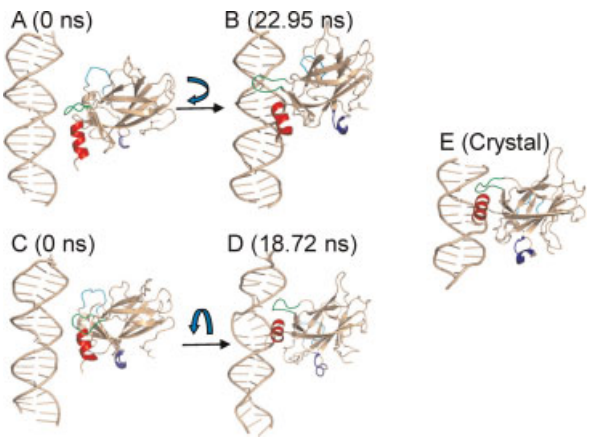

Figure 5. Comparison of the starting conformation with the end conformation. A and B are for simulation I and C and D are for simulation II. Several loops and helix H2 are colored differently to facilitate the visualization of the p53 rotation with respect to DNA. The orientation change of the p53 core domain with respect to DNA is similar to a rotation indicated with the arrows.

Specific hydrogen bonding properties were also monitored to compare the similarity between the simulation and the crystal structure. As shown in Figure 3C, throughout the simulation, none of the donors from Arg280 or Lys120 in complex I were able to form hydrogen bonds with the specific bases. The shortest distance between the donor and the acceptor was around 5 Å. In this case, both Lys120 and Arg280 interacted with the DNA backbone atoms. However, upon breaking of these interactions with the backbone and simple rotation of the side chains of Arg280 or Lys120, these two residues can easily position at the right place to form specific hydrogen bonds. Encouragingly, in the second simulation, these hydrogen bonds were observed at different times (Figure 3F). At 12.5 ns and between 15 and 16 ns, the distance of the hydrogen bond of Lys120 reached 2.5 Å. The distance of the hydrogen bond of Arg280 also reached around 3 Å near the end of the trajectory (from 18 ns to 20 ns).

The driving force for the p53 movement on the DNA

Interactions between p53 and DNA

The initial contact between the p53 and the DNA was triggered by both the electrostatic and VDW interactions shown in Figures 3A, D. Electrostatic attractions played a more pronounced role in both cases, as the two molecules were oppositely charged at the interacting surfaces. In simulation I, the total interaction energy sharply decreased from the beginning until around 5 ns when it became relatively stabilized (Figure 3A). Interaction energies for individual residues were also plotted in Figure 3B and S1-A. It can be seen that most other residues contributed less significantly than Lys120 and Arg280, except Lys139. On the surface of p53, Lys139 was located in the vicinity of loop L1, and interacted with the DNA backbone. All other residues interacted with DNA because their locations were directly related to the residues that were in contact with DNA in the crystal structure—Lys120, Ser241, Arg273, Ala276, Cys277, Arg248, Arg280, and Arg283 (Cho *et al*., 1994). All the residues shown in Figure S1 were in close vicinity to one of the residues that were in contact with the DNA in the crystal structure.

Among the many residues interacting with the DNA, Lys120, Arg280, and Arg248 were inspected in detail because of their direct role in specific DNA binding. The importance of these residues in specific binding is also supported by additional experimental evidence derived from NMR, mutational analysis (Klein *et al*., 2001b; Ishimaru *et al*., 2003), and X-ray crystallography. Figure 3B shows that in the first 5 ns Arg280 continuously gained favorable interaction while Lys120 and Arg248 contributed only transiently in that period. After 5 ns, the Arg280 and Lys120 interactions became persistent, although the interaction energy for Lys120 still varied significantly between 10 to 20 ns. The fluctuation in the Lys120 interaction energy was caused by the movement of the Lys120 side-chain which mainly interacted with the DNA backbone and at the same time, frequently deviated from the DNA backbone as shown in Figure 2. However, the dynamic binding process indicates that while Arg280 quickly gained favorable interactions with the DNA, the flexibility of the Lys120 orientation seemed to have contributed to the conformational search of the molecule. This flexibility of Lys120 led to the movement toward the major

groove by reaching the other edge of the major groove, which eventually resulted in the further movement of Arg280 into the major groove. Arg248 made more favorable interactions around 2 ns in the early stage of the trajectory. However, this interaction diminished for much of the rest of the trajectory. It regained significant interactions at around 18 ns, due to the overall favorable orientation of the molecule with respect to the DNA.

In simulation II, the overall interaction energy also sharply decreased in the first 2.5 ns (Figure 3D). However, the overall interaction energy change was not as smooth as in simulation I, although the total energy at the end of the simulation reached a comparable level to that in the first simulation (Figure 3A, D). The involvement of the three residues in the binding process was also different from simulation I. Arg280 still gained favorable interaction quickly at the start of the trajectory and made significant interactions with the DNA throughout the run (Figure 3E). However, different from simulation I, Arg248 also played an important role during the whole process, while Lys120 essentially did not interact with the DNA in the early stage (Figure 3E). After 14 ns, Lys120 established a persistent interaction with DNA; the interaction energy for this residue was stabilized and maintained at about −20 kcal/mol, similar to that observed in simulation I (Figure 3E). Structural analysis shows that during that period, Lys120, as well as Arg280, were inserted into the major groove (Figure 4H). Interaction energies for all the residues that were involved in the protein–DNA interactions were also plotted (Figure S1-B). Among residues other than Lys120, Arg280 or Arg248, Lys139 gained significant interaction in the first 7 ns of the simulation. Interestingly, this interaction disappeared soon after that point.

Regardless of the differences between the two simulations, those results show that the positively charged surface residues allowed p53 to quickly access the DNA and make favorable interactions. The positive charges of the residues (Lys120, Arg280, and Arg248) and their specific locations on the surface played critical roles in the positioning of p53 with respect to DNA. When properly positioned, these residues dominated the movement of the molecule. In particular, when Lys120 and Arg280 were properly positioned in the major groove, Arg248 was in an advantageous position that allows it to tighten up the interaction with the minor groove (simulation I). Similarly, when Arg280 and Arg248 were properly bound to the DNA, as in simulation II, Lys120 was also in a favorable position to bind in the major groove. Therefore, these charged residues seemed to serve collectively as a binding motif and to be able to orient p53, resulting in the insertion of Arg280 and Lys120 in the major groove and the anchoring of Arg248 in or near the minor groove. Interestingly, the Arg248 interaction with DNA differed in the crystal structures. It either directly inserts into the minor groove or interacts with the minor groove through bridging water molecules (Cho *et al*., 1994; Kitayner *et al*., 2006). In our simulations, we only observed that Arg248 attached to the edge of the minor groove, without penetration.

The role of ions

At the beginning of the simulations sodium ions were evenly distributed around the complex. With the progression of the simulations, more ions surrounded the DNA molecule than the protein (data not shown). We also observed that the interactions of the charged residues, particularly Lys120, with the DNA were very dynamic, which allowed it to search broader conformational space. It is known that charge–charge interactions are normally very persistent, which was also observed in these simulations. Such strong interactions can trap the residues in a non-optimal conformation. For example, in the first simulation, Lys120 was unable to find its way to project deeper into the major groove due to its interactions with the charged DNA backbone. While in the second simulation, Lys120 was able to break away from the backbone, possibly due to the intervention of sodium ions competing for DNA interactions. Such a dynamic process may or may not happen in the short time-scale simulations. On the other hand, the presence of a negatively charged ion such as chloride may also help mediate the interactions between the repulsive Lys120 and Arg280. As was observed in the simulation, the two residues were separated by a distance much larger than that observed in the crystal structure, and the loop L1 motif conformation was also very different due to the deviation of Lys120. A negatively charged chloride ion can bring them closer so that both Arg280 and Lys120 can interact with the bases in the major groove. Regardless of the presence of ions, Lys120 may still just float around the surface of the backbone, resulting in its weak interaction with DNA and unresolved conformation in some crystal structures (Ho *et al*., 2006).

DISCUSSION AND CONCLUSIONS

Here we employ MD simulations to gain insight into the dynamic behavior at the early stage of the p53-DNA recognition process. Our simulations initiate when the p53 core domain already accesses the DNA with the "right" surface due to time scale limitations. It behooves us to also emphasize that as always in simulations, the observed dynamic properties may be influenced by the simulation conditions such as the starting structure, molecular (sequence) size—protein and in our case particularly DNA length, presence of ions, etc. More over, in reality, the p53 core domain may contact the DNA surface through *any* accessible surface; however, simulations cannot afford to explore all possibilities. Local minima can also constitute caveats: once the p53 core domain is in contact with DNA, the strength of the salt bridge can trap a certain conformation even if the two binding partners are not in the optimal specific configuration. Such problems are typical in simulations, and they are expected to be reduced if longer simulation time scales are applied. Mimicking physiological conditions, like salt concentrations can also alleviate such problems: in our case, the stability of the p53 core domain is known to be affected by the change of the salt concentrations (Ishimaru *et al*., 2009; Xue *et al*., 2009a, 2009b). Further, modified bases or residues will affect the protein-DNA interaction: methylation was shown to alter the recognition (Petrovich and Veprintsev, 2009) and acetylation of Lys120 affects the transcription-independent apoptosis (Sykes *et al*., 2009) function of the p53.

In the context of the entire p53 protein, the core domain can be expected to exist in equilibrium between their isolated states and the dimeric organization. A recent work that presented a low-resolution p53 tetramer model supports such a premise (Tidow *et al*., 2007). The p53 core domain binding to DNA can thus be illustrated by three scenarios (Figure 6). In the first scenario (Figure 6A), the p53 dimers need to undergo significant conformational transformations before they can bind to their cognate response elements specifically. In the second (Figure 6B),

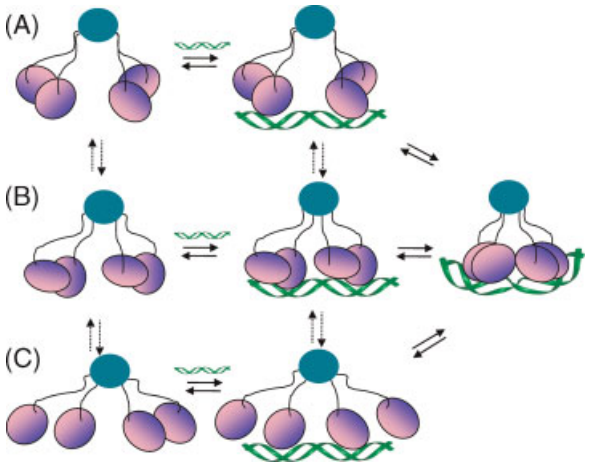

Figure 6. Possible scenarios for the binding mechanisms of p53 dimer of dimers. (A) p53 dimers exist in a form that needs significant conformational transformation before it can bind to its response elements specifically. (B) p53 dimers exist in a form that requires simple rotation of one domain to assume the DNA binding-ready conformation. (C) p53 exists in a conformation without significant interactions between the core domains, and the core domains bind to DNA individually before they form a specific dimerization interface between the core domains. The p53 DNA-binding domains and the tetramerization domains are shown as oval and circular shapes in pink and cyan and DNA is shown as double helix in green.

p53 dimers exist in an organization that requires simple rotation of one domain to assume the DNA binding-ready conformation. In the third scenario (Figure 6C), p53 exists in a conformation without significant interactions between the core domains. Each core domains bind to the DNA individually. Subsequently, they form a specific dimerization interface.

There are data supporting these scenarios. Previous experimental structural data revealed the conformational diversity of p53 dimers: the crystal structure from the Pavletich's group revealed a dimeric organization in which subunits A and B have extensive favorable interactions with each other (Cho *et al*., 1994). In this organization, one subunit was in specific contact with DNA while the other was not. Such a dimer would follow the first scenario and may need to overcome a large energy barrier. The DNA-free p53 dimer model of the Fersht group based on NMR data appears to follow the second scenario: when one core domain was bound to the DNA, the other p53 core domain appeared to simply rotate by 70° to bind DNA in the same specific manner without steric clashes and avoiding a huge energy penalty (Veprintsev *et al*., 2006) (Figure 6B). In the third scenario, prior to the binding to DNA, the two core domains of p53 may be present as monomers, without contact between them (Figure 6C). Upon binding DNA, they dimerize through the salt bridges involving the Glu180 and Arg181 residues. In this scenario, the dimerization assists in the stabilization of the bound state but does not help in the conformational search prior to the binding. In each of the scenarios, binding of the first p53 core domain is crucial for the overall recognition process. On the other hand, dimers with specific pre-organization can easily recognize the DNA and bind with high affinity via a conformation selection mechanism, with an equilibrium shift toward such a binding-ready conformation (Ma *et al*., 1999; Tsai *et al*., 1999a; Tsai *et al*., 1999b; Kumar *et al*., 2000; Boehr and Wright, 2008; Lange *et al*., 2008; Ma *et al*., 2002). Our results suggest that the existence of pre-organized p53 dimers with the binding surface readily available is important to the binding efficiency. The mechanism followed by dimeric p53 can be a similar cooperative process as observed in our monomeric simulations: binding non-specifically to the DNA, quickly drifting along the DNA and anchoring in the major groove.

Jiao *et al*. have shown the dynamic interactions between p53 and DNA via AFM experiments (Jiao *et al*., 2001). Their results indicate that p53 and DNA can make specific and non-specific interactions directly, through a diffusion-contact mechanism. When bound to the DNA, the p53 molecule can dissociate, flip and slide along the DNA chain. Several studies have suggested that diffusion-non specific contact-sliding along the major groove can be a general feature of DNA-binding proteins searching for their respective binding sites (Cao *et al*., 2009; Givaty and Levy, 2009; Halford, 2009; Tkacik and Bialek, 2009). The p53 sliding along DNA major groove is evidenced by its low friction energy (Tafvizi *et al*., 2008). Such a search process is likely to become faster when assisted by C-terminus interactions with DNA. While earlier experimental results argued that the specific role of the C-terminus is not fully understood (Jayaraman and Prives, 1995), several reports regarding the controversial role of the p53 C-terminus have now seemingly converged to the consensus opinion that it is a positive regulator, particularly in enhancing p53 sliding along the DNA track (McKinney *et al*., 2004). McKinney *et al*. have further pointed out that on its own, the C-terminus slides even more efficiently. Liu and Kulesz-Martin proposed that the role of the C-terminus is to recognize an altered DNA conformation upon DNA breakage (Liu and Kulesz-Martin, 2006). It is also likely that the C-terminus binds non-specifically so that the core domains are kept in close vicinity to the DNA, increasing the local concentrations of the protein.

Proteins have been shown to bind to DNA non-specifically but in a manner similar to specific binding (Iwahara *et al*., 2006). The p53 core domain also binds the DNA non-specifically during the search process. Our results show that when the p53 was bound initially at the minor groove, it was not stable and quickly moved to the major groove. Moving to the major groove seemed to be the preferred way for p53-DNA interaction, even when the DNA sequence is not specific for p53 binding. Since a non-specific interaction with DNA would not be as stable as the specific binding, the movement of the core domains along the DNA track is thermodynamically viable. And, in this scenario, residues Lys120 and Arg280 function as a sensor. Once they find the hydrogen bonding partners, they lock in. Our results yield insights into the preferred way of DNA recognition by the p53 core domain and provide details of the recognition process at the early stage of p53 binding to DNA.

Acknowledgements

This research was supported (in part) by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under contract number NO1-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government. This study

utilized the high-performance computational capabilities of the Biowulf PC/Linux cluster at the National Institutes of Health, Bethesda, MD (http://biowulf.nih.gov).

REFERENCES

Balagurumoorthy P, Lindsay SM, Harrington RE. 2002. Atomic force microscopy reveals kinks in the p53 response element DNA. *Biophys. Chem.* **101–102**: 611–623.

Balagurumoorthy P, Sakamoto H, Lewis MS, Zambrano N, Clore GM, *et al.* 1995. Four p53 DNA-binding domain peptides bind natural p53-response elements and bend the DNA. *Proc. Natl Acad. Sci. USA* **92**: 8591–8595.

Bargonetti J, Friedman PN, Kern SE, Vogelstein B, Prives C. 1991. Wild-type but not mutant p53 immunopurified proteins bind to sequences adjacent to the SV40 origin of replication. *Cell* **65**: 1083–1091.

Boehr DD, Wright PE. 2008. Biochemistry. How do proteins interact? *Science* **320**: 1429–1430.

Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, *et al*. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**: 187–217.

Cao XQ, Zeng J, Yan H. 2009. Physical signals for protein-DNA recognition. *Phys. Biol.* **6**: 36012.

Ceskova P, Chichger H, Wallace M, Vojtesek B, Hupp TR. 2006. On the Mechanism of Sequence-specific DNA-dependent Acetylation of p53: the acetylation Motif is exposed upon DNA binding. *J. Mol. Biol.* **357**: 442–44456.

Cherny DI, Striker G, Subramaniam V, Jett SD, Palecek E, *et al*. 1999. DNA bending due to specific p53 and p53 core domain-DNA interactions visualized by electron microscopy. *J. Mol. Biol.* **294**: 1015–1026.

Cho Y, Gorina S, Jeffrey PD, Pavletich NP. 1994. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science* **265**: 346–355.

Clore GM, Omichinski JG, Sakaguchi K, Zambrano N, Sakamoto H, *et al.* 1994. High-resolution structure of the oligomerization domain of p53 by multidimensional NMR. *Science* **265**: 386–391.

Dehner A, Klein C, Hansen S, Muller L, Buchner J, *et al*. 2005. Cooperative binding of p53 to DNA: regulation by protein-protein interactions through a double salt bridge. *Angew. Chem. Int. Ed. Engl.* **44**: 5247–5251.

el-Deiry WS. 1998. Regulation of p53 downstream genes. *Semin. Cancer Biol.* **8**(5): 345–357.

el-Deiry WS, Kern SE, Pietenpol JA, Kinzler KW, Vogelstein B. 1992. Definition of a consensus binding site for p53. *Nat. Genet.* **1**: 45–49.

Fields S, Jang SK. 1990. Presence of a potent transcription activating sequence in the p53 protein. *Science* **249**: 1046–1049.

Funk WD, Pak DT, Karas RH, Wright WE, Shay JW. 1992. A transcriptionally active DNA-binding site for human p53 protein complexes. *Mol. Cell Biol.* **12**: 2866–2871.

Givaty O, Levy Y. 2009. Protein sliding along DNA: dynamics and structural characterization. *J. Mol. Biol.* **385**: 1087–1097.

Halford SE. 2009. An end to 40 years of mistakes in DNA-protein association kinetics? *Biochem. Soci. Transact.* **37**: 343–348.

Ho WC, Fitzgerald MX, Marmorstein R. 2006. Structure of the p53 core domain dimer bound to DNA. *J. Biol. Chem.* **281**: 20494–20502.

Horvath MM, Wang X, Resnick MA, Bell DA. 2007. Divergent evolution of human p53 binding sites: cell cycle versus apoptosis. *PLoS Genet.* **3**: e127.

Ishimaru D, Maia LF, Maiolino LM, Quesado PA, Lopez PC, *et al*. 2003. Conversion of wild-type p53 core domain into a conformation that mimics a hot-spot mutant. *J. Mol. Biol.* **333**: 443–451.

Ishimaru D, Ano Bom AP, Lima LM, Quesado PA, Oyama MF, *et al*. 2009 Cognate DNA Stabilizes the Tumor Suppressor p53 and Prevents Misfolding and Aggregation. *Biochemistry* **26**: 6126–6135.

Iwahara J, Zweckstetter M, Clore GM. 2006. NMR structural and kinetic characterization of a homeodomain diffusing and hopping on non-specific DNA. *Proc. Natl Acad. Sci. USA* **103**: 15062–15067.

Jagelska EB, Brazda V, Pecinka P, Palecek E, Fojta M. 2008. DNA topology influences p53 sequence-specific DNA binding through structural transitions within the target sites. *Biochem. J.* **412**: 57–63.

Jayaraman J, Prives C. 1995. Activation of p53 sequence-specific DNA binding by short single strands of DNA requires the p53 C-terminus. *Cell* **81**: 1021–1029.

Jiao Y, Cherny DI, Heim G, Jovin TM, Schaffer TE. 2001. Dynamic interactions of p53 with DNA in solution by time-lapse atomic force microscopy. *J. Mol. Biol.* **314**: 233–243.

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

Kastan MB, Onyekwere O, Sidransky D, Vogelstein B, Craig RW. 1991. Participation of p53 protein in the cellular response to DNA damage. *Cancer Res.* **51**: 6304–6311.

Kitayner M, Rozenberg H, Kessler N, Rabinovich D, Shaulov L, *et al.* 2006. Structural basis of DNA recognition by p53 tetramers. *Mol. Cell* **22**: 741–753.

Klein C, Georges G, Kunkele KP, Huber R, Engh RA, *et al*. 2001a. High thermostability and lack of cooperative DNA binding distinguish the p63 core domain from the homologous tumor suppressor p53. *J. Biol. Chem.* **276**: 37390–37401.

Klein C, Planker E, Diercks T, Kessler H, Kunkele KP, *et al*. 2001b. NMR spectroscopy reveals the solution dimerization interface of p53 core domains bound to their consensus DNA. *J. Biol. Chem.* **276**: 49020–49027.

Kumar S, Ma B, Tsai CJ, Sinha N, Nussinov R. 2000. Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci.* **9**: 10–19.

Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KF, *et al*. 2008. Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* **320**: 1471–1475.

Liu Y, Kulesz-Martin MF. 2006. Sliding into home: facilitated p53 search for targets by the basic DNA binding domain. *Cell Death Differ.* **13**: 881–884.

Luo J, Li M, Tang Y, Laszkowska M, Roeder RG, *et al*. 2004. Acetylation of p53 augments its site-specific DNA binding both in vitro and in vivo. *Proc. Natl Acad. Sci. USA* **101**: 2259–2264.

Ma B, Kumar S, Tsai CJ, Nussinov R. 1999. Folding funnels and binding mechanisms. *Protein Eng.* **12**: 713–720.

Ma B, Shatsky M, Wolfson HJ, Nussinov R. 2002. Multiple diverse ligands binding at a single protein site: a matter of pre-existing populations. *Protein Sci.* **11**: 184–197.

MacKerell AD, Jr, Bashford D, Jr, Bellott M, Dunbrack RL, Jr, Evanseck JD, *et al*. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**: 3586–3616.

McKinney K, Prives C. 2002. Efficient specific DNA binding by p53 requires both its central and C-terminal domains as revealed by studies with high-mobility group 1 protein. *Mol. Cell Biol.* **22**: 6797–6808.

McKinney K, Mattia M, Gottifredi V, Prives C. 2004. p53 linear diffusion along DNA requires its C terminus. *Mol. Cell* **16**: 413–424.

McLure KG, Lee PW. 1998. How p53 binds DNA as a tetramer. *Embo J.* **17**: 3342–3350.

Nagaich AK, Appella E, Harrington RE. 1997a. DNA bending is essential for the site-specific recognition of DNA response elements by the DNA binding domain of the tumor suppressor protein p53. *J. Biol. Chem.* **272**: 14842–14849.

Nagaich AK, Zhurkin VB, Durell SR, Jernigan RL, Appella E, *et al*. 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. USA* **96**: 1875–1880.

Nagaich AK, Zhurkin VB, Sakamoto H, Gorin AA, Clore GM, *et al*. 1997b. Architectural accommodation in the complex of four p53 DNA binding domain peptides with the p21/waf1/cip1 DNA response element. *J. Biol. Chem.* **272**: 14830–14841.

Okorokov AL, Sherman MB, Plisson C, Grinkevich V, Sigmundsson K, *et al.* 2006. The structure of p53 tumour suppressor protein reveals the basis for its functional plasticity. *Embo J.* **25**: 5191–5200.

Pan Y, Nussinov R. 2007. Structural basis for p53 binding-induced DNA bending. *J. Biol. Chem.* **282**: 691–699.

Pan Y, Nussinov R. 2008. p53-Induced DNA bending: the interplay between p53-DNA and p53-p53 interactions. *J. Phys. Chem.* **112**: 6716–6724.

Pan Y, Nussinov R. 2009. Cooperativity Dominates the Genomic Organization of p53-Response Elements: A Mechanistic View. *PLoS Comput. Biol.* **7**: e1000448.

Petrovich M, Veprintsev DB. 2009. Effects of CpG methylation on recognition of DNA by the tumour suppressor p53. *J. Mol. Biol.* **386**: 72–80.

Poon GM, Brokx RD, Sung M, Gariepy J. 2007. Tandem dimerization of the human p53 tetramerization domain stabilizes a primary dimer inter-

mediate and dramatically enhances its oligomeric stability. *J. Mol. Biol.* **365**: 1217–1231.
Riley T, Sontag E, Chen P, Levine A. 2008. Transcriptional control of human p53-regulated genes. *Nat. Rev.* **9**: 402–412.
Rippin TM, Freund SM, Veprintsev DB, Fersht AR. 2002. Recognition of DNA by p53 core domain and location of intermolecular contacts of cooperative binding. *J. Mol. Biol.* **319**: 351–358.
Sauer M, Bretz AC, Beinoraviciute-Kellner R, Beitzinger M, Burek C, *et al.* 2008. C-terminal diversity within the p53 family accounts for differences in DNA binding and transcriptional activity. *Nucleic Acids Res.* **36**: 1900–1912.
Smeenk L, van Heeringen SJ, Koeppel M, van Driel MA, Bartels SJ, *et al.* 2008. Characterization of genome-wide p53-binding sites upon stress response. *Nucleic Acids Res.* **36**: 3639–3654.
Sun XZ, Vinci C, Makmura L, Han S, Tran D, *et al.* 2003. Formation of disulfide bond in p53 correlates with inhibition of DNA binding and tetramerization. *Antioxid. Redox Signal.* **5**: 655–665.
Sykes SM, Stanek TJ, Frank A, Murphy ME, McMahon SB. 2009. Acetylation of the DNA binding domain regulates transcription-independent apoptosis by p53. *J. Biol. Chem.* **284**: 20197–20205.
Tafvizi A, Huang F, Leith JS, Fersht AR, Mirny LA, *et al.* 2008. Tumor suppressor p53 slides on DNA with low friction and high stability. *Biophys. J.* **95**: L01–03.
Tidow H, Melero R, Mylonas E, Freund SM, Grossmann JG, *et al.* 2007. Quaternary structures of tumor suppressor p53 and a specific p53 DNA complex. *Proc. Natl Acad. Sci. USA* **104**: 12324–12329.
Tkacik G, Bialek W. 2009. Diffusion, dimensionality, and noise in transcriptional regulation. *Phys. Rev.* **79**: 051901.
Tsai CJ, Kumar S, Ma B, Nussinov R. 1999a. Folding funnels, binding funnels, and protein function. *Protein Sci.* **8**: 1181–1190.
Tsai CJ, Ma B, Nussinov R. 1999b. Folding and binding cascades: shifts in energy landscapes. *Proc. Natl Acad. Sci. USA* **96**(18): 9970–9972.
Veprintsev DB, Freund SM, Andreeva A, Rutledge SE, Tidow H, *et al.* 2006. Core domain interactions in full-length p53 in solution. *Proc. Natl Acad. Sci. USA* **103**: 2115–2119.
Vousden KH. 2002. Activation of the p53 tumor suppressor protein. *Biochim. Biophys. Acta* **1602**: 47–59.
Vousden KH, Lu X. 2002. Live or let die: the cell's response to p53. *Nat. Rev. Cancer* **2**: 594–604.
Wang Y, Rosengarth A, Luecke H. 2007. Structure of the human p53 core domain in the absence of DNA. *Acta Crystallogr. D Biol. Crystallogr.* **63**: 276–281.
Wang Y, Schwedes JF, Parks D, Mann K, Tegtmeyer P. 1995. Interaction of p53 with its consensus DNA-binding site. *Mol. Cell Biol.* **15**: 2157–2165.
Waterman JL, Shenk JL, Halazonetis TD. 1995. The dihedral symmetry of the p53 tetramerization domain mandates a conformational switch upon DNA binding. *Embo J.* **14**: 512–519.
Wei CL, Wu Q, Vega VB, Chiu KP, Ng P, *et al.* 2006. A global map of p53 transcription-factor binding sites in the human genome. *Cell* **124**: 207–219.
Xue Y, Wang S, Feng X. 2009a. Effect of Metal Ion on the Structural Stability of Tumor Suppressor Protein p53 DNA Binding Domain. *J. Biochem.* **146**: 193–200.
Xue Y, Wang S, Feng X. 2009b. Influence of Magnesium Ion on the Binding of p53 DNA Binding Domain to DNA Response Elements. *J. Biochem.*
Zeng J, Yan J, Wang T, Mosbrook-Davis D, Dolan KT, *et al.* 2008. Genome wide screens in yeast to identify potential binding sites and target genes of DNA-binding proteins. *Nucleic Acids Res.* **36**: e8.
Zhao K, Chai X, Johnston K, Clements A, Marmorstein R. 2001. Crystal structure of the mouse p53 core DNA-binding domain at 2.7 A resolution. *J. Biol. Chem.* **276**: 12120–12127.