Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

Sliding of p53 along DNA Can Be Modulated by Its Oligomeric State and by Cross-Talks between Its Constituent Domains

Netaly Khazanov and Yaakov Levy*

Department of Structural Biology, Weizmann Institute of Science, Rehovot 76100, Israel

The p53 protein is a homotetrameric transcription factor whose monomers comprise several domains. Although its organization with and without DNA was elucidated recently, characterizing the p53–DNA complex at the atomic level remains challenging because of its many disordered regions. Here we use computational models to predict the wiring of the four chains composing p53 and study its sliding dynamics along DNA in different oligomeric states. We find that helical sliding along the major groove is the most feasible DNA search mechanism for a large range of salt concentrations. Tighter packing of the tetrameric core domain is associated with a greater nonspecific affinity for DNA and the slowest linear diffusion dynamics along DNA. C-tails facilitate linear diffusion but restrict the association of two primary dimers into a tetramer. This restriction can disappear at higher salt concentrations, which decrease the affinity of C-tails for DNA, or upon interaction of the C-tail with other DNA segments. Our results support evidence for the positive regulation of p53 function by the C-tails and suggest that posttranslational charge modifications may alter the affinity of the tails for DNA. Conversely, the N-termini have little effect on sliding characteristics. Changes in the electrostatic potentials of the core domain via missense mutations corresponding to cancer development can also affect sliding by p53. Our study provides molecular insight into the role of various p53 domains during DNA search and indicates that the complex interdomain and protein–DNA cross-talks in which p53 engages may be related to its repertoire of cellular functions.

Introduction

The transcription factor protein p53 is critical for tumor suppression and, despite having been discovered about 30 years ago,[1] remains the center of attention in the fields of cancer research and tumor biology.[2] In response to genotoxic agents, radiation, stress, hypoxia, and other inappropriate signals, p53 initiates cell cycle arrest, DNA repair, or apoptosis, thus leading researchers to bestow on p53 the moniker "the guardian of the genome."[3–5] Loss of p53 activity (often through missense mutations) is a critical event in cancer development. p53 is a homotetramer in which each monomer is composed of about 400 residues (393 residues for human p53) and consists of four major domains: the DNA binding core domain (residues 98–303), the tetramerization (Tet) domain (residues 324–355), the N-terminal transcription activation domain at the N-terminus (residues 1–67), and the basic tail borne by the Tet domain at the C-terminus (C-tails; residues

---

Corresponding author. E-mail address:
koby.levy@weizmann.ac.il.

Abbreviations used: Tet, tetramerization; EM, electron microscopy; SAXS, small-angle X-ray scattering; 1D, one-dimensional; 3D, three-dimensional; TI, tetrameric interface; PDB, Protein Data Bank.

363–393)[6] (see Fig. 1). While the first two domains are folded, the N-terminus and C-tails are intrinsically disordered. The two folded domains (i.e., the core and Tet domains) are connected to each other by unstructured linkers of 20 residues, and both form a tetramer composed of two primary dimers (i.e., they can be defined as dimers of dimers).

The structures of the separated core and Tet domains in the presence and in the absence of DNA were determined by both X-ray crystallography and
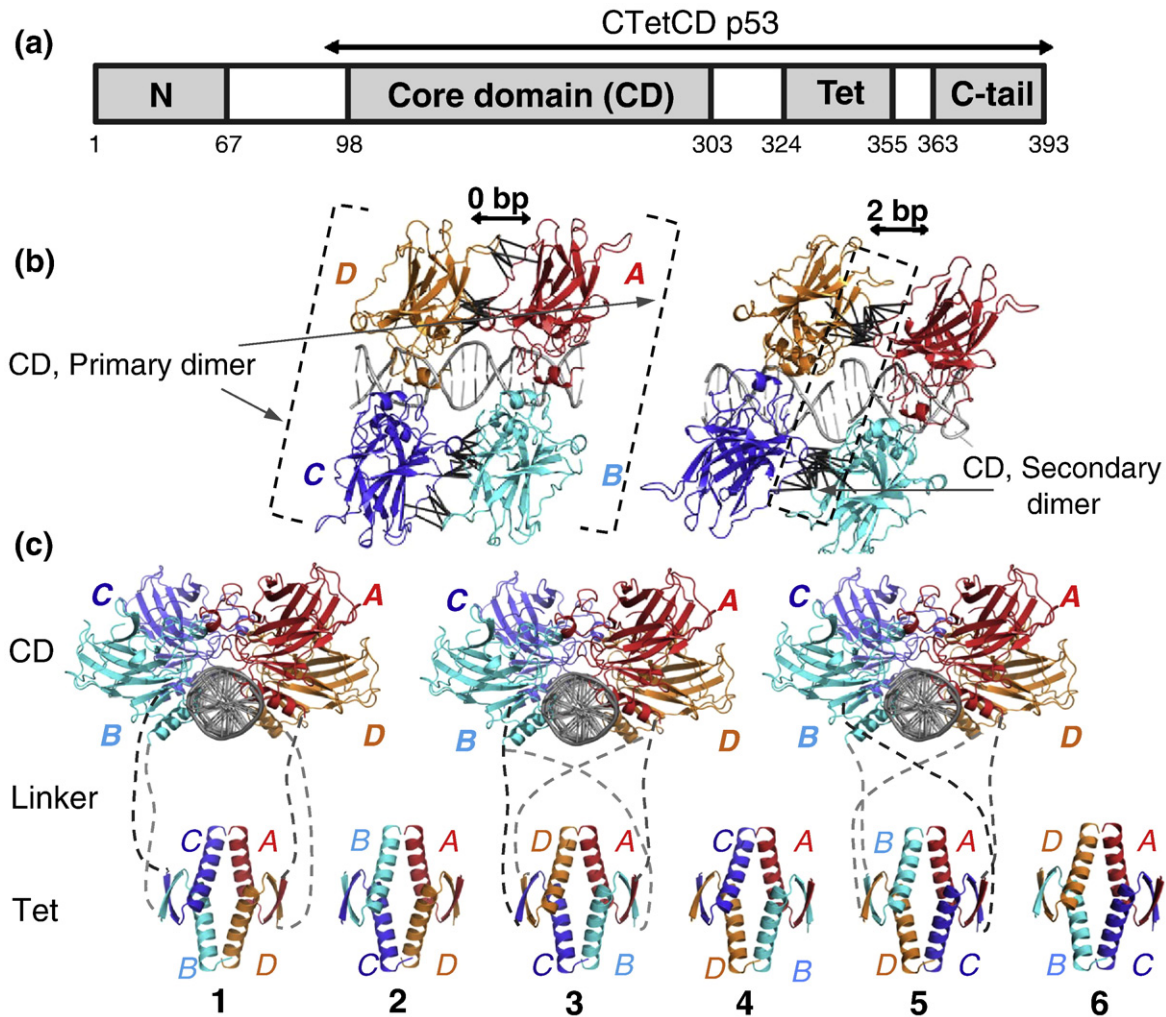

Fig. 1. Predicting the relative orientations of the Tet domain and the core domain (CD) of p53. (a) The four main domains that comprise the p53 protein: the transactivation domain (N; residues 1–67), the DNA binding CD (residues 98–303), the Tet domain (Tet; residues 324–355), and the C-terminus (C-tail; residues 363–393). (b) The crystal structures of a tetrameric CD (composed of monomers *A*, *B*, *C*, and *D*) with the two decameric half-sites separated by 0 bp (PDB ID: 3KZ8) and 2 bp (PDB ID: 2AC0). The tetramer, which is a dimer of dimers, is composed of two primary dimers (*AB* and *CD*; rectangles) that interact with separate response elements on the DNA and can themselves interact via the secondary interface. (c) Representation, for full-length p53, of the CD tetramer shown in (b) binding to a DNA molecule (gray), together with the six possible variants of the Tet domain that can be created by the four constituent monomers of the CD as they wrap around the DNA molecule. The orientation of the Tet domain relative to the CD depends on their connectivity, which is achieved by flexible linkers (shown for variants 1, 3, and 5). For example, in variants 1 and 2, the Tet dimers are created by monomers *A* and *D* or by monomers *B* and *C*, which are from different primary dimers of the CD (i.e., they interact with the same decameric half-site of the DNA) and are located on the same side of the DNA molecule. In variant 1, the major interface between the constituent dimers of the tetramer (i.e., the major TI) is formed between *A* and *C* (interface *AC*) or between *B* and *D* (interface *BD*), while in variant 2, the TI comprises *AB* and *CD*. In variants 3 and 4, the dimers that form the Tet domain are the primary dimers of the CD, but they are placed on opposite sides of the DNA molecule to create the TI variants *AD* and *BC* (variant 3) and *AC* and *BD* (variant 4). In variants 5 and 6, the Tet dimers are from different primary dimers of the CD and are placed on different sides of the DNA, with two possible TI variants: *AB* and *CD* (variant 5) and *AD* and *BC* (variant 6).

NMR methods.[7–14] The core domain adopts an Ig-like β-sandwich structure that provides a scaffold for recognizing DNA. The Tet domain forms a symmetrical tetramer made of two tight dimers stabilized by an anti-parallel β-sheet and helix–helix interactions (Fig. 1). Four core domains (one from each p53 monomer) interact with cognate DNA (comprising two decameric half-sites) as a dimer of dimers (the primary dimers are *AB* and *CD*; see Fig. 1). Genomic analysis showed that the two primary dimers can specifically interact with DNA response elements when spaced 0–20 bp apart.[15,16] The structure of the tetrameric core domain interacting with the cognate DNA sequence has been resolved recently for DNA with 0-bp or 2-bp spacers between the decameric half-sites and showed differences in the size of the secondary interface formed between the primary dimers.[12,17–19] The buried surface area in the secondary interface, which approximately doubles in size when the half-sites on the DNA are contiguous, also restricts the rotation of the dimers relative to each other.[17] We note that, while the crystal structure of p53 bound to contiguous half-sites showed tight binding to the DNA, a recent electron microscopy (EM) study revealed different binding modes for p53 and, in some of them, a loosely arranged tetrameric p53.[20]

It is difficult to crystallize full-length p53[21] or to determine its structure using NMR because of its low stability and because ~40% of the protein comprises intrinsically disordered regions.[22] Recently, the structure of full-length p53 was studied using cryo-EM of negatively stained samples and small-angle X-ray scattering (SAXS).[14,23] In the complex of full-length p53 with DNA, the interactions of the core domains with specific DNA response elements are very similar to those seen in the crystal structure of the isolated core domains with DNA. The EM model for the p53–DNA complex suggests that the DNA is located between the core domain and the Tet domain.[12,23] The EM structure supports an orientation of the Tet domain parallel with the DNA—an orientation that can be stabilized by interactions of the basic C-tails with the DNA. It was proposed that the Tet domain may instead lie perpendicular to the DNA axis.[12] By contrast, an effort to study the structure of full-length p53 with DNA using EM suggested that tetrameric p53 is formed through contacts between the core domain and the N-terminus/C-terminus, while the Tet domain is dissociated and the p53 molecule resembles a hollow skewed cube.[24]

Unlike other transcription factors, p53 contains two discrete DNA binding domains: the core domain and the C-tail of the Tet domain. Although both these domains bind tightly to DNA, they do so in very different ways. The central core domain contains the sequence-specific DNA binding domain. It is believed that cancer-related mutations (at

Arg175, Gly245, Arg248, Arg249, Arg273, and Arg282), which are mostly located within the core domain region that directly interacts with the DNA, occur at 'hot-spot' amino acids that affect the affinity of the core domain for the DNA.[25,26] By contrast, the basic C-tails are nonspecific DNA binding domains that bind with high affinity to a wide variety of DNA sequences via the electrostatic interactions of several lysine residues. Many studies aimed to address questions regarding the role of the C-tails in DNA recognition and p53 activity and their role in regulating the interactions of the core domain with DNA. Some research groups propose that this domain has no function in sequence-specific DNA binding,[27–29] whereas others suggest that the C-tails play the role of a negative regulator for binding[30,31] or promote linear diffusion on DNA.[32,33] Thus, the functional role of the C-tails remains unclear and is the subject of much controversy. However, advancing the understanding of the binding mode of full-length p53 with DNA should help to elucidate the matter.

Our approach to studying the role of various domains in controlling the regulatory activity of their protein is to study the cross-talk between constituent domains while the protein searches the DNA. After the facilitated diffusion model proposed by Berg *et al.*[34] and Vonhippel and Berg,[35] it has been well accepted that DNA search by a protein comprises different mechanisms: one-dimensional (1D) search (sliding), hopping, three-dimensional (3D) search, and intersegment transfer.[36–41] During sliding, the protein uses the same interface defined by its specific binding to perform a helical motion along the major groove.[42–47] In hopping, the protein performs multiple dissociation events over a short period of time during which the protein rebinds to the DNA in the vicinity of its dissociation point, while in 3D search, the protein reassociates at a point uncorrelated with the original location of the dissociation event. In intersegment transfer, the protein jumps between two segments of DNA that are separated by many sequences but geometrically proximate. Many mechanistic details of these DNA search modes and their interplay remain unclear,[45,46,48,49] yet their characterization has advanced in recent years, thanks to experiments at both the ensemble and the single-molecule level. The 1D sliding of a protein along DNA has been visualized[50] for diverse biological systems: RNA polymerase,[51] *lac* repressor,[52,53] DNA repair,[54] and p53 transcription factor.[55] Furthermore, different proteins search DNA using different strategies and search mode combinations. We have previously shown that the molecular characteristics of proteins can affect their sliding properties. The existence of a disordered tail in structurally homologous homeodomains can modulate their DNA search efficiency.[48] Similarly,

tethering domains by a flexible linker can also improve DNA search compared to that performed by the isolated domains.[48] The great structural complexity of p53, especially its high proportion of disordered regions, suggests that its inherent molecular features support its interaction with DNA.

A recent single-molecule study of p53 demonstrated that the only plausible mechanism for nonspecific search on DNA is 1D translocation.[55] While sliding of p53 was directly observed, the details at the molecular level and, in particular, the role of the various domains of p53 and its oligomeric state during interactions with DNA remain unclear. To address these questions, we have modeled full-length p53, as well as a truncated tetrameric p53 protein composed of the core domain, the Tet domain, and a C-tail (residues 94–393) following Tidow *et al.*, which we refer to as CTetCD.[23] On the basis of the predicted model of p53 and using coarse-grained molecular dynamics simulations of protein–DNA interactions, we visualized the linear diffusion of p53 along DNA and explored how this motion is supported by the unique molecular architecture of p53. We then investigated cross-talks between the p53 domains and quantified how they may be affected by changes in salt concentration, DNA organization, or point mutations in p53, and thus modulate DNA search.

Results

Predicted structure of the complex of full-length p53 with DNA

To study the dynamics of full-length p53 as it searches DNA, we first focused on predicting the structure and wiring (i.e., interaction pattern) of p53 with DNA. Figure 1 presents six possible variants of the Tet domain in a p53–DNA complex depending

on how the core and Tet domains are connected to each other via the linkers. Using the native-topology-based model, we characterized the thermodynamic and kinetic stabilities of the six different variants of CTetCD p53. During the simulations, which covered a wide range of temperatures, the core domain was kept bound to the DNA. These simulations included several folding/unfolding transitions and could capture the equilibrium ensemble at these simulated temperatures and the thermodynamic characteristics of each variant.

The thermodynamics of the six variants was explored by plotting the profiles of their folding free energy as a function of Q, which is the number of native contacts (Fig. S1). Q is a reaction coordinate characterized by three barriers: two barriers correspond to dimerization, and a third barrier corresponds to the formation of the Tet domain. The energetics of the assembly of the Tet domain was evaluated by plotting the specific heats of the six variants (Fig. S1). The folding temperature (T_f), at which the folded and unfolded conformations are equally populated (and, thus, the free-energy difference between the folded state and the unfolded state is zero), is represented by the peak of the specific heat curve. A larger value of T_f indicates greater thermodynamic stability, while a lower folding barrier indicates faster folding kinetics. Variant 1 has the smallest folding barrier and a relatively high T_f. Variant 2 has the largest T_f. Accordingly, the variants in which dimers are formed by monomers located at the same side of the DNA (*A* and *D*, or *C* and *B*; see Fig. 1b) have the greatest folding stability or the fastest kinetics.

The relative populations of the six variants were studied using numerous annealing simulations in which the Tet domains could fold to any of the six possible configurations. Variants 1 and 2 were obtained in 55% of the simulations (see Table 1). The variants in which the dimers were formed by two monomers from opposite sides of the DNA

Table 1. Results of annealing simulations using six possible variants of the Tet domain in the CTetCD model of p53 (in %) in which the core domain was modeled as a stable tetramer (i.e., with a secondary interface) or as two primary domains (i.e., without a secondary interface)

Variant	Dimeric interface	Major TI	With second interface[a]		No second interface[b]	
			% Annealing simulations	Total	% Annealing simulations	Total
1	**AD** and **BC**	**AC** and **BC**	24	55	12	39
2		**AB** and **CD**	31		17	
3	AB and CD	AD and BC	8	17	15	33
4		AC and BD	9		18	
5	AC and BD	AB and CD	16	28	14	28
6		AD and BC	12		15	

Entries in boldface indicate the most populated states obtained.
  [a] The annealing simulations performed in the presence of the tetrameric core domain utilized rigid core domains in which the primary dimers are spaced by 2 bp (using the coordinates of the domain as given in PDB ID: 2AC0).
  [b] The annealing simulations involved two primary dimeric core domains and utilized rigid domains separated by 12 bp.

were less populated: for example, variants 3 and 4, in which the monomers face each other from opposite sides of the DNA, were formed in 17% of the annealing simulations. Similarly, variants 5 and 6, in which the monomers are placed along the diagonal, were formed in 28% of the simulations. Between variants 1 and 2, the more stable and populated variant was variant 2 (31%), and this result is supported by its higher folding temperature. Thus, a tradeoff between the thermodynamic stability of variant 2 and the faster folding kinetics of variant 1 is obtained; notice that the only difference between these variants are the monomers involved in the tetrameric interface (TI) (i.e., the major interface between the two constituent dimers of the tetramer).

The core domain can also increase the heterogeneity of the structure of p53,[56,57] mostly by changing the size of the secondary interface between the two primary dimers. X-ray studies showed that the secondary interface is larger when the spacer between the two primary dimers is reduced from 2 bp to 0 bp. The secondary interface may disappear when the two primary dimers interact with response elements spaced, for example, by 10 bp. Although less than 20% of the response elements of p53 in the human genome are spaced more than 2 bp apart,[16,58] it is likely that the dimeric core domains search DNA while the spacer between them is large. To explore the structure of the p53 protein in terms of the organization of the Tet domain relative to the core domain, we performed annealing simulations of the six variants when the two primary dimers were placed 12 bp apart. We found that the linkers are sufficiently long to accommodate all six variants and, because of the larger distance between the primary dimers, there is a higher probability that the primary dimers of the Tet domain will be composed of subunits from the primary dimers of the core domain (Table 1).

We decided to continue our investigation into the interaction of p53 with DNA using variant 1, which has the lower folding barrier, but variant 2 may be equally as good. Variants 1 and 2 are similar in that the Tet domain is formed by monomers from different primary core domain dimers (i.e., by monomers that interact with different decameric response elements) located on the same side of the DNA molecule, and we assumed that this is the most dominant factor appropriate for the purpose of our research.

Sliding properties of p53 along DNA

The sliding of full-length and CTetCD p53 along DNA was studied using a computational model that we have previously used to explore the molecular details of the sliding of various transcription factors along DNA.[46,48,59] Recent experimental and theoretical studies[45–47,60] have indicated that, during sliding along DNA, many geometrical features of the protein–DNA interface are reminiscent of those found in the specific protein–DNA complex, which is not in contradiction with the conformation changes that are often observed when comparing the crystal structures of specific and nonspecific complexes.[61,62] The linear diffusion of p53 along DNA was detected recently by single-molecule experiments.[55] While these experiments very powerfully illuminate diffusion kinetics, their ability to provide insights into molecular phenomena (such as cross-talk between the domains of p53) is limited. To quantify the sliding mechanism of p53 and, in particular, the oligomeric state of the core domain in the mechanisms used to search DNA, we studied the sliding of three versions of the variant 1 (Fig. 1) of p53: a version in which the secondary interface in the core domain was fully formed (with a spacer of 0 bp), a version with a partially formed secondary interface (with a spacer of 2 bp), and a version in which the secondary interface was completely eliminated. The three versions were simulated for both full-length and CTetCD models of p53 at a wide range of salt concentrations using a coarse-grained model. In this model, the p53 molecule was entirely flexible, and its interactions with DNA were guided solely by electrostatic interactions.

To examine the configuration of the core domain bound to DNA as a dimer (i.e., without a secondary interface) and as a tetramer (i.e., with a secondary interface) during simulations at a low salt concentration, we compared the average distance between any bead of the core domain p53 and its closest DNA bead during sliding with the corresponding distance in the crystal of the specific complex of p53 with its cognate DNA. These profiles show a high degree of correlation between the binding mode of p53 with DNA during sliding and the binding mode of p53 in the specific complex for both full-length and truncated p53 (Fig. 2a). Similar characteristics for the sliding of the core domain of p53 were recently observed in atomistic simulations.[63] For the tetramer, we studied the sliding of CTetCD p53 with an extensive secondary interface (i.e., the two response elements are contiguous) and when the secondary interface is smaller (i.e., the two response elements are separated by 2 bp). The two variants of the CTetCD model of p53 slide using the same interface as that used for the specific interactions with DNA; nonetheless, the variant with the extensive secondary interface is closer to the DNA, and its sliding conformation is more similar to the specific binding than the variant with the smaller secondary interface (Fig. 2b). When the core domain was modeled without the secondary interface, a high level of similarity between nonspecific and specific interactions was still observed for full-length and CTetCD p53 (Fig. 2a), although the similarity
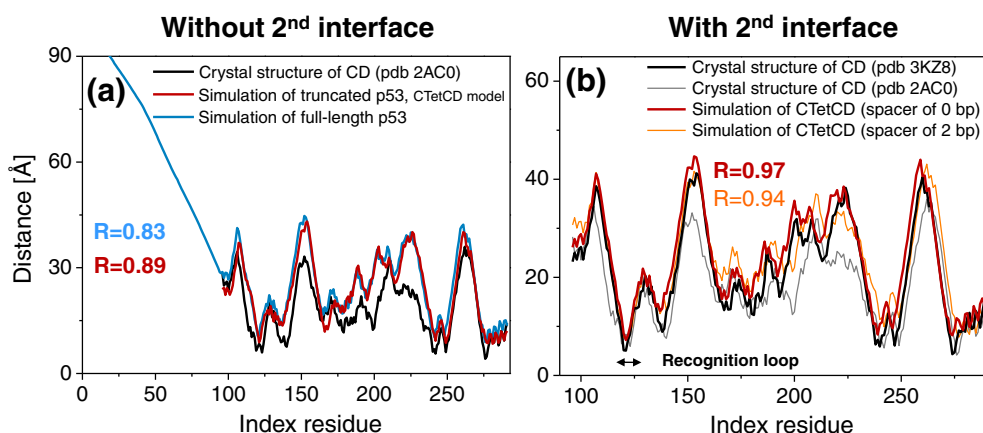
Fig. 2. Interactions of the core domain (CD) of p53 with DNA during sliding. The interaction of the CD with DNA (a) when formation of the secondary interface was eliminated or (b) when the secondary interface was formed with a spacer of 0 bp or 2 bp (see Fig. 1). The p53 transcription factor was modeled using the full-length sequence and with a truncated N-terminus (CTetCD model). The average distance between each protein bead (representing a residue) of the CD modeled using the full-length (blue) and truncated (CTetCD; red or orange) models and the closest DNA bead during the simulation sampled at a salt concentration of 0.01 M is compared to the distance in the crystal structure of the specific complex between the CD of p53 and its cognate DNA (black line). Correlation coefficients between each p53 model and the crystal structure of the CDs indicate that truncation did not influence the binding of the CD; however, reduction or elimination of the secondary interface has a larger effect.

was less than that observed using models that include the secondary interface.

To illustrate the sliding of full-length and CTetCD models of p53 along DNA, we present snapshots from the simulations (Fig. 3). The snapshots illustrate that full-length and truncated p53 follow a helical 1D path along the DNA (as shown elsewhere for the sliding of other DNA binding proteins[46,48]) both with and without the secondary interface. The C-tails, which are mostly positively charged (each C-tail includes nine positively charged residues and four negatively charged residues), engage in strong electrostatic interactions with the negatively charged phosphate groups of the DNA molecule. During sliding, the Tet domain is oriented parallel with the DNA, and the positioning of the Tet domain is supported by the interactions of the C-tails with the DNA (Fig. 3a and c). When the primary dimers of the core domain freely diffuse along the DNA (under the constraints imposed by the linkers, which may limit the diffusion of the core domain), the C-tails interact with the intervening DNA and further restrict the relative motion of the dimers (Fig. 3b and d). Formation of the secondary interface in this case will not occur solely by a linear diffusion of the dimers but will require the dissociation of the C-tails. Importantly, the N-termini (mostly negatively charged) in full-length p53 do not interact with the C-tails (mostly positively charged) but are exposed to the solvent. While the N-termini may affect the affinity of p53 for DNA, they seem not to affect the interaction of the core domain and C-tails with DNA. We therefore continued our study by focusing on the truncated

CTetCD model of p53 to decipher potential cross-talks between the core domain, the Tet domain, and the C-tails.

To better quantify the effect of the C-tails on the dynamics of tetrameric p53 along DNA, we repeated our study of the linear diffusion of the two versions of CTetCD p53 variant 1 (i.e., with a 0-bp or 2-bp spacer between the response elements) with electrostatically neutralized C-tails. In addition to studying tetrameric p53, we studied monomeric and dimeric p53, with the C-tail being either charged or neutralized.

Figure 4a (left) shows that tetrameric p53 (CTetCD model) mostly slides along the DNA and that the sliding propensity decreases slightly at higher salt concentrations. This result is in agreement with a recent single molecular experiment that leaves 1D search as the only probable mechanism for nonspecific search on DNA.[55] Truncating the N-termini results in a greater sliding propensity, presumably because the N-termini are repelled from the DNA and their truncation increases the affinity of p53 for DNA. Truncating the C-tails or neutralizing their charged residues scarcely affects the sliding propensity of p53, and these variants even have a slightly higher sliding propensity than wild-type CTetCD p53. The sliding propensity of p53 is lower when the core domain does not diffuse as a tetramer but as two independent dimers. A lower sliding propensity is observed for monomeric and dimeric p53, which also perform hopping dynamics along the DNA or 3D diffusion in the bulk. Accordingly, a higher-order oligomeric core domain correlates with an increased nonspecific affinity for the DNA and, consequently, the core domain more tightly interacts
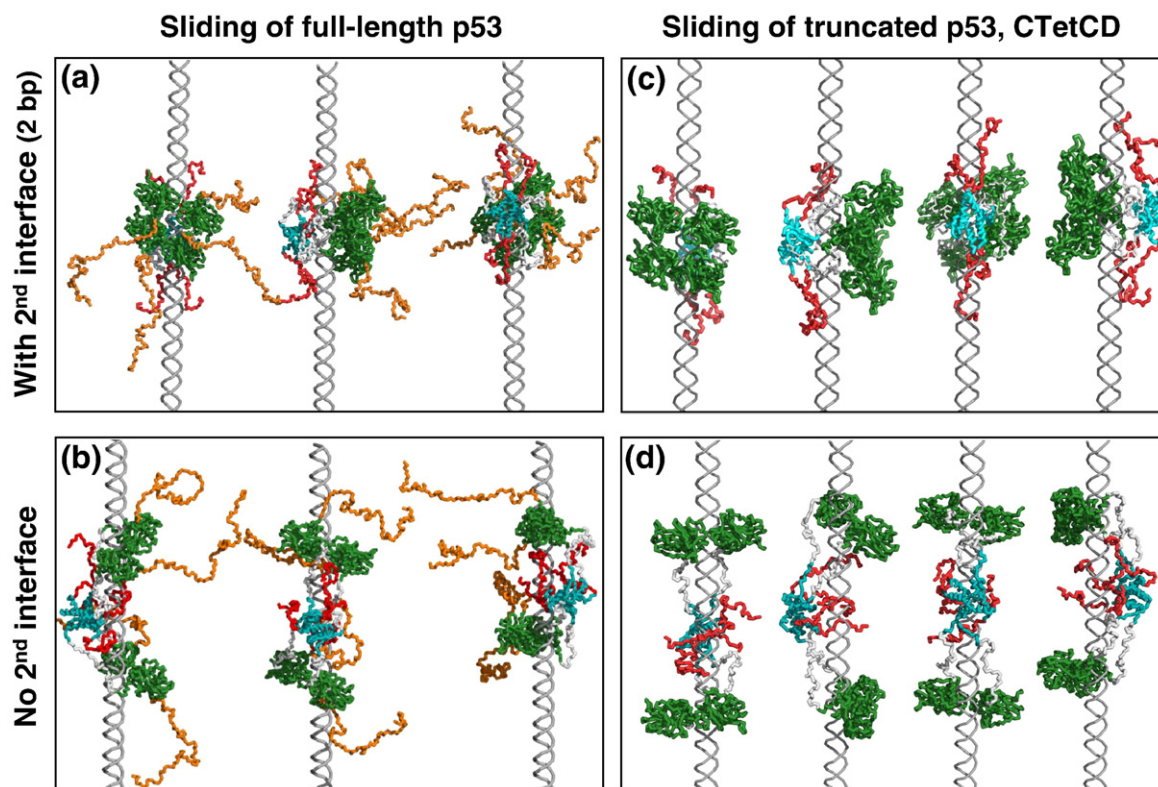
Sliding of full-length p53 **Sliding of truncated p53, CTetCD**

Fig. 3. Conformations of p53 during sliding along DNA. Snapshots of the sliding of (a and b) the full-length model of p53 and (c and d) the CTetCD model of p53 along DNA. (a and c) The two primary dimers diffuse after the formation of the secondary interface (spaced by 2 bp) in the core domain (i.e., as a tetramer). (b and d) The dimers of the core domain diffuse without the secondary interface (i.e., as two dimers). The N-termini are shown in orange, the core domain is shown in green, the linkers are shown in gray, the Tet domain is shown in cyan, and the tails are shown in red. The snapshots illustrate that p53 performs a helical motion along the DNA in all models. Tails are positively charged (nine positively charged residues); thus, during sliding, they tightly interact with the negatively charged phosphate of the DNA. The C-tails are stretched along the DNA when the core domain diffuses as a tetramer, but they are placed in between the two primary dimers when the secondary interface is not formed. The parallel orientation of the Tet domain is determined by the position of the positively charged tails closest to the DNA.

with the DNA. The higher sliding propensity of tetrameric p53 compared with its monomeric and dimeric forms is reflected in the 1D diffusion coefficient D_1 (Fig. 4a, right). The D_1 of tetrameric p53 with a secondary interface in the core domain is relatively low, and this could be rationalized by the larger interface between p53 and DNA. The D_1 is larger for dimeric p53 and much higher for the monomeric variant. The D_1 values of tetrameric and dimeric p53 are almost independent of salt concentration, unlike that of the monomeric variant[48] (Fig. 4a, right), and this can be explained by the small fraction of hopping at low and high salt concentrations. Neutralization of the C-tails of all oligomeric variants of p53 results in a slower linear diffusion, indicating that C-termini promote the motion of p53 on the DNA. Accordingly, the charged tails slightly increase searching via hopping (at the expense of sliding) and thus increase the sliding diffusion coefficient, as was also observed experimentally.[64] While the oligomeric state, the C-tails, and the

secondary interface affect the value of D_1, the N-termini have a minor effect, if any.

The efficiency of DNA search by the core and Tet domains in a CTetCD model of p53

The p53 protein has two distinct domains—the core domain and the positively charged C-tails of the Tet domain—that can independently interact with DNA. This unique feature distinguishes p53 from many other DNA binding proteins, which have a single domain that interacts with DNA. To better understand the cross-talk between these two domains when searching DNA, we compared the efficiencies of their DNA searches. Both domains can slide on DNA, and Fig. 4b (left) indicates that the core domain performs mostly sliding, but the propensity of the Tet domain to slide via the flexible C-tails is smaller and decreases more sharply with increasing salt concentration. The higher propensity of the core domain to slide is accompanied by its much slower
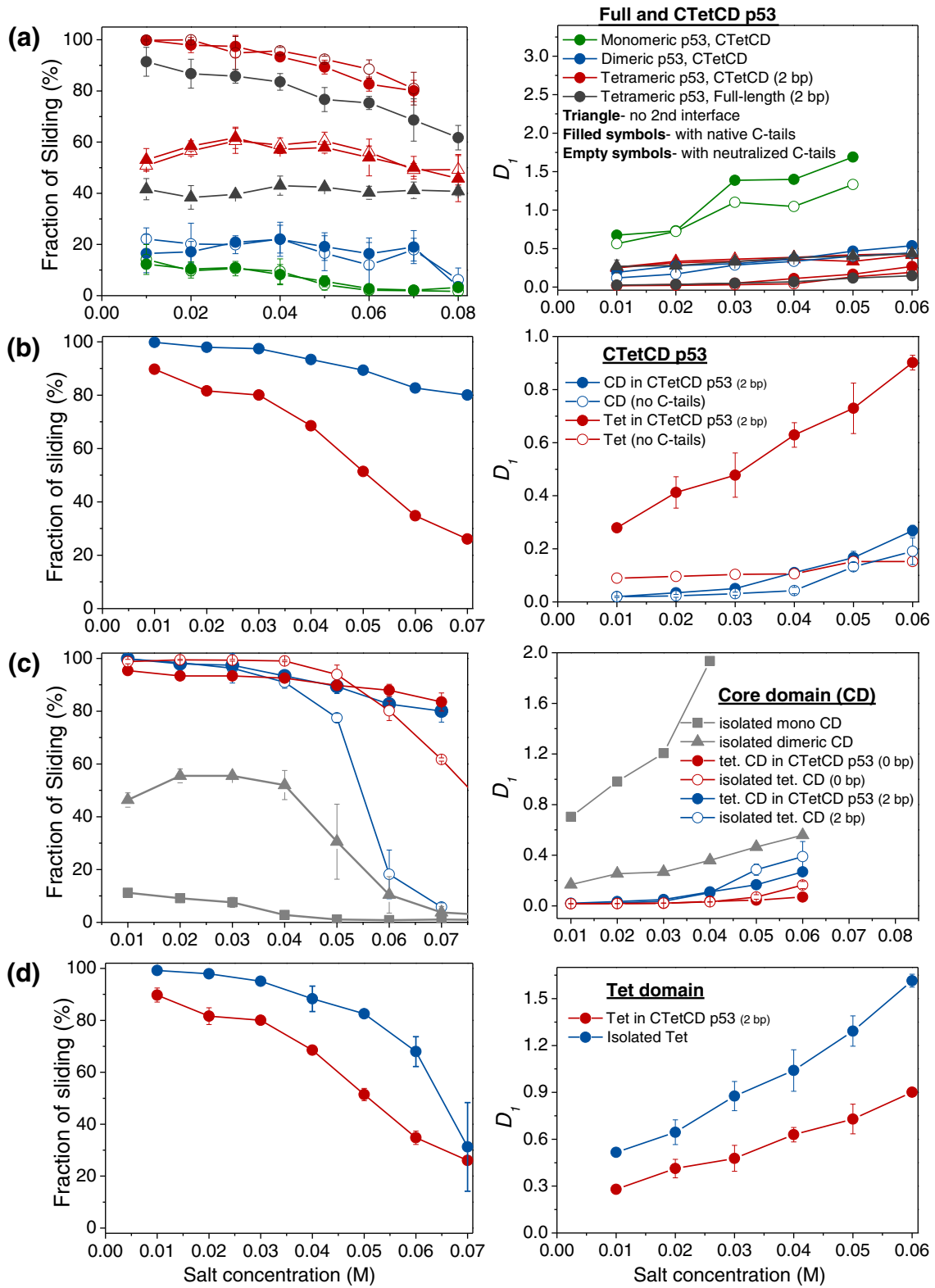
Fig. 4 (*legend on next page*)

linear diffusion compared to the Tet domain (Fig. 4b, right). The tetrameric core domain cannot dissociate from the DNA because of structural restrictions introduced by the connectivity of p53. By contrast, the positively charged tails are much more sensitive to changes in salt concentration; they dissociate from the DNA as salt concentration increases (0.06 M; Fig. 4b), and they mostly hop along the DNA (i.e., the tails remain close to the DNA but are not tightly bound at the major groove). In the absence of the C-tails, the Tet domain does not interact with DNA (Fig. 4b, right), and its linear diffusion is dictated solely by its linkers to the core domains.

Sliding of the core and Tet domains: Tethered *versus* isolated forms

To better understand the cross-talk between the core domain and the Tet domain, we studied their nonspecific interactions with DNA at different salt concentrations in two forms: as part of CTetCD p53 (tethered form) and as isolated domains. The isolated core domain was studied in monomeric, dimeric, and tetrameric forms. Furthermore, the tetrameric form was studied when the secondary interface was large or small (corresponding to a spacer between the DNA response elements of 0 bp and 2 bp). Examining first the forms of the core domain (Fig. 4c), we observed that the dynamics of the isolated core domains along DNA is more sensitive to salt concentration than the tethered core domain for both tighter and weaker secondary interfaces. Up to a salt concentration of 0.04 M, there is no significant difference between the isolated tetrameric core domain and the tethered tetrameric core domain as both mostly slide along the DNA, with the core domain performing a helical motion during which the interface between the tetrameric core domain and the DNA remains highly similar to the interface of the specific complex. The size of the secondary interface affects sliding speed. The core domain with the tight secondary interface slides more slowly than the variant with the smaller interface as it interacts more tightly with the DNA

(see Fig. 2b). Due to a weaker nonspecific affinity for the DNA, the monomeric and dimeric core domains perform less sliding dynamics even at low salt concentrations (Fig. 4c). With increasing salt concentration, all oligomeric forms of the isolated core domain search DNA using hopping (resulting in a faster linear diffusion) and especially 3D diffusion in the bulk, which is not an option for the tethered core domain since it is linked to the Tet domain. The combination of sliding, hopping, and 3D diffusion results in a more efficient DNA search by the isolated forms of the core domain than the tethered tetrameric variant at most salt concentrations.

We then sought to distinguish and characterize the capability of the Tet domain to search DNA by its C-tails as an isolated domain and when tethered to the core domain (Fig. 4d). Interestingly, the isolated Tet domain can linearly diffuse along DNA. When the Tet domain is tethered to the core domain, its sliding is clearly restrained (Fig. 4d, left) by the tight binding of the core domain to the DNA and by the fact that the core domain slides continuously in a helical fashion along the DNA. When the Tet domain diffuses in the absence of geometrical constraints imposed by the core domain and linker, its translocation along the DNA is faster because of the great flexibility of the tail. Translocation can still be performed in a helical fashion, but less continuously than is observed in CTetCD p53. Indeed, the linear diffusion coefficient is higher for the isolated Tet domain than for its tethered counterpart (Fig. 4d, right). The isolated Tet domain also utilizes sliding more than the tethered Tet domain (Fig. 4d). We may speculate that the linkers connecting the core and Tet domains affect the sliding propensity of the Tet domain and that, in their absence, the Tet domain can better interact with the DNA and also move faster on the DNA.

Relative orientation of the Tet and core domains during sliding is mediated by disordered C-tails

In addition to determining the wiring of p53 around the DNA and the positions of the linkers that

Fig. 4. Dynamics of sliding and linear diffusion coefficient for different oligomeric states of p53 and for its isolated domains at various salt concentrations. The effect of the oligomeric state and the composition of p53 on (left) the fraction of the search undertaken by sliding (note that the protein can search also by hopping and 3D diffusion) and on (right) the linear diffusion coefficient D_1. (a) Full-length and CTetCD p53 with native C-tails (filled circles) and neutralized C-tails (empty circles) in its monomeric, dimeric, and tetrameric forms. For the tetrameric form, the results are shown for p53 with and without the secondary interface in the core domain. The linear diffusion coefficient of monomeric p53 is most sensitive to salt concentration, but neutralizing the C-tails affects the D_1 of most variants. (b) The core domain (blue line) and the Tet domain (red line) of CTetCD p53 in which the core domain is modeled with the secondary interface. The interaction of the Tet domain with DNA is achieved via the C-tails. Note that even at a high salt concentration, the core domain cannot dissociate from the DNA (it mostly performs sliding) because of structural constraint. (c) The isolated core domain as a monomer (squares), dimer (triangles), and tetramer (circles). The tetrameric core domain was studied in its isolated form or tethered (i.e., as part of CTetCD p53) as a function of salt concentration. Two models of CTetCD p53 were tested: with a large secondary interface (2 bp) or a small secondary interface (0 bp) between the dimeric core domains, which are shown in red and blue, respectively. (d) Isolated Tet domain (blue) and Tet domain as part of CTetCD p53 (red) as a function of salt concentration.
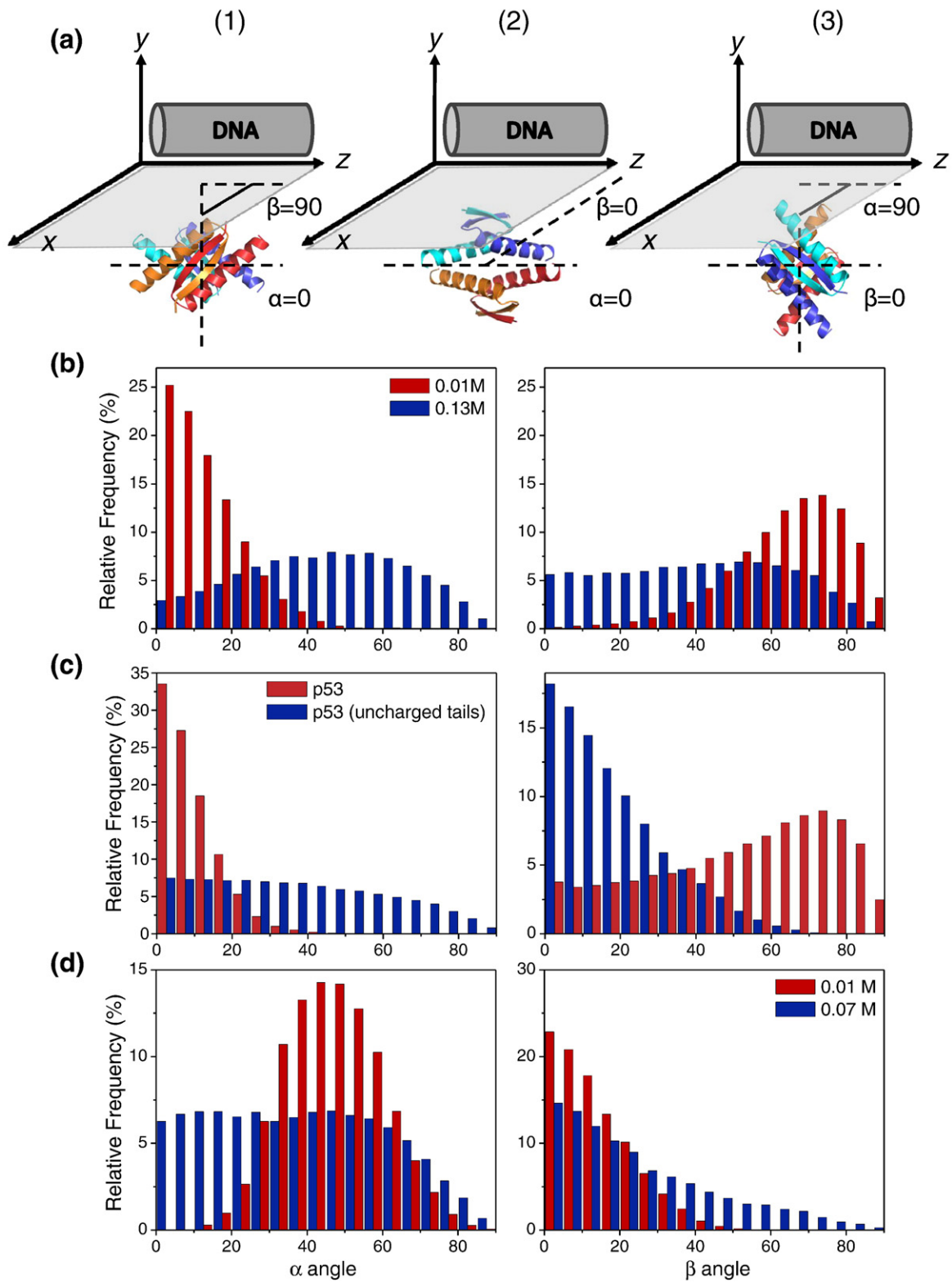
Fig. 5 (*legend on next page*)

connect the core and Tet domains, it is necessary to characterize the orientation of the Tet domain relative to the DNA molecule. Two alternative orientations of the Tet domain relative to the DNA molecule have been proposed.[12,22] In one model, the Tet domain is perpendicular to the DNA; in the other model, which is supported by a recent EM study,[23] the Tet domain is parallel with the DNA. To provide a detailed understanding of the relative orientations of the core and Tet domains when they are bound to DNA, we defined two angles, α and β, between the Tet domain and the DNA plane (Fig. 5a). There are three possible combinations for the arrangement of the Tet domain in terms of these angles. In two combinations, the Tet domain lies parallel with the DNA: $\alpha = 0°$ and $\beta = 0°$ or 90° (see Fig. 5a). In the third variant, the Tet domain lies perpendicular to the DNA ($\alpha = 90°$ and $\beta = 0°$). The values of the α and β angles of variant 1 (when the core domain is predominantly bound on DNA) at low (0.01 M) and high (0.13 M) salt concentrations are shown in Fig. 5b. At a low salt concentration, the angle α is about 0°, and the angle β is about 70°, indicating an arrangement of the Tet domain parallel with the plane of the DNA. At a high salt concentration, however, a wide distribution of angles is obtained, indicating loss of structural specificity.

The effect of the neutralization of the positive charges of the C-tails on the orientation between the Tet domain and the DNA plane is similar to the effect of increasing the salt concentration (Fig. 5b and c). In a model with no positive charges on the C-tails, angle α has a wider distribution of values that obscures the orientation of the Tet domain. To confirm the interactions of the charged tails with the DNA, we measured the distance between the DNA molecule and the center of mass of the charged and neutralized tails during sliding. While the natural (charged) tails are very close to the DNA (both with and without the secondary interface in the core domain; see Fig. 3), the tails with no positive charges

show a wider distribution of distances from the DNA (the distance between the DNA molecule and the center of mass of the charged and neutralized tails during sliding is 10 ± 2 Å and 50 ± 15 Å, respectively). The orientation of the Tet and core domains in the presence of two DNA molecules (Fig. 5d) is described later.

The effect of DNA concentration on the structure and dynamics of p53

The p53 protein does not necessarily interact with a single stretch of DNA segment. It can interact with a more complex conformation of DNA and, indeed, it is known that p53 is involved in interactions with histones, which are involved in DNA packing. In the crowded environment of the DNA, the tails can be envisaged interacting with different DNA segments, and it seems likely that DNA recognition by p53, and therefore its function, will be altered accordingly. To study the capability of the C-tails to interact with different DNA segments, we modeled the dynamics of CTetCD p53 along two parallel DNA molecules separated by a distance of 100 Å when the core domain diffuses as a tetramer or as two primary dimers that are connected via the Tet domain (i.e., with or without the secondary interface).

The effect of the additional DNA molecule on the translocation of p53 along the two DNAs at low and high salt concentrations can be observed in Fig. 6. For p53 with a secondary interface, the presence of an additional DNA molecule significantly slows translocation along the DNA, such that the p53 molecule with a secondary interface is anchored in one place at a low salt concentration. When the core domain of p53 is tightly packed (has a large secondary interface), its diffusion is very slow and inhibited in the presence of two DNA molecules (Fig. 6). The slower diffusion of p53 on two DNA molecules, compared to sliding on a single DNA, is also observed when the two primary dimers of the core domains diffuse independently (although their

Fig. 5. Orientation of the Tet domain of full-length p53 to the DNA plane at different salt concentrations. (a) The orientation of the Tet domain relative to the DNA plane is defined by the angles α and β, which describe three different orientations of the Tet domain with respect to the xz plane of the DNA. In cases 1 and 2, the Tet domain adopts an orientation parallel with the plane of DNA, while in case 3, the Tet domain lies in an orientation perpendicular to the DNA. (b) The effect of salt concentration on the distribution of orientation angles α and β during the sliding of p53 along DNA. In these simulations, the core domain was static, and only the linkers and the Tet domain, with its flexible tails, were fully dynamic. At a low salt concentration (0.01 M), the Tet domain lies in an orientation parallel with the plane of the DNA (case 1), but at a high salt concentration (0.13 M), a wide distribution of α and β angles, indicative of a nonspecific orientation of the Tet domain to the DNA, is observed. (c) The effect of the C-tails on the orientation of the Tet domain to the DNA during sliding. During sliding, the Tet domain of the CTetCD p53, which was entirely flexible, lies parallel with the DNA (angel α ~0°). In the model with no positive charges on the C-tails, the Tet domain has a wide distribution of angle α, while angle β is close to zero (i.e., the Tet domain loses its specific orientation with respect to the DNA). (d) Sliding of p53 in the presence of two parallel DNA molecules. At a salt concentration of 0.01 M, the distribution of angle α centers on ~50°, indicating an orientation of the Tet domain perpendicular to the DNA plane. A wider distribution of angle α is observed at a salt concentration of 0.07 M. The distributions of angle β (~0°) are similar in the two salt concentrations.
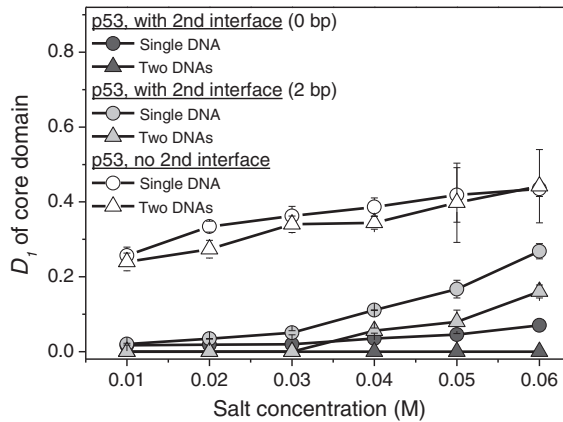
Fig. 6. Influence of an additional DNA molecule on the translocation of p53 along the DNA. The linear diffusion coefficient of the core domain (CD) when it moves on a single DNA molecule (circles) or on two DNA molecules (triangles) is shown for various salt concentrations. The CD is modeled with a large secondary interface (0 bp; dark-gray symbols), with a smaller secondary interface (2 bp; light-gray symbols), and without it (empty symbols).

sliding is much faster than that of tetrameric p53), suggesting that the interactions of the tails with different DNA regions restrict the dynamics of the core domains.

To better compare the dynamics of p53 on a single DNA molecule and the dynamics of p53 on two DNA molecules, we calculated the average distance of the tail of p53 lacking the secondary interface in the core domain from a DNA molecule (Fig. 7). When p53 slides along a single DNA, the average distance of the C-tails from the DNA increases with salt concentration, indicating that the tails are less bound to the DNA because of their weaker attraction to it. Consequently, the core domains diffuse faster (Fig. 6), and the separation distance between the two primary dimers of the core domain becomes smaller (Fig. 7b). When the diffusion occurs on two DNA molecules, the average distance of the tail from the DNA is larger at all salt concentrations as; at low salt concentrations, the tail can be attracted by both DNA molecules (and, therefore, the average distance to a given DNA molecule is larger), while at higher concentrations, the tail is less attracted to any DNA. Interestingly, the separation distance between the two primary dimers of the core domain is smaller when sliding takes place on two DNA molecules than on a single DNA molecule even at a low salt concentration. The interactions of the C-tails with other DNA enable the two primary dimers to come closer (Fig. 7c). The snapshots in Fig. 7c illustrate the interplay between the interactions of the C-tails with the DNA and the separation distance between the primary dimers of the core domains.

When p53 interacts with two DNA molecules at a low salt concentration, the values of angles α and β center around 50° and 0° (Fig. 5d), respectively, indicating an orientation of the Tet domain perpendicular to the DNA. The snapshot in Fig. S3 illustrates the interactions between each pair of tails and different DNA molecules, as well as the transition of the orientation of the Tet domain from parallel (Fig. 3) to perpendicular (Fig. 7c) when the p53 molecule is in a higher DNA concentration. Note that the distribution of angle α is around 50° and not around 90°, as one would expect, since the Tet domain is very flexible and fluctuates greatly. While angle α is widely distributed at a salt concentration of 0.07 M, there is little difference in the distribution of angle β ($\sim$0°) between the two salt concentrations (Fig. 5d).

The effect of missense mutations on the sliding dynamics of p53

Most missense mutations of p53 are clustered in the core domain, where they are often involved in DNA interactions. We selected the most frequent mutations that occur at the p53–DNA interface and involve the neutralization of a positive charge or its reversal to a negative charge. The mutational hot spots were checked for the CTetCD model of p53 with the secondary interface in the core domain by neutralizing a positively charged residue (Lys120, Arg248, Arg273, Arg280, or Arg283; see Fig. 8a for the location of these residues at the p53–DNA interface) in each monomer of p53. All the mutations affect the sliding dynamics of p53. The linear diffusion of all the mutants is faster than that of the wild-type core domain (Fig. 8b). The faster diffusion correlates with a lower sliding propensity but a higher hopping propensity (Fig. 8c). Similar qualitative results were observed when the mutations were studied using the CTetCD model of p53 without the secondary interface in the core domain and for an isolated tetrameric core domain.

Discussion

The functions and regulatory activity of p53 clearly depend on its interaction with DNA. The unique molecular characteristics of p53 as a tetrameric transcription factor—in which each subunit is composed of about four domains that include many disordered regions and where the tetramer possesses two DNA binding domains—make its specific and nonspecific interactions with DNA quite heterogeneous, since both protein–protein and protein–DNA interfaces can adopt various forms.[2,14,20,65,66]

While there is no known structure at the atomic level of resolution of full-length p53 in complex with DNA, the structure of the core domain–DNA
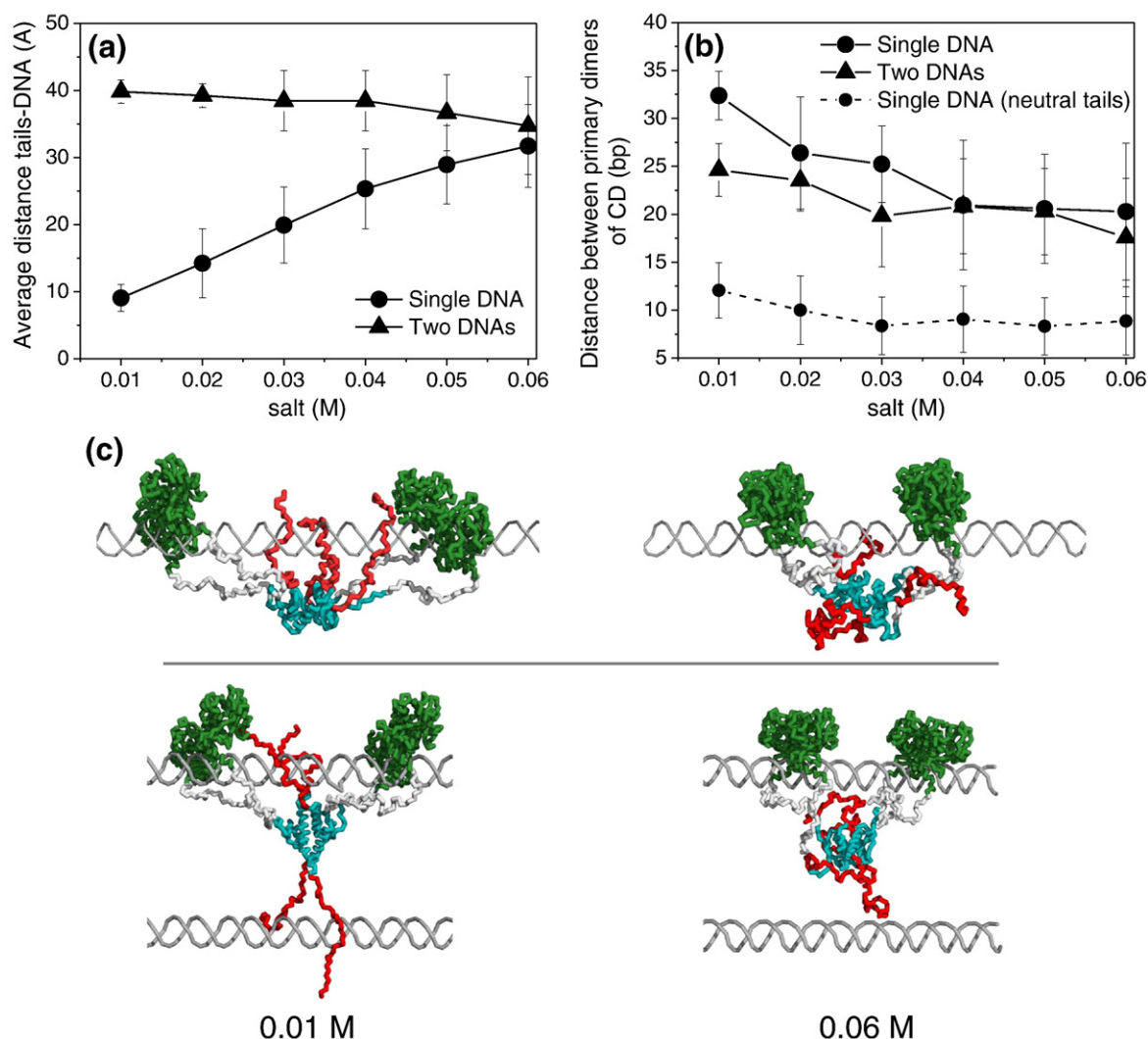
Fig. 7. The interplay between the location of the C-tails on DNA and the separation distance between the primary dimers of the core domains (CDs). (a) The average distance of the C-tails of p53 (with no secondary interface within the CD) from a DNA molecule when sliding takes place on a single DNA molecule (circles) or on two DNA molecules (triangles) at various salt concentrations. (b) The separation distance between the centers of mass of the two primary dimers of the CD while it slides on a single DNA molecule or on two DNA molecules. As control, the separation distance is also shown for a case in which the C-tails are neutralized electrostatically. (c) Snapshots of the sliding of p53 at salt concentrations of 0.01 M and 0.06 M illustrating the interactions of the tails with DNA and the separation distance between the dimers.

complex has been resolved recently by X-ray crystallography[7,12,13,18,19] and in solution in the absence of DNA by NMR.[10] The structure of the Tet domain has been resolved by both NMR and X-ray crystallography.[8,9,11] The EM and SAXS determinations of the structure of full-length p53 with and without DNA response elements provide valuable insights into the organization of the tetramer.

This study explored the likely wiring of p53, the mechanism by which it searches DNA, and how its oligomeric state and potential cross-talks between the domains may affect its nonspecific interaction with DNA. It adds to the existing body of

knowledge by doing so at the atomistic resolution required to describe the various intradomain and interdomain interfaces in the oligomeric state of p53 and to detail the interfaces it forms with DNA. The results demonstrate that the variants with the lowest free-energy barrier to folding or the highest thermodynamic stability are those in which the primary dimers constituting the Tet domain are formed by core domain monomers located on the same side of the DNA (i.e., that interact with different response elements). Sliding in a helical motion while interacting with the DNA major groove is the only feasible DNA search mechanism for a relatively
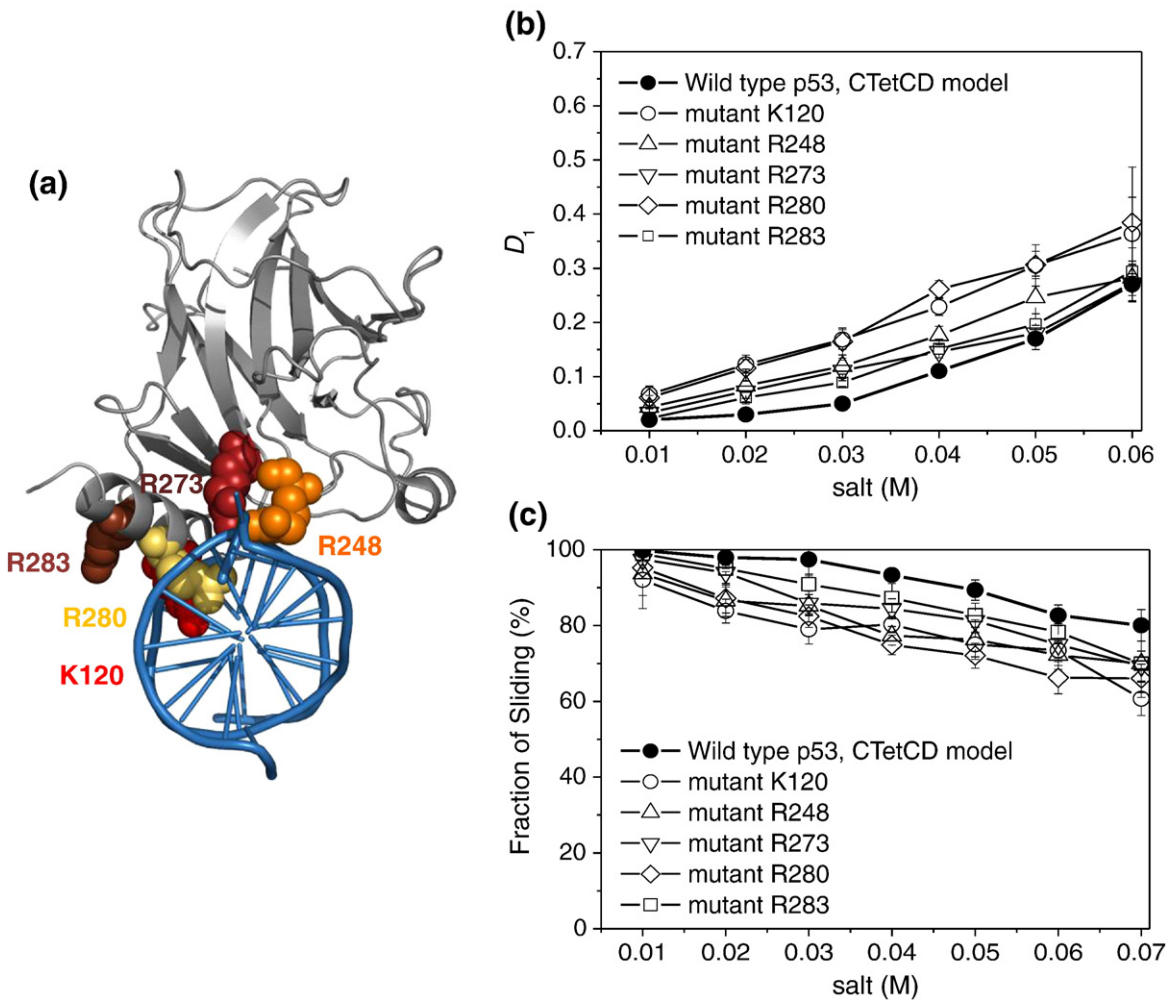
Fig. 8. The effect of missense mutations on the nonspecific interaction of p53 with DNA. (a) The positions of the studied missense mutations (K120, R248, R278, R280, and R283), highlighting that they are placed at the interface that p53 forms with DNA. (b) The diffusion coefficients of the mutants of p53 (modeled as CTetCD, with a secondary interface spaced at 2 bp in the CD). (c) The sliding propensity of the studied p53 mutants.

large range of salt concentrations when the core domain diffuses as a stable tetramer, supporting a recent single-molecule experiment of the sliding of p53 along DNA.[55]

Since p53 can interact with two DNA response elements spaced by 0–20 bp, we studied the interplay between the size of the secondary interface between the two primary dimeric core domains and sliding speed. The sliding of tetrameric p53 was studied with large, small, and no secondary interface, corresponding to a spacer of 0 bp, 2 bp, and >10 bp between the response elements. We found that increasing the packing of p53 (i.e., increasing the secondary interface) slows its sliding speed. A p53 with a larger secondary interface has a higher nonspecific affinity for the DNA since the dimeric core domain is better oriented and, therefore, its electrostatic attraction to the DNA is larger, permit-

ting a more committed sliding along the major groove. Clearly, monomeric and dimeric p53 slide much faster than the tetrameric variant, since the absence of the strong topological constraint introduced by the Tet domain in the tetrameric p53 enables hopping, which can accelerate sliding kinetics. One can envisage how, for p53, being in a lower-order oligomeric state enhances DNA search efficiency. Monomeric and dimeric p53 lack the packed secondary interface, and better orientation with respect to the DNA, tighter binding, greater friction, and slower sliding associate with tetrameric p53. Thus, one may speculate that the search for the DNA binding site is conducted by dimeric p53, with formation of the secondary interface to create a p53 tetramer occurring only at a later stage.

The N-termini, which are negatively charged and essential for the functioning of p53, interact with

neither the protein nor the DNA, and therefore have little effect on the sliding mechanism. The C-termini, which are positively charged and thus strongly interact with DNA, can affect sliding features very significantly. DNA search is faster and involves more sliding when the C-tails are charged rather than neutralized, suggesting that the deletion of C-tails or their posttranslational modification (e.g., phosphorylation and acetylation[3,67]) may reduce their ability to bind nonspecifically to DNA. Our results indicate that the disordered and charged C-termini accelerate search speed, in agreement with a recent single-molecule measurement of the diffusion of full-length p53 and of a truncated version without the C-tails.[68] Furthermore, the large capture radius of the C-tails allows interactions with distant DNA molecules, which may increase p53's localization to and affinity for DNA. The positively charged residues of the C-tails are therefore essential for modulating interactions with DNA. Our evidence supports previous studies on the importance of net charge to the biophysics and function of intrinsically disordered proteins, in general, and to the modulation of protein–DNA interactions, in particular.[69–74]

The C-tails linked to the Tet domain influence the orientation of the latter to the DNA molecule. The orientation changes from parallel in the presence of a single DNA molecule to perpendicular in the presence of two DNA molecules. Increasing the salt concentration results in a loss of the parallel orientation of the Tet domain to the DNA, as does increased crowding (i.e., a higher DNA concentration). Our simulations show that the flexible C-tails mediate the interactions between the core domain and the DNA, and dictate the orientation of the Tet domain relative to the DNA. We consider the perpendicular orientation of the Tet domain to the DNA in the case of two DNA molecules, as well as the parallel orientation of the Tet domain to a single DNA, at low salt concentrations as being linked to the dependency of the orientation of the Tet domain on the local cellular environment.

When p53 includes a tetrameric core domain, the C-tails speed the search. However, during the quicker search performed by two core domain dimers operating independently, the C-tails interfere with the subsequent interaction between the primary dimers of the core domains (via formation of the secondary interface) and impede binding to contiguous response elements. A higher salt concentration or a denser DNA environment alleviates this situation as it weakens the interaction between the C-tails and the DNA. Our results on the interactions of the charged tail with the DNA and its role in modulating the orientation of the Tet domain relative to the DNA support the hypothesis that the tails can further promote conformational change in the p53–DNA complex. Thus, the tails are important not only in order to search for the location

of cognate sites but also to support specific interactions with the DNA (e.g., DNA bending, as observed in the crystal structure of the specific complex of p53 with some of its response elements[75–77]).

Cross-talk between the Tet domain and the core domain was investigated by studying their dynamics with DNA both as part of the CTetCD model of p53 (i.e., tethered) and as isolated domains. The isolated core domains can engage in a combination of sliding, hopping, and 3D diffusion, resulting in a more efficient DNA search. By contrast, its linkage to the Tet domain means that the tethered tetrameric core domain searches DNA solely by sliding at most salt concentrations. With respect to the Tet domain, the isolated domain interacts more tightly with DNA and slides faster than its tethered counterpart.

Changes in the electrostatic potential of the core domain can also affect sliding on p53. The five cancer-causing mutants of p53 that involve a single charge neutralization at the interface between the core domain and DNA show a significantly faster linear diffusion because they search DNA by hopping rather than by sliding. The faster search may also result in a weaker nonspecific affinity for DNA. For example, Lys120, which directly interacts with DNA and whose neutralization significantly increases D_1, is only rarely mutated in human cancer but undergoes acetylation that leads to apoptosis.[78,79] One may suggest that the rapid translocation of the core domain mutants during nonspecific interactions with DNA prevents the mutant from locating the specific DNA binding site and results in a lower affinity for the DNA, which may affect the functioning of p53 as a transcription factor. Our overall observation that neutralizing a positively charged residue at the DNA interface causes faster linear diffusion is expected, given that interactions in our model are guided solely by electrostatics. However, our results indicate that the magnitude of the effects is not identical for all mutations, and we may propose that the magnitude of the effect of the mutation on the diffusion coefficient is related to the malfunction of the mutant and that the effects of mutation will depend on the extent to which these biophysical features are changed.

In the current study, the p53 model was based on the complex of the specific interactions of the core domain with DNA (in which the two half-sites are separated by 2 bp) and the structural insights obtained from cryo-EM and SAXS of full p53.[23] However, the richness of the p53 system may allow other modeling as well.[56,66] Indeed, we investigated a single wiring mode among six valid possibilities that are more evenly populated when the core domains are modeled without the secondary interface and that may show different cross-talks between the domains compared to those reported here for the selected variant. Our study included a

Table 2. The variants of p53 simulated in this study

	Full length (residues 1–393)	CTetCD (residues 94–393)	CTetCD neutral N-tail (residues 94–393)	CTetCD truncated N-tail (residues 94–362)	C only (residues 94–303)	TetCD only (residues 324–393)
Monomer	—	✓	✓	—	✓	—
Dimer	—	✓	✓	—	✓	—
Tetramer (0-bp spacer)	✓	✓, 2	—	—	—	—
Tetramer (2-bp spacer)	✓	✓, M, 2	✓, 2	✓	✓, M	✓
Tetramer (no second interface)	✓	✓, M, 2	✓, 2	—	—	—

The rows indicate the oligomeric state of p53, and the columns indicate the domains that were included. The symbol "✓" indicates that the interaction of these variants of p53 with a single 100-bp DNA was studied. "M" indicates that this construct was studied for both the wild-type version of the model and the five mutants of p53 generated by neutralizing the positive charge of the Lys120, Arg248, Arg273, Arg280, or Arg283 residue in the core domain (C) section of the model. "2" indicates that the wild-type construct was also studied with two parallel DNA molecules, each of 100 bp. Tetrameric p53 with the two dimeric core domains interacting with response elements spaced by 0 bp or 2 bp was modeled using PDB IDs 3KZ8 and 2AC0, respectively. Tetrameric p53 with no second interface was modeled by the structure given in PDB ID 2AC0, but eliminated the formation of any interactions at the second interface.

crude model of the flexible regions (i.e., N-termini, C-tails, and linkers) whose structural features may, in some instances, raise questions as to the validity of their representation as random coils. Furthermore, another cryo-EM study has suggested an alternative organization of p53 with DNA.[24] Future studies of p53 should therefore address the interplay between the structural heterogeneity and complexity of p53, the cellular environment, and the crosstalk between its various domains toward achieving molecular and biophysical quantification of the functioning of p53 as "the guardian of the genome."

Materials and Methods

The studied variants of p53

The p53 protein is a complex transcription factor that is composed of about four domains that form a tetramer (Fig. 1). To explore cross-talks between the domains and the effect of the oligomeric state of p53, we studied several variants of the protein involving different oligomeric states of p53 and including different domains. Simulations of tetrameric full-length p53 were complemented by simulations using the tetrameric CTetCD model (residues 94–393, which lack the N-termini). In its tetrameric form, the two primary dimers of the core domain (see Fig. 1) can be separated by 0–20 bp. We therefore simulated three variants of full-length p53 (as well as of the CTetCD model of p53). In two variants, there was a secondary interface between the two primary dimers (based on the crystal structures 3KZ8 and 2AC0 in which there is a spacer of 0 bp and 2 bp, respectively). The third was a tetrameric variant in which there was no secondary interface between the two primary dimers, so the distance between the two primary dimers is governed by the linker connecting the core domain to the Tet domain. The primary dimer of the core domain is defined based on the larger interface found in the crystal structure of the core

domains with DNA (e.g., the interface between A and B is larger than that between A and D).[12] To pinpoint the effect of the C-tails, we studied the same variants but with the tails uncharged or entirely removed. We also studied the monomeric and dimeric variants of p53, either as the CTetCD model or as the core domain only. An isolated Tet domain with C-tails was studied as well. In all these simulations, the protein was entirely flexible.

Cancer-causing mutants bearing mutations in the core domain at the interface formed with the DNA were simulated using the tetrameric variants [either with or without the secondary interface (found in p53 when the response elements are separated by 2 bp), as well as the isolated tetrameric core domain] to examine how the mutations affect nonspecific interactions with DNA. The mutations were introduced to residues Lys120, Arg248, Arg273, Arg280, and Arg283 by neutralizing the positive charge at their corresponding bead in the core domain. Table 2 summarizes all the variants of p53 studied in this work.

Construction of the p53 structure

While several experimental techniques have recently provided a structural understanding of the organization of the various domains of p53 in its tetrameric form, the fact that p53 includes a substantial fraction of disordered regions hinders efforts to obtain a description of the structure of full p53 at an atomic resolution. To study the nonspecific interaction of p53 with DNA, we need a model of p53 that reveals the connectivity of the core and Tet domains. We therefore constructed the full-length model of p53 involving the N-terminus, core domain, Tet domain, and C-tails. To construct this model, we used the coordinates of the high-resolution structures of the p53 core domain [Protein Data Bank (PDB) ID: 3KZ8 or 2AC0] and of the Tet domain (PDB ID: 1SAE). In these models, the core domain has a secondary interface between the primary dimers spaced 0 bp or 2 bp apart. The length of the linkers between the Tet domain and the core domain allows the two primary dimers to be spaced even farther apart during DNA search. We therefore constructed

another model of tetrameric p53 in which there is no secondary interface and the two primary dimers can more independently diffuse along the DNA. The N-termini (residues 1–93), the unstructured linkers (residues 303–324), and the C-terminal tails (residues 363–393) were incorporated by PyMOL.[80] The N-tails and C-tails were modeled simply as random coils. Although both termini are highly charged (17 negatively charged residues and 2 positively charged residues in each N-terminus, and 9 positively charge residues in each C-terminus) and hence are intrinsically unstructured, they may have some structural features.[69,70] A preference for a helical structure was observed for peptides from the C-tails in complex with a number of p53 binding proteins.[81] The N-terminus includes a proline-rich region, which may have some conformational preferences. A more refined modeling of unstructured regions in p53 may be included in future studies. The CTetCD model of p53 (core domain, Tet domain, and C-terminal domain) was constructed by truncating the N-terminus (1–93 residues) from the full-length model.

The first task in constructing the quaternary organization of p53 is to explore the wiring between the tetrameric core domain and the tetrameric Tet domain. As shown in Fig. 1, there are six possible wiring patterns. The thermodynamics and kinetics of these variants were probed when studying the folding characteristics of each of the six variants while keeping the core domain frozen on the DNA. We then designed a model in which all the six variants could be formed, and we looked for the most populated variant. The details of the prediction of the organization of p53 are described below.

Simulation model

To study long timescale processes such as sliding, hopping, and 3D diffusion, we simulated our system using a coarse-grained model.[46,59,74] In the protein, each amino acid was represented by a single bead for each residue located at the C^α of that residue. In the DNA, each nucleotide was represented by three beads: for the phosphate, sugar, and base groups. Beads representing charged amino acids (Arg, Lys, Asp, and Glu) were charged in the model, and a negative point charge was assigned to beads representing the DNA phosphate groups. Unlike the protein, the DNA remained frozen during the simulations and could not undergo folding and unfolding events. We assumed that the flexibility of DNA, which can have a major effect on the specific affinity of a protein for DNA, has a minor effect on nonspecific protein–DNA interactions.[82–85] The protein and the DNA were placed in a box with dimensions of 350 Å × 350 Å × 350 Å, with the DNA being placed at the center of the box along its z-axis.

The native-topology-based model corresponded to a perfectly funneled energy landscape, where native protein interactions were attractive and all other interactions were repulsive.[86,87] In addition to the native interactions, electrostatic interactions between all charged residues of the protein and the phosphate bead of the DNA were included and modeled by the Debye–Huckel potential, which accounts for the ionic strength of a solute immersed in aqueous solution.[88] Using this structure-based model, we and others have

studied the effect of charge–charge interactions on protein folding and protein stability.[89,90] We would like to emphasize that protein–DNA interactions were modeled solely by electrostatic forces and excluded-volume interactions, so there is no bias toward the specific interactions between the proteins and the DNA. The role of electrostatic interactions in nonspecific and specific protein–DNA complexes is widely acknowledged,[91,92] and we have shown for various protein–DNA systems that the sliding dynamics and DNA search mechanism can be captured by electrostatic and excluded-volume interactions.[46,48,59,93]

The sliding dynamics was explored at salt concentrations in the range of 0.01–0.07 M using a dielectric constant of 80. The simulations were performed at a constant temperature that was chosen to be low enough to preclude dissociation of p53, which could make its dynamics with DNA more complex, while allowing thermal fluctuations. Accordingly, while p53 could exhibit some intramolecular and intermolecular dissociations, in this study, we focused on its association and dissociation dynamics with DNA and its internal protein variability. This approach has the advantage of enabling us to dissect the molecular complexity of p53 and to pinpoint the components that govern its nonspecific interactions with DNA. A 100-bp B-DNA was used to study nonspecific interactions of the protein with a single double-stranded DNA; two 100-bp B-DNA molecules, separated by a distance of 100 Å and oriented parallel with each other along their z-axes, were used to study the nonspecific interactions of p53 with two DNAs. The Debye–Huckel model is valid for low salt concentrations and has been used previously to study the energetics and dynamics of various biomolecular systems such as RNA folding,[94] chromatin assembly,[95] and protein–DNA binding.[93] More details on the simulation can be found in Givaty and Levy.[46] The search mechanism of p53 was studied for wild-type p53 and for cancer-causing mutants bearing mutations in the core domain at the interface formed with the DNA (Lys120, Arg248, Arg273, Arg280, and Arg283).

To compare the folding thermodynamics of the six variants of p53, we studied their folding using a native-topology-based model constructed separately for each variant. Numerous trajectories simulated using Langevin dynamics[96] ($\gamma = 0.01$) were collected, so that multiple unfolding/folding transitions would be available at temperatures around the folding temperature. The folding thermodynamics was inferred using the weighted histogram analysis method.[97] During these simulations, the core domain of p53 and the DNA molecule were kept static. The folding properties of p53 were introduced by five terms: bonds, angles, torsion angles, the Lennard–Jones term for native interactions, and an excluded-volume term for nonnative interactions. All the energy terms were identical for different variants of the Tet domain of p53, and the variants differed only in terms of the interconnectivity of the core and Tet domains.

In addition to the folding simulations of the six designed variants of p53 shown in Fig. 1, we ran multiple simulated-annealing trajectories in which all the six variants could be formed. The purpose of the annealing simulations was to examine the population of each variant in light of its thermodynamics. During the simulated-annealing trajectories, we started with a dissociated Tet domain and gradually reduced the temperature of each

simulation step (from $T_{start} = 0.92T_f$ to $T_{end} = 0.6T_f$). In the annealing simulations, the core domain was frozen and directly interacted with the DNA. To examine the effect of the spacer between the primary dimers of the core domain on the wiring of the Tet domain, we performed the annealing simulation with a spacer of 2 bp or 12 bp.

Structural classification of protein sliding, hopping, and 3D diffusion

DNA search by p53 was simulated at various salt concentrations, which were introduced by the Debye–Huckel model. Trajectories were analyzed to measure the mode of DNA search adopted (sliding, hopping, or 3D diffusion), the structural features during sliding, and the linear diffusion coefficient (D_1), using the procedures described by Givaty and Levy.[46] The p53 protein can slide along the DNA using two different domains: the recognition loop of the core domain and the positively charged tails of the Tet domain. Since these two recognition regions have different flexibilities, different definitions for sliding are needed for each domain. For the core domain, a snapshot was defined as taking part in a sliding search if at least 65% of its recognition loop (residues 116–121, in which Lys120 contributes to a tight interaction with the major groove) was in contact with the major groove and the distance of the center of mass of the recognition loop from the DNA was up to 14 Å longer than that in the crystal structure. For the C-tail, the only difference was that the distance of the center of mass of the recognition region of the tail from the DNA was required to be up to 10 Å longer than that in the crystal structure. Since the tails are very flexible and disordered, the positions probed using sliding, D_1, and the z-displacement of the C-tails were analyzed according to the center of mass of the Tet domain to which the tails were attached. The structural requirements for sliding had to be fulfilled for at least a single subunit of these C-tails. The simulation frame was considered as performing 3D diffusion if the center of mass of the recognition loop/region of the tail of each of the four monomers was farther than 30 Å from the main DNA axis, since, at this distance, the average electrostatic energy drops to about 2% of the energy in the sliding conformations at low salt concentration.

Supplementary materials related to this article can be found online at doi:10.1016/j.jmb.2011.01.059

References

1. Levine, A. J. & Oren, M. (2009). The first 30 years of p53: growing ever more complex. *Nat. Rev. Cancer*, **9**, 749–758.
2. Vousden, K. H. & Prives, C. (2009). Blinded by the light: the growing complexity of p53. *Cell*, **137**, 413–431.
3. Bode, A. M. & Dong, Z. G. (2004). Post-translational modification of p53 in tumorigenesis. *Nat. Rev. Cancer*, **4**, 793–805.
4. Appella, E. & Anderson, C. W. (2001). Post-translational modifications and activation of p53 by genotoxic stresses. *Eur. J. Biochem.* **268**, 2764–2772.
5. Brooks, C. L. & Gu, W. (2003). Ubiquitination, phosphorylation and acetylation: the molecular basis for p53 regulation. *Curr. Opin. Cell Biol.* **15**, 164–171.
6. Joerger, A. C. & Fersht, A. R. (2007). Structure–function-rescue: the diverse nature of common p53 cancer mutants. *Oncogene*, **26**, 2226–2242.
7. Cho, Y. J., Gorina, S., Jeffrey, P. D. & Pavletich, N. P. (1994). Crystal-structure of a p53 tumor-suppressor DNA complex—understanding tumorigenic mutations. *Science*, **265**, 346–355.
8. Jeffrey, P. D., Gorina, S. & Pavletich, N. P. (1995). Crystal-structure of the tetramerization domain of the p53 tumor-suppressor at 1.7 angstroms. *Science*, **267**, 1498–1502.
9. Clore, G. M., Omichinski, J. G., Sakaguchi, K., Zambrano, N., Sakamoto, H., Appella, E. & Gronenborn, A. M. (1995). Interhelical angles in the solution structure of the oligomerization domain of p53: correction.; (vol. 265, p. 386, 1994). *Science*, **267**, 1515–1516.
10. Canadillas, J. M. P., Tidow, H., Freund, S. M. V., Rutherford, T. J., Ang, H. C. & Fersht, A. R. (2006). Solution structure of p53 core domain: structural basis for its instability. *Proc. Natl Acad. Sci. USA*, **103**, 2109–2114.
11. Lee, W. T., Harvey, T. S., Yin, Y., Yau, P., Litchfield, D. & Arrowsmith, C. H. (1994). Solution structure of the tetrameric minimum transforming domain of p53. *Nat. Struct. Biol.* **1**, 877–890.
12. Kitayner, M., Rozenberg, H., Kessler, N., Rabinovich, D., Shaulov, L., Haran, T. E. & Shakked, Z. (2006). Structural basis of DNA recognition by p53 tetramers. *Mol. Cell*, **22**, 741–753.
13. Ho, W. C., Fitzgerald, M. X. & Marmorstein, R. (2006). Structure of the p53 core domain dimer bound to DNA. *J. Biol. Chem.* **281**, 20494–20502.
14. Joerger, A. C. & Fersht, A. R. (2008). Structural biology of the tumor suppressor p53. *Annu. Rev. Biochem.* **77**, 557–582.
15. el-Deiry, W., Kern, S., Pietenpol, J., Kinzler, K. & Vogelstein, B. (1992). Definition of a consensus binding site for p53. *Nat. Genet.* **1**, 45–49.
16. Wei, C., Wu, Q., Vega, V., Chiu, K., Ng, P., Zhang, T. *et al.* (2005). A global map of p53 transcription-factor binding sites in the human genome. *Cell*, **124**, 207–219.
17. Kitayner, M., Rozenberg, H., Rohs, R., Suad, O., Rabinovich, D., Honig, B. & Shakked, Z. (2010). Diversity in DNA recognition by p53 revealed by crystal structures with Hoogsteen base pairs. *Nat. Struct. Mol. Biol.* **17**, 423–429.

18. Chen, Y. H., Dey, R. & Chen, L. (2010). Crystal structure of the p53 core domain bound to a full consensus site as a self-assembled tetramer. *Structure*, **18**, 246–256.

19. Malecka, K. A., Ho, W. C. & Marmorstein, R. (2009). Crystal structure of a p53 core tetramer bound to DNA. *Oncogene*, **28**, 325–333.

20. Melero, R., Rajagopalan, S., Lazaro, M., Joerger, A. C., Brandt, T., Veprintsev, D. B. *et al.* (2011). Electron microscopy studies on the quaternary structure of p53 reveal different binding modes for p53 tetramers in complex with DNA. *Proc. Natl Acad. Sci. USA*, **108**, 557–562.

21. Okorokov, A. L. & Orlova, E. V. (2009). Structural biology of the p53 tumour suppressor. *Curr. Opin. Struct. Biol.* **19**, 197–202.

22. Shakked, Z. (2007). Quaternary structure of p53: the light at the end of the tunnel. *Proc. Natl Acad. Sci. USA*, **104**, 12231–12232.

23. Tidow, H., Melero, R., Mylonas, E., Freund, S. M. V., Grossmann, J. G., Carazo, J. M. *et al.* (2007). Quaternary structures of tumor suppressor p53 and a specific p53–DNA complex. *Proc. Natl Acad. Sci. USA*, **104**, 12324–12329.

24. Okorokov, A. L., Sherman, M. B., Plisson, C., Grinkevich, V., Sigmundsson, K., Selivanova, G. *et al.* (2006). The structure of p53 tumour suppressor protein reveals the basis for its functional plasticity. *EMBO J.* **25**, 5191–5200.

25. Hollstein, M., Rice, K., Greenblatt, M. S., Soussi, T., Fuchs, R., Sorlie, T. *et al.* (1994). Database of p53 gene somatic mutations in human tumors and cell-lines. *Nucleic Acids Res.* **22**, 3551–3555.

26. Hainaut, P. & Hollstein, M. (2000). p53 and human cancer: the first ten thousand mutations. *Adv. Cancer Res.* **77**, 81–137.

27. Wolcke, J., Reimann, M., Klumpp, M., Gohler, T., Kim, E. & Deppert, W. (2003). Analysis of p53 "latency" and "activation" by fluorescence correlation spectroscopy—evidence for different modes of high affinity DNA binding. *J. Biol. Chem.* **278**, 32587–32595.

28. Kaeser, M. D. & Iggo, R. D. (2002). Chromatin immunoprecipitation analysis fails to support the latency model for regulation of p53 DNA binding activity *in vivo*. *Proc. Natl Acad. Sci. USA*, **99**, 95–100.

29. Espinosa, J. M. & Emerson, B. M. (2001). Transcriptional regulation by p53 through intrinsic DNA/chromatin binding and site-directed cofactor recruitment. *Mol. Cell*, **8**, 57–69.

30. Ahn, J. & Prives, C. (2001). The C-terminus of p53: the more you learn the less you know. *Nat. Struct. Biol.* **8**, 730–732.

31. Weinberg, R. L., Freund, S. M., Veprintsev, D. B., Bycroft, M. & Fersht, A. R. (2004). Regulation of DNA binding of p53 by its C-terminal domain. *J. Mol. Biol.* **342**, 801–811.

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

33. Liu, Y., Lagowski, J. P., Vanderbeek, G. E. & Kulesz-Martin, M. F. (2004). Facilitated search for specific genomic targets by p53 C-terminal basic DNA binding domain. *Cancer Biol. Ther.* **3**, 1102–1108.

34. Berg, O. G., Winter, R. B. & Vonhippel, P. H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic-acids: 1. Models and theory. *Biochemistry*, **20**, 6929–6948.

35. Vonhippel, P. H. & Berg, O. G. (1989). Facilitated target location in biological-systems. *J. Biol. Chem.* **264**, 675–678.

36. Wunderlich, Z. & Mirny, L. A. (2008). Spatial effects on the speed and reliability of protein–DNA search. *Nucleic Acids Res.* **36**, 3570–3578.

37. Slutsky, M. & Mirny, L. A. (2004). Kinetics of protein–DNA interaction: facilitated target location in sequence-dependent potential. *Biophys. J.* **87**, 4021–4035.

38. Shimamoto, N. (1999). One-dimensional diffusion of proteins along DNA—its biological and chemical significance revealed by single-molecule measurements. *J. Biol. Chem.* **274**, 15293–15296.

39. Klenin, K. V., Merlitz, H., Langowski, J. & Wu, C. X. (2006). Facilitated diffusion of DNA-binding proteins. *Phys. Rev. Lett.* **96**, 018104–1–018104-4.

40. Halford, S. E. & Marko, J. F. (2004). How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res.* **32**, 3040–3052.

41. Halford, S. E. (2009). An end to 40 years of mistakes in DNA–protein association kinetics? *Biochem. Soc. Trans.* **37**, 343–348.

42. Schurr, J. M. (1979). One-dimensional diffusion-coefficient of proteins absorbed on DNA—hydrodynamic considerations. *Biophys. Chem.* **9**, 413–414.

43. Iwahara, J., Zweckstetter, M. & Clore, G. M. (2006). NMR structural and kinetic characterization of a homeodomain diffusing and hopping on nonspecific DNA. *Proc. Natl Acad. Sci. USA*, **103**, 15062–15067.

44. Iwahara, J. & Clore, G. M. (2006). Direct observation of enhanced translocation of a homeodomain between DNA cognate sites by NMR exchange spectroscopy. *J. Am. Chem. Soc.* **128**, 404–405.

45. Iwahara, J. & Clore, G. M. (2006). Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature*, **440**, 1227–1230.

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

47. Blainey, P., Luo, G., Kou, S., Mangel, W., Verdine, G., Bagchi, B. & Xie, X. S. (2009). Nonspecifically bound proteins spin while diffusing along DNA. *Nat. Struct. Mol. Biol.* **16**.

48. Vuzman, D., Azia, A. & Levy, Y. (2010). Searching DNA via a "Monkey Bar" mechanism: the significance of disordered tails. *J. Mol. Biol.* **396**, 674–684.

49. Gowers, D. M., Wilson, G. G. & Halford, S. E. (2005). Measurement of the contributions of 1D and 3D pathways to the translocation of a protein along DNA. *Proc. Natl Acad. Sci. USA*, **102**, 15883–15888.

50. Gorman, J. & Greene, E. C. (2008). Visualizing one-dimensional diffusion of proteins along DNA. *Nat. Struct. Mol. Biol.* **15**, 768–774.

51. Bustamante, C., Guthold, M., Zhu, X. S. & Yang, G. L. (1999). Facilitated target location on DNA by individual *Escherichia coli* RNA polymerase molecules observed with the scanning force microscope operating in liquid. *J. Biol. Chem.* **274**, 16665–16668.

52. Wang, J., Lu, Q. & Lu, H. P. (2006). Single-molecule dynamics reveals cooperative binding–folding in protein recognition. *PLoS Comput. Biol.* **2**, 842–852.

53. Elf, J., Li, G. W. & Xie, X. S. (2007). Probing transcription factor dynamics at the single-molecule level in a living cell. *Science*, **316**, 1191–1194.

54. Graneli, A., Yeykal, C. C., Robertson, R. B. & Greene, E. C. (2006). Long-distance lateral diffusion of human Rad51 on double-stranded DNA. *Proc. Natl Acad. Sci. USA*, **103**, 1221–1226.

55. Tafvizi, A., Huang, F., Leith, J. S., Fersht, A. R., Mirny, L. A. & van Oijen, A. M. (2008). Tumor suppressor p53 slides on DNA with low friction and high stability. *Biophys. J.* **95**, L1–L3.

56. Joerger, A. C., Rajagopalan, S., Natan, E., Veprintsev, D. B., Robinson, C. V. & Fersht, A. R. (2009). Structural evolution of p53, p63, and p73: implication for heterotetramer formation. *Proc. Natl Acad. Sci. USA*, **106**, 17705–17710.

57. Natan, E., Hirschberg, D., Morgner, N., Robinson, C. V. & Fersht, A. R. (2009). Ultraslow oligomerization equilibria of p53 and its implications. *Proc. Natl Acad. Sci. USA*, **106**, 14327–14332.

58. Pan, Y. P. & Nussinov, R. (2009). Cooperativity dominates the genomic organization of p53-response elements: a mechanistic view. *Plos Comput. Biol.* **5**, e1000448.

59. Vuzman, D., Polonsky, M. & Levy, Y. (2010). Facilitated DNA search by multidomain transcription factors: cross-talk via a flexible linker. *Biophys. J.* **99**, 1202–1211.

60. Iwahara, J., Schwieters, C. D. & Clore, G. M. (2004). Characterization of nonspecific protein–DNA interactions by ^{1}H paramagnetic relaxation enhancement. *J. Am. Chem. Soc.* **126**, 12800–12808.

61. Viadiu, H. & Aggarwal, A. K. (2000). Structure of BamHI bound to nonspecific DNA: a model for DNA sliding. *Mol. Cell*, **5**, 889–895.

62. Kalodimos, C. G., Biris, N., Bonvin, A. M. J. J., Levandoski, M. M., Guennuegues, M., Boelens, R. & Kaptein, R. (2004). Structure and flexibility adaptation in nonspecific and specific protein–DNA complexes. *Science*, **305**, 386–389.

63. Pan, Y. P. & Nussinov, R. (2010). Preferred drifting along the DNA major groove and cooperative anchoring of the p53 core domain: mechanisms and scenarios. *J. Mol. Recognit.* **23**, 232–240.

64. Sauer, M., Bretz, A. C., Beinoraviciute-Kellner, R., Beitzinger, M., Burek, C., Rosenwald, A. *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.

65. Chitayat, S. & Arrowsmith, C. (2010). Four p(53)s in a pod. *Nat. Struct. Mol. Biol.* **17**, 390–391.

66. Huang, F., Rajagopalan, S., Settanni, G., Marsh, R. J., Armoogum, D. A., Nicolaou, N. *et al.* (2009). Multiple conformations of full-length p53 detected with single-molecule fluorescence resonance energy transfer. *Proc. Natl Acad. Sci. USA*, **106**, 20758–20763.

67. Jayaraman, L. & Prives, C. (1999). Covalent and noncovalent modifiers of the p53 protein. *Cell Mol. Life Sci.* **55**, 76–87.

68. Tafvizi, A., Huang, F., Fersht, A. R., Mirny, L. A. & van Oijen, A. M. (2011). A single-molecule character-ization of p53 search on DNA. *Proc. Natl Acad. Sci. USA*, **108**, 563–568.

69. Mao, A. H., Crick, S. L., Vitalis, A., Chicoine, C. L. & Pappu, R. V. (2010). Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc. Natl Acad. Sci. USA*, **107**, 8183–8188.

70. Marsh, J. A. & Forman-Kay, J. D. (2010). Sequence determinants of compaction in intrinsically disor-dered proteins. *Biophys. J.* **98**, 2383–2390.

71. Borg, M., Mittag, T., Pawson, T., Tyers, M., Forman-Kay, J. D. & Chan, H. S. (2007). Polyelectrostatic interactions of disordered ligands suggest a physical basis for ultrasensitivity. *Proc. Natl Acad. Sci. USA*, **104**, 9650–9655.

72. Muller-Spath, S., Soranno, A., Hirschfeld, V., Hofmann, H., Ruegger, S., Reymond, L. *et al.* (2010). From the cover: charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc. Natl Acad. Sci. USA*, **107**, 14609–14614.

73. England, J. L. & Haran, G. (2010). To fold or expand—a charged question. *Proc. Natl Acad. Sci. USA*, **107**, 14519–14520.

74. Vuzman, D. & Levy, Y. (2010). DNA search efficiency is modulated by charge composition and distribution in the intrinsically disordered tail. *Proc. Natl Acad. Sci. USA*, **107**, 21004–21009.

75. Pan, Y. P. & Nussinov, R. (2008). p53-induced DNA bending: the interplay between p53–ONA and p53–p53 interactions. *J. Phys. Chem. B*, **112**, 6716–6724.

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

77. Nagaich, A. K., Zhurkin, V. B., Durell, S. R., Jernigan, R. L., Appella, E. & 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. USA*, **96**, 1875–1880.

78. Tang, Y., Luo, J. Y., Zhang, W. Z. & Gu, W. (2006). Tip60-dependent acetylation of p53 modulates the decision between cell-cycle arrest and apoptosis. *Mol. Cell*, **24**, 827–839.

79. Sykes, S. M., Mellert, H. S., Holbert, M. A., Li, K., Marmorstein, R., Lane, W. S. & McMahon, S. B. (2006). Acetylation of the p53 DNA-binding domain regu-lates apoptosis induction. *Mol. Cell*, **24**, 841–851.

80. DeLano, W. L. (2006). *The PyMOL Molecular Graphics System.* DeLano Scientific, San Carlos, CA.

81. Rustandi, R. R., Baldisseri, D. M. & Weber, D. J. (2000). Structure of the negative regulatory domain of p53 bound to S100B(beta beta). *Nat. Struct. Biol.* **7**, 570–574.

82. Rohs, R., West, S., Liu, P. & Honig, B. (2009). Nuance in the double-helix and its role in protein–DNA recognition. *Curr. Opin. Struct. Biol.* **19**, 171–177.

83. Rohs, R., West, S., Sosinnsky, A., Liu, P., Mann, R. & Honig, B. (2009). The role of DNA shape in protein–DNA recognition. *Nature*, **461**, 1248–1253.

84. Bouvier, B. & Lavery, R. (2009). A free energy pathway for the interaction of the SRY protein with its binding site on DNA from atomistic simulations. *J. Am. Chem. Soc.* **131**, 9864–9865.

85. Prevost, C., Takahashi, M. & Lavery, R. (2009). Deforming DNA: from physics to biology. *Chem-PhysChem*, **10**, 1399–1404.

86. Levy, Y., Onuchic, J. N. & Wolynes, P. G. (2007). Fly-casting in protein–DNA binding: frustration between protein folding and electrostatics facilitates target recognition. *J. Am. Chem. Soc.* **129**, 738–739.

87. Onuchic, J. N. & Wolynes, P. G. (2004). Theory of protein folding. *Curr. Opin. Struct. Biol.* **14**, 70–75.

88. Schlick, T. (2000). *Molecular Modeling and Simulation: An Interdisciplinary Guide.* Springer, New York, NY.

89. Azia, A. & Levy, Y. (2009). Nonnative electrostatic interactions can modulate protein folding: molecular dynamics with a grain of salt. *J. Mol. Biol.* **393**, 527–542.

90. Su, J., Chen, W. & Wang, C. (2010). Role of electrostatic interactions for the stability and folding behavior of cold shock protein. *Proteins*, **78**, 2157–2169.

91. Misra, V. K., Hecht, J. L., Yang, A. S. & Honig, B. (1998). Electrostatic contributions to the binding free energy of the lambda cl repressor to DNA. *Biophys. J.* **75**, 2262–2273.

92. Temiz, A., Benos, P. & Camacho, C. J. (2010). Electrostatic hot spot on DNA-binding domains mediates phosphate desolvation and the pre-organization of specificity determinant side chains. *Nucleic Acids Res.* **38**, 2134–2144.

93. Marcovitz, A. & Levy, Y. (2009). Arc-repressor dimerization on DNA: folding rate enhancement by colocalization. *Biophys. J.* **96**, 4212–4220.

94. Hyeon, C. & Thirumalai, D. (2005). Mechanical unfolding of RNA hairpins. *Proc. Natl Acad. Sci. USA*, **102**, 6789–6794.

95. Beard, D. A. & Schlick, T. (2001). Computational modeling predicts the structure and dynamics of chromatin fiber. *Structure*, **9**, 105–114.

96. Mor, A., Ziv, G. & Levy, Y. (2008). Simulations of proteins with inhomogeneous degrees of freedom: the effect of thermostats. *J. Comput. Chem.* **29**, 1992–1998.

97. Kumar, S., Bouzida, D., Swendsen, R. H., Kollman, P. A. & Rosenberg, J. M. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules: 1. The method. *J. Comput. Chem.* **13**, 1011–1021.