A single-molecule characterization of p53 search on DNA

Anahita Tafvizi[a,b], Fang Huang[c,3], Alan R. Fersht[c,2], Leonid A. Mirny[b,2,1], and Antoine M. van Oijen[a,2,1,4]

[a]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 250 Longwood Avenue, Seeley G. Mudd 204A, Boston, MA 02115; [b]Harvard-Massachusetts Institute of Technology Division of Health Sciences and Technology, 77 Massachusetts Avenue, E25-526C Cambridge, MA 02139; and [c]Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 0QH, United Kingdom

The tumor suppressor p53 slides along DNA while searching for its cognate site. Central to this process is the basic C-terminal domain, whose regulatory role and its coordination with the core DNA-binding domain is highly debated. Here we use single-molecule techniques to characterize the search process and disentangle the roles played by these two DNA-binding domains in the search process. We demonstrate that the C-terminal domain is capable of rapid translocation, while the core domain is unable to slide and instead hops along DNA. These findings are integrated into a model, in which the C-terminal domain mediates fast sliding of p53, while the core domain samples DNA by frequent dissociation and reassociation, allowing for rapid scanning of long DNA regions. The model further proposes how modifications of the C-terminal domain can activate "latent" p53 and reconciles seemingly contradictory data on the action of different domains and their coordination.

recognition | response element | transcription factor

An essential transcription factor in multicellular organisms, the tumor suppressor p53 regulates cell-cycle arrest and apoptosis. Genetic alteration and mutations of p53 have been found in more than 50% of all human cancers (1). Most of these mutations affect the core domain responsible for recognition and binding to cognate sites on DNA. In contrast to many other well characterized transcription factors, p53 contains two distinct DNA-binding domains (Fig. 1). Whereas the core domain binds DNA in a sequence-specific fashion, the C-terminal domain of p53 interacts with DNA in a manner independent of the sequence (2–4) and is subject to multiple acetylations and phosphorylations that activate p53. How the two DNA-binding domains coordinate their actions and influence dynamics of p53 has been a subject of great interest and controversy.

Initial studies of the interactions between p53 and DNA suggested that the C-terminal domain of p53 negatively regulates the binding of core domain to its specific site on DNA. Hupp et al. reported that the C-terminal domain of p53 has a remarkable effect on the ability of the core domain to bind to its target site on DNA (5). The authors reported that the deletion of the C-terminal domain, phosphorylation of a serine residue in the C-terminal domain, or the use of an antibody (PAb 421) directed to the C-terminal domain increased the binding of the core domain to short DNA oligonucleotides. Further, phosphorylation and acetylation of various residues in the C-terminal domain leads to an increase in the binding of core domain in vitro (5). In addition, modifications of the C-terminal domain are widely observed in cells in which DNA damage has activated a p53 response (6). These observations led to the hypothesis that the nonspecific interaction of the C-terminal domain with the DNA interferes with the ability of core domain to bind to the cognate site until relevant signals cause modifications in the C-terminal domain, and alleviate its negative regulation of core DNA binding (4, 5, 7).

More recent studies showed that p53 requires its C-terminal domain for efficient recognition of the target site in long or circular DNA (8). Further, the C-terminal domain is important
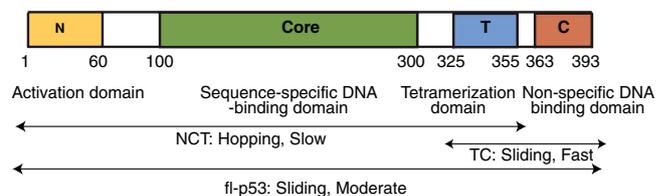

Fig. 1. Multidomain structure of p53 and constructs used in the study. Tumor suppressor p53, a 393-residue long protein, is composed of four different domains: The activation domain at the N-terminal end of the protein, the sequence-specific DNA-binding domain at the core of the protein, the tetramerization domain, and the nonspecific DNA-binding domain at the C-terminal end of the protein. p53 can bind DNA through both sequence-specific (core domain) and nonspecific (C-terminal domain) protein-DNA interactions.

in one-dimensional translocation of the protein along DNA (8). McKinney et al. also demonstrated that the C-terminally truncated protein is considerably less efficient at binding and trans-activating targets in vivo. Taken together, these results suggest a positive regulatory role of p53′s C-terminal domain for DNA binding (8).

Here we use single-molecule imaging tools to examine the complex role of the C-terminal domain in p53 recognition. We demonstrate that seemingly contradictory earlier studies can be reconciled into a comprehensive model of p53-DNA recognition if both kinetics of the search process and competition between specific and nonspecific binding are considered at high DNA concentrations present in the cell nucleus.

Central to this mechanism is the process by which p53 searches for its sites on long genomic DNA. Such a search process was suggested to constitute multiple rounds of three-dimensional diffusion and effectively one-dimensional sliding along DNA (9). Experimental studies of other, mostly bacterial, transcription factors have confirmed this mechanism in vitro and in vivo (10, 11). Key to this process is sliding along the duplex while bound to DNA nonspecifically. Because the C terminus binds DNA non-specifically, it is implied to be responsible for mediating sliding of p53 along DNA. Importantly, an efficient search process requires fast translocation along DNA and optimal affinity for nonspecific DNA. Excessive affinity leads to sequestration of the protein by nonspecific DNA, while insufficient affinity makes sliding rounds

short and search slow. Intriguingly, while the C terminus can provide sliding, it is the core domain that recognizes the cognate sequence (12–14). While structural studies have ruled out allosteric models of direct interactions between C terminus and core domains (15, 16), interplay between the two domains remains a subject of great interest.

Results

C-Terminal Domain of p53 Translocates on DNA Much Faster than the Full-Length p53, While the Core Domain Is Unable to Slide on DNA.

Aiming to understand the role of individual domains and to investigate the molecular mechanism underlying one-dimensional diffusion of p53 protein on DNA, we visualized and quantitatively characterized the motion of individual p53 proteins in vitro along flow-stretched DNA. In previously reported single-molecule studies, we visualized the interaction between fluorescently labeled p53 and DNA, and showed that the full-length p53 is capable of a diffusive translocation along DNA (17). In order to determine the role of the individual DNA-binding domains of p53 we perform single-molecule experiments on the following three constructs: the TC domain (tetramerization domain + C-terminal domain), the NCT domain (N-terminal domain + core domain + tetramerization domain) and the full-length p53 molecule. We fluorescently labeled NCT, TC, and full-length p53 constructs and used total internal reflection fluorescence (TIRF) microscopy to visualize their movement on flow-stretched lambda phage DNA molecules.

The 48.5-kb long double-stranded DNA was coupled at one end to the top surface of a microscope cover slip and hydrodynamically stretched by applying a laminar flow of aqueous buffer (Fig. 2A) (11, 17). The fluorescence emitted by the labeled protein was collected by a CCD camera and its position determined by fitting the fluorescence intensity profile to a two-dimensional Gaussian distribution (17). The protein positions for each captured frame were then tracked and linked to determine the trajectory of each protein.

Fig. 2C shows a time series of fluorescence images of representative full-length p53, the C-terminal, and the core domain of p53

moving along stretched DNA. Fig. 2D shows the trajectory of the same three protein constructs as determined from the images in Fig. 2C. Fig. 2E shows the Mean Square Displacement (MSD) vs. time for the three trajectories in two-dimensions. As can be seen from the trajectories and the MSD plots, the C-terminal domain is capable of translocating much faster on DNA than the full-length protein, while the core domain does not show a significant translocation on the same time scale.
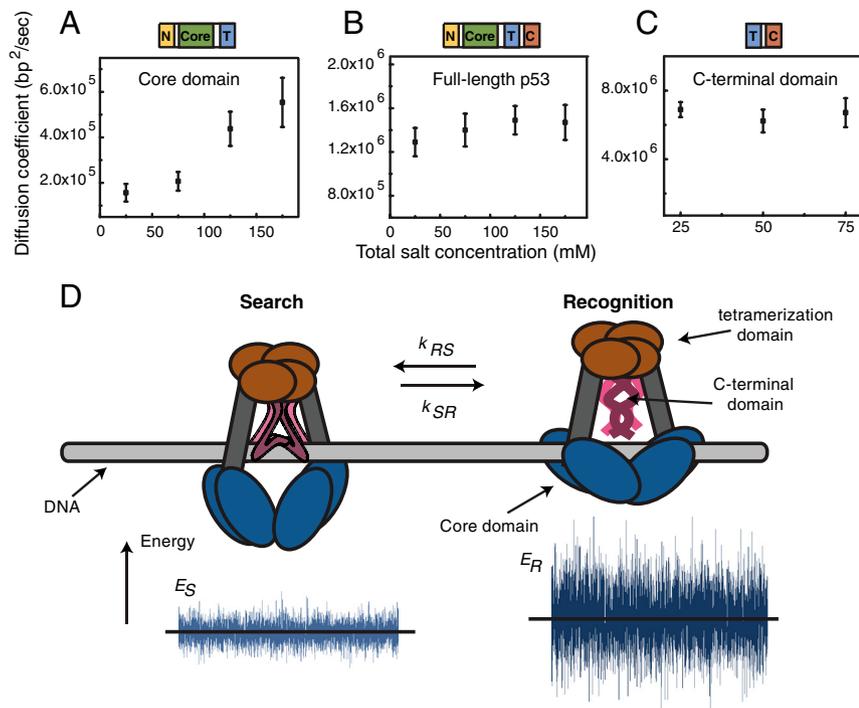
To characterize dynamics of individual molecules, we measure diffusion coefficients of their one-dimensional sliding. A key challenge is to factor out drift due to the flow and fluctuations of the DNA itself. The drift was determined from individual trajectories (Fig. S1) and subtracted (see *SI Text*). To take into account DNA fluctuations, we developed a method that uses DNA-attached quantum dots as reference points. We attached quantum dots to three different locations on DNA (Fig. 2B; positions corresponding to one-third, two-thirds, and full length of the lambda phage DNA) and measured the trajectories of the quantum dots in the same flow condition as used in our sliding experiments. The corresponding MSD vs. t plots, both for the movement along (longitudinal) and perpendicular (transversal) to the direction of the flow, are shown in Fig. S2. The MSD of the DNA-bound quantum dots increases at short time scales, but remains constant in longer time scales as is expected for bounded diffusion.

Fig. 3 compares the MSD at time $t = 0.5$ s of individual proteins (core domain, full-length protein, and C-terminal domain) in the longitudinal direction to its MSD in the transverse direction and in different salt concentrations. The black dashed line indicates the MSD of bound quantum dots on different locations on DNA and the gray area under the dashed line represents the intrinsic DNA fluctuations (Fig. S2). The MSDs obtained for the single-molecule trajectories using the core domain lie close to the MSD of the DNA fluctuations, suggesting that core domain is incapable of moving along DNA. But, for full-length p53 as well as the C-terminal domain, the MSD in the longitudinal direction is distinctly larger than that of the DNA fluctuations, suggesting diffusive movement of those constructs on DNA.
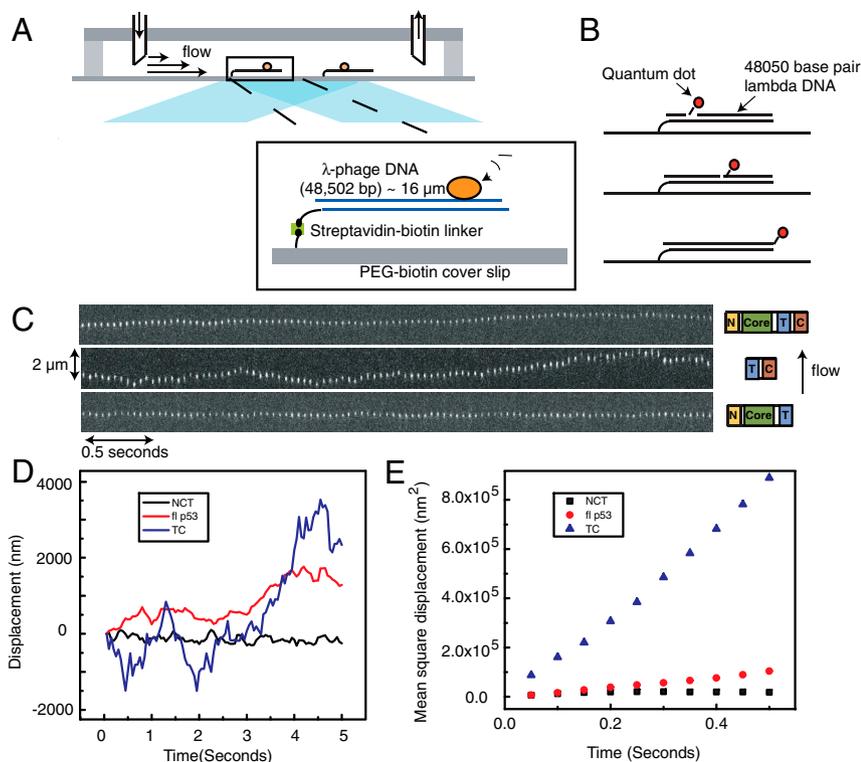
Fig. 2. Experimental set up and sliding trajectories of different p53 constructs. (*A*) Set up of the experiment. Several lambda-phage DNA molecules are attached to the glass bottom surface of a microfluidic flow cell through a streptavidin–biotin linker. The molecules are stretched by the drag force of a laminar flow of aqueous buffer through the flow cell. A laser is reflected off of the interface between the aqueous buffer and the glass surface generating an evanescent field decaying exponentially into the buffer. By mechanically stretching the DNA molecules within the evanescent field, the fluorophore labels of the p53 proteins on DNA can be selectively illuminated and excited while keeping the illumination from the background to a minimum. (*B*) To visualize the fluctuations due to DNA Brownian motion, quantum dots are attached to three different locations on lambda DNA, 14.7 kbp from the tethered point, 33,8 kbp from the tethered point, and at the end of the 48,5 kbp-long lambda DNA. (*C*) Kymograph of individual full-length p53, C-terminal domain (Tetramerization + C-terminal domain) and core domain (N-terminal + Core + Tetramerization domain) moving along flow-stretched DNA. All kymographs are superimposed on the same time and length axis. (*D*) Diffusion trajectory for the same three constructs. (*E*) MSD vs. time of the same trajectories.

Fig. 3. Mobility of different p53 constructs. MSD for drift-corrected trajectories of different constructs compared to MSD of DNA fluctuations in time 0.5 s and different salt concentration for core domain (NCT) (Blue), full-length p53 (Red), and C-terminal (TC) domain (Green) of p53. The gray area represents the MSD as a result of the DNA fluctuations measured by the quantum dot attached to DNA in both longitudinal and transverse direction. Full-length protein and C-terminal domain have higher MSD in the longitudinal direction, suggesting a one-dimensional diffusional translocation along DNA. For the core domain this is only the case at high salt concentration.

The diffusion coefficients for the three different protein constructs are determined at 75 mM total salt concentration after subtracting the effect of DNA fluctuation (see *SI Text*) and shown in Table 1. The C-terminal domain is capable of translocating much faster on DNA than the full-length protein, while the core domain displays a diffusion coefficient that is an order of magnitude smaller than either the full-length protein or the C-terminal domain.

The Core Domain of p53 Translocates on DNA via a *Hopping* Mechanism. Two distinct mechanisms are suggested for a protein that diffuses along DNA—sliding and hopping. A sliding protein remains in contact with DNA while translocating along it. On the other hand, a hopping protein is suggested to make microscopic associations and dissociations on and off the DNA. To differentiate between these mechanisms, we use the dependence of the diffusion coefficient on the salt concentration in the solvent. It is expected that the diffusion coefficient of sliding is independent of the salt concentration. However, in the case of hopping, the fraction of time spent off DNA depends on protein affinity for nonspecific DNA. A higher salt concentration reduces proteins affinity for nonspecific DNA making it dissociate more frequently, thus spending more time in solution subject to three-dimensional diffusion and thus yielding a higher diffusion coefficient (18).

In earlier work, we observed that the diffusion coefficient of full-length p53 protein is independent of salt concentration. Those data suggested a sliding mechanism in which the protein moves along the DNA while maintaining constant contact with the duplex. To understand better the mechanism of sliding of the p53 protein on DNA, we measured the diffusion coefficient for the core domain, C-terminal domains, and full-length p53 in (total) salt concentrations ranging from 25 mM to 175 mM. Fig. S3 *A–C* shows the distributions of diffusion coefficients for the three constructs and Fig. 4 *A–C* shows the means and standard errors of the means of the distributions. The widths of the distributions reflect both uncertainties due to the short length of the photo-bleaching-limited trajectories and the intrinsic heterogeneity within the population of the studied single molecules. Also, the short length of some trajectories results in apparent negative diffusion coefficients. However, the large number of molecules present in the distributions allows us to determine their means with high precision. In particular, we are able to detect small shifts as a function of salt concentration. At low salt concentration, the diffusion coefficients of core domain are negligibly small, suggesting that under these conditions the core domain is effectively immobile on the DNA. At higher salt concentrations, the mean diffusion coefficient increases, suggesting a hopping mechanism for translocation of core domain along DNA. The residence time of the core domain on DNA can be calculated in different salt concentrations by comparing our observed ex-

Table 1. Diffusion coefficient of different p53 constructs

Mean diffusion coefficient ± Standard error of the mean in 75 mM total salt concentration

Full-length p53	C-terminal domain (TC)	Core domain (NCT)
$(1.62 \pm 0.17) \times 10^5$ nm^2/sec	$(7.76 \pm 0.98) \times 10^5$ nm^2/sec	$(2.39 \pm 0.48) \times 10^4$ nm^2/sec
$(1.40 \pm 0.15) \times 10^6$ bp^2/sec	$(6.71 \pm 0.58) \times 10^6$ bp^2/sec	$(2.07 \pm 0.42) \times 10^5$ bp^2/sec

Mean diffusion coefficient and standard error of the mean for full-length p53, TC domain, and NCT domain of p53. All measurements are in 75 mM total salt concentration. The C-terminal domain moves much faster on DNA than the full-length protein, while the core-domain is almost immobile on DNA.

Fig. 4. Disentangling sliding and hopping: a two-state model of p53 search on DNA. (*A*) Diffusion coefficient of core domain increases with salt concentrations suggesting a *hopping* mechanism for the core domain along DNA. The error bars in the figure are the standard error of the mean. The diffusion coefficient for full-length protein (*B*) and C-terminal domain (*C*) stays constant with salt concentrations suggesting a *sliding* mechanism for these protein constructs. (*D*) Cartoon demonstrating the two different DNA-binding modes when the p53 protein is translocating along DNA. In the *search* mode, the protein is only bound to DNA with its C-terminal domain resulting in fast sliding along DNA. In the *recognition* mode, the protein is bound to DNA with both C-terminal and core domain resulting in a slower translocation along DNA. The energy landscape for recognition mode E_R has a higher variance than the energy landscape of the search mode E_S resulting in a slower rate of translocation along DNA. p53 combines the two different binding modes when searching for its target site on DNA.

perimental diffusion coefficient with that obtained for a protein freely diffusing in solution (see *SI Text*). These values suggest that the core domain spends only 10^{-4}–10^{-3} of the total time in solution in 75 mM total salt concentration and thus is bound to the nonspecific DNA the majority of the time (>99.9%).

The C-Terminal Domain of p53 Translocates on DNA via a *Sliding* Mechanism. Fig. S3*C* shows distributions of the diffusion coefficient of the C-terminal domain diffusing along DNA in different salt concentration. Fig. 4*C* shows the dependence of the mean and standard error of the mean of the diffusion coefficient, on salt concentration. The diffusion coefficient of the C-terminal domain of p53 is independent of salt concentration and remains constant over the range of 25 mM to 75 mM total salt concentration, suggesting a sliding mechanism for translocation of the p53 C-terminal domain on DNA.

Discussion

The Sequence-Specific Core Domain of p53 Experiences a Rugged Energy Landscape, While the Nonspecific C-Terminal Domain Slides on a Smooth Energy Landscape. Our single-molecule data suggest that both full-length p53 and its C-terminal domain diffuse rapidly along DNA. Further, we demonstrate that the core domain is essentially immobilized on nonspecific DNA. These observations suggest a model in which the p53 protein slides on DNA via its C-terminal domain. The core domain, however, does not constantly maintain contact with DNA, but rather stochastically associates and dissociates on and off the DNA.

Results of the single-molecule experiments are in very good agreement with the theory of one-dimensional/three-dimensional facilitated diffusion (9) (19) (20). First, our recent theoretical study (9) predicted that the diffusion coefficient of sliding depends strongly on the ability of the protein to bind DNA in a sequence-specific manner. A DNA-binding domain with a high sequence-specificity is predicted to experience strong sequence-dependent binding energy, even on noncognate DNA, and thus unable to slide along such rugged energy landscape. However, a domain that binds with moderate specificity ($\sim 1\ k_B T$ sequence-specific energy) or nonspecifically is expected to have a relatively smooth sliding landscape and can slide fast (Fig. S4). In agree-

ment with the theory, the C-terminal domain that binds DNA nonspecifically demonstrates a rapid translocation with $\sigma = 0.6\ k_B T$, where sigma is a measure of the sequence-specific binding energy, while the sequence-specific core domain diffuses very slowly with $\sigma > 2\ k_B T$ (see *SI Text*).

Full-Length p53 Moves on DNA Through a *Two-State Mechanism* of *Search* and *Recognition*. Theoretical studies (9, 20) suggest a *two-state search mechanism*, which provides a rationale for our single-molecule measurements and demonstrates how the DNA-binding domains of p53 are coordinated. The two-state mechanism suggests that both fast search and sequence-specific recognition can be achieved if the protein has two distinct conformational states: a *search state* characterized by largely nonspecific binding and fast sliding, and a *recognition state* in which a protein binds DNA in a sequence-specific manner while unable to slide. Simulations and analytical treatment demonstrated that the target site can be rapidly found and recognized if the protein spends most of the time in the search state while frequently interrogating DNA by going into the recognition state. These two states and fast transitions between them have been observed in a range of DNA-binding proteins (19, 21) and correspond to different conformations of the same DNA-binding domain. The multidomain structure of p53 allows it to distribute the roles of these two states between the two DNA-binding domains. We propose that, in tetrameric p53, the search state corresponds to a conformation in which the C termini are bound to DNA and the core domains are undocked, thus allowing for nonspecific binding and fast translocation. The recognition state, in turn, involves docking of the core domains to DNA and specific recognition (Fig. 4*D*). Conformational switching between the two states allows both sufficiently fast translocation and specific binding. In agreement with such a two-state mechanism, the rate of translocation of the full-length proteins is a factor of five lower than that of the C terminus construct because the protein spends a certain fraction of time in the immobile recognition conformation. Using the structure of p53 revealed by electron microscopy (16), the ratio of the corresponding rotational diffusion coefficients, controlled for increased size of the full-length p53 as compared to C terminus construct, can be estimated. This approximation results in an estimate of

40–50% for the fraction of time spent in the search state (see *SI Text*). Because analysis of the full-length single-molecule trajectories was focused on mobile particles, this estimate provides an upper bound.

It is also possible to estimate the minimal rate of the conformational transition in p53 required for fast search and specific binding. Using the measured diffusion coefficient for full-length p53 and in vivo time on DNA (22, 23), we obtain a rate constant of about 10^3 s^{-1} (see *SI Text*). Thus, if the conformational transition happens on a submillisecond time scale, the protein can efficiently search for its cognate site. This prediction may be tested by H/D-exchange or similar techniques (24).

Finally, we are able to relate our single-molecule measurements to published in vivo fluorescence recovery studies (22, 23) and p53 copy-number measurements (25) to calculate the time it takes p53 to find a specific site (e.g., p21) on DNA (see *SI Text*). Assuming that only about 5% of genomic DNA is accessible due to chromatinization and that about 1,000 copies of p53 are activated, we obtain a search time in the range of 3–30 min. This estimate is consistent with about an hour for initial expression of downstream genes (25). Moreover this reasoning suggests that the latent p53 with a lifetime of about 20 min and its slow oligomerization kinetics (26, 27) is unlikely to yield significant occupancy in tetrameric form at hundreds of target promoters. The search, however, is fast enough to allow a long-lived activated form (lifetime of ~200 min) (28, 29) to bind most of target promoters.

Our framework of a one-dimensional/three-dimensional search process and our single-molecule data allow a reconciliation of seemingly contradictory studies of the role of the C terminus in p53 recognition. From a thermodynamic point of view, the C-terminal domain functions as a negative regulator for p53 by sequestering it onto nonspecific DNA. From a kinetic perspective, the C-terminal region functions as positive regulator for p53 by facilitating the search process. An optimal affinity is required for fast search and a stable cognate complex. This model explains how experimental alterations of the C-terminal domain have both positive and negative effects on p53 function. Truncation of the C terminus or binding by specific antibodies eliminates sequestration and leads to better binding to the cognate sites on short DNA fragments (5), while making binding to long DNA molecules kinetically inefficient. Sequestration to nonspecific DNA also explains the fact that long nonspecific DNA molecules inhibit binding of the full-length p53 to short cognate DNA molecules, but have no effect on C terminally truncated form (5). Moreover, modulation of affinity for nonspecific DNA can serve as a regulatory mechanism. For example, activation of p53 by acetylation of the C-terminal domain reduces its affinity for nonspecific DNA several fold, and thus can activate p53 by allowing rapid search for target sites.

In summary, we used single-molecule experiments to visualize and quantitatively characterize diffusive motion of individual p53 proteins along DNA molecules. We demonstrated that the C-terminal domain is nonspecifically bound to DNA and is capable of sliding very rapidly along DNA, while the full-length protein moves on DNA at a much slower rate. We demonstrated that single-molecule measurements are consistent with the theory of sliding, and the two-state mechanism of sliding/recognition (9, 20, 30) and proposed that while on DNA p53 rapidly interconverts between two conformations. This rapid switching allows the protein to sample sequences for specific, core-domain mediated binding, while enabling rapid search through the interaction between the C-terminal domain and DNA.

Materials and Methods

DNA Preparation and Flow Stretching. Purified DNA from λ phage (New England Biolabs) was linearized and biotinylated at one end by annealing a 3′ biotin-modified oligo (5′AGGTCGCCGCCC3′-biotin; Integrated DNA Technologies) to the complementary λ-phage 5′ overhang. Flow cells (0.1 mm height, 2.0 mm width) with a streptavidin-coated surface were prepared as described previously (17, 31, 32). The streptavidin-coated flow-cell surfaces were blocked by incubation with blocking buffer (Tris 20 mM, EDTA 2 mM, NaCl 50 mM, BSA 0.2 mg/mL, Tween 20 0.005%; pH 7.5) for 20 min. Biotin-modified DNA constructs were introduced into the flow cell at a rate of 0.1 mL/min at a concentration of 10 pM for 20 min. These conditions resulted in an average density of ~100 surface-tethered DNA molecules per field of view (~50 × 50 μm²).

The single-molecule imaging experiments were performed in an imaging buffer, containing 20 mM Hepes, 0.5 mM EDTA, 2 mM MgCl₂, 0.5 mM DTT, 0.05 mg/mL BSA (pH 7.9), and varying amounts of KCl. Imaging buffer was drawn into the channel by a syringe pump at a flow rate of 0.1 mL/min, creating shear flow near the coverslip surface (11). Single-molecule imaging was done with 30–100 pM TC (Tetramerization + C-terminal) p53 and 10–50 pM NCT (N-terminal + Core domain + Tetramerization) in imaging buffer. The proteins were kept at low-micromolar concentration, and were diluted right before the single-molecule experiment. The single-molecule experiments were done within less than 1 h from the time of dilution. Due to the slow kinetics of the tetramer-dimer transition (26), all constructs are assumed to be in the tetrameric form during the single-molecule experiment.

Protein Preparation and Labeling. *Expression and purification of TC.* P53 Tet + C (293–393) with an N-terminal cysteine was cloned in PET 24-HLTev using BamHI and EcoRI sites. The resulting plasmid encodes a fusion protein with an N-terminal 6xHis tag, followed by a lipoyl domain, a TeV protease cleavage site and the p53 Tet + C (293–393) sequence of interest. The proteins were expressed in *Escherichia coli* strain BL21 and purified by a Ni-affinity column followed by cleavage with TeV overnight. Subsequent purification by cation exchange chromatography on SP Sepharose and gel filtration on Superdex 75 yielded a purity of >99% (33). To measure the oligomerization state of the TC domain, we measured the lifetime of different TC domains as well as only the C-terminal domain on DNA. The average lifetime of the C-terminal domain on DNA is 0.88 ± 0.05 seconds, whereas the TC domain has the lifetime of 2.41 ± 0.08 seconds on DNA (Fig. S5). Both experiments are done in 25 mM total salt concentration. Because the tetramerization domain does not interact with DNA, we conclude than the TC domain in our single-molecule experiment conditions must be a dimer or tetramer.

Labeling of TC. The labeling was carried out in phosphate buffer (20 mM sodium phosphate, 150 mM NaCl, pH 7.0) with a protein concentration of 100 μM on ice. 10-fold excess Alexa Fluor 555 maleimide was added in the presence of 1 mM of tris(2-carboxyethyl) phosphine (TCEP). The labeling progress was followed by matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). The reaction was quenched with 10 mM β-mercaptoethenol after ~1 h. The mixture was then loaded onto a G-25 desalting column to separate excess dye.

Purification and labeling of NCT. The superstable mutant of NCT p53 (N-terminal + core domain + Tetramerization domain, residues 1–363) with mutations M133L, V203A, N239Y, and N268D (34) was used. The protein was expressed in *Escherichia coli* and purified as described (33, 35).

The NCT was labeled with Alexa Fluor 555 carboxylic acid succinimidyl ester (Invitrogen) through the N terminus amine. The labeling was carried out in phosphate buffer (50 mM sodium phosphate, 150 mM NaCl, pH 6.4). Alexa Fluor 555 of equal molarity was added to 1 mL of NCT solution (30 μM). The labeling progress was followed by MALDI-TOF MS. The reaction was quenched after about 1 h with 0.2 mL of 1 M Tris (hydroxymethyl) aminomethane-HCl (pH 7.4) and the labeled protein was separated from the free dye on a G-25 desalting column.

1. Vogelstein B, Lane D, Levine AJ (2000) Surfing the p53 network. *Nature* 408:307–310.
2. Wu WJ, et al. (1995) Allelic frequency of p53 gene codon 72 polymorphism in urologic cancers. *Jpn J Cancer Res* 86:730–736.
3. Oberosler P, Hloch P, Ramsperger U, Stahl H (1993) p53-catalyzed annealing of complementary single-stranded nucleic acids. *EMBO J* 12:2389–2396.

BIOPHYSICS AND COMPUTATIONAL BIOLOGY

CHEMISTRY

4. Halazonetis TD, Kandil AN (1993) Conformational shifts propagate from the oligomerization domain of p53 to its tetrameric DNA binding domain and restore DNA binding to select p53 mutants. *EMBO J* 12:5057–5064.

5. Hupp TR, Meek DW, Midgley CA, Lane DP (1992) Regulation of the specific DNA binding function of p53. *Cell* 71:875–886.

6. Prives C, Hall PA (1999) The p53 pathway. *J Pathol* 187:112–126.

7. Muller-Tiemann BF, Halazonetis TD, Elting JJ (1998) Identification of an additional negative regulatory region for p53 sequence-specific DNA binding. *Proc Natl Acad Sci USA* 95:6079–6084.

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

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

10. Elf J, Li GW, Xie XS (2007) Probing transcription factor dynamics at the single-molecule level in a living cell. *Science* 316:1191–1194.

11. Blainey PC, van Oijen AM, Banerjee A, Verdine GL, Xie XS (2006) A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. *Proc Natl Acad Sci USA* 103:5752–5757.

12. 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.

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

14. Kitayner M, et al. (2006) Structural basis of DNA recognition by p53 tetramers. *Mol Cell* 22:741–753.

15. Huang F, 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.

16. Tidow H, et al. (2007) Quaternary structures of tumor suppressor p53 and a specific p53 DNA complex. *Proc Natl Acad Sci USA* 104:12324–12329.

17. Tafvizi A, et al. (2008) Tumor suppressor p53 slides on DNA with low friction and high stability. *Biophys J* 95:L01–03.

18. Berg OG, Winter RB, von Hippel PH (1981) Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* 20:6929–6948.

19. Kalodimos CG, et al. (2004) Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science* 305:386–389.

20. Mirny LA, Wunderlich Z, Tafvizi A, Leith J, Kosmrlj A (2009) How a protein searches for its site on DNA: the mechanism of facilitated diffusion. *J Phys A-Math Theor* 42:434013.

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

22. Mueller F, Wach P, McNally JG (2008) Evidence for a common mode of transcription factor interaction with chromatin as revealed by improved quantitative fluorescence recovery after photobleaching. *Biophys J* 94:3323–3339.

23. Hinow P, et al. (2006) The DNA binding activity of p53 displays reaction-diffusion kinetics. *Biophys J* 91:330–342.

24. Kalodimos CG, Boelens R, Kaptein R (2002) A residue-specific view of the association and dissociation pathway in protein—DNA recognition. *Nat Struct Biol* 9:193–197.

25. Wang YV, et al. (2007) Quantitative analyses reveal the importance of regulated Hdmx degradation for p53 activation. *Proc Natl Acad Sci USA* 104:12365–12370.

26. Natan E, Hirschberg D, Morgner N, Robinson CV, Fersht AR (2009) Ultraslow oligomerization equilibria of p53 and its implications. *Proc Natl Acad Sci USA* 106:14327–14332.

27. Friedler A, Veprintsev DB, Hansson LO, Fersht AR (2003) Kinetic instability of p53 core domain mutants: implications for rescue by small molecules. *J Biol Chem* 278:24108–24112.

28. McVean M, Xiao H, Isobe K, Pelling JC (2000) Increase in wild-type p53 stability and transactivational activity by the chemopreventive agent apigenin in keratinocytes. *Carcinogenesis* 21:633–639.

29. Liu M, Dhanwada KR, Birt DF, Hecht S, Pelling JC (1994) Increase in p53 protein half-life in mouse keratinocytes following UV-B irradiation. *Carcinogenesis* 15:1089–1092.

30. Hu L, Grosberg AY, Bruinsma R (2008) Are DNA transcription factor proteins maxwellian demons? *Biophys J* 95:1151–1156.

31. Lee JB, et al. (2006) DNA primase acts as a molecular brake in DNA replication. *Nature* 439:621–624.

32. van Oijen AM, et al. (2003) Single-molecule kinetics of lambda exonuclease reveal base dependence and dynamic disorder. *Science* 301:1235–1238.

33. Weinberg RL, Freund SM, Veprintsev DB, Bycroft M, Fersht AR (2004) Regulation of DNA binding of p53 by its C-terminal domain. *J Mol Biol* 342:801–811.

34. Nikolova PV, Henckel J, Lane DP, Fersht AR (1998) Semirational design of active tumor suppressor p53 DNA binding domain with enhanced stability. *Proc Natl Acad Sci USA* 95:14675–14680.

35. Veprintsev DB, et al. (2006) Core domain interactions in full-length p53 in solution. *Proc Natl Acad Sci USA* 103:2115–2119.

Supporting Information

Tafvizi et al. 10.1073/pnas.1016020107

SI Text

DNA Preparation and Flow Stretching. Purified DNA from λ phage (New England Biolabs) was linearized and biotinylated at one end by annealing a 3′ biotin-modified oligo (5′AGGTCGCCG-CCC3′-biotin; Integrated DNA Technologies) to the complementary λ-phage 5′ overhang. Flow cells (0.1 mm height, 2.0 mm width) with a streptavidin-coated surface were prepared as described previously (1–3). The streptavidin-coated flow-cell surfaces were blocked by incubation with blocking buffer (Tris 20 mM, EDTA 2 mM, NaCl 50 mM, BSA 0.2 mg/mL, Tween 20 0.005%; pH 7.5) for 20 min. Biotin-modified DNA constructs were introduced into the flow cell at a rate of 0.1 mL/ min at a concentration of 10 pM for 20 min. These conditions resulted in an average density of ∼100 surface-tethered DNA molecules per field of view (∼50×50 μm^2).

The single-molecule imaging experiments were performed in an imaging buffer, containing 20 mM Hepes, 0.5 mM EDTA, 2 mM MgCl$_2$, 0.5 mM DTT, 0.05 mg/mL BSA (pH 7.9), and varying amounts of KCl. Imaging buffer was drawn into the channel by a syringe pump at a flow rate of 0.1 mL/ min, creating shear flow near the coverslip surface (4). Single-molecule imaging was done with 30–100 pM TC (Tetramerization + C-terminal) p53 and 10–50 pM NCT (N-terminal + Core domain + Tetramerization) in imaging buffer. The proteins were kept at low-micromolar concentration, and were diluted right before the single-molecule experiment. The single-molecule experiments were done within less than 1 h from the time of dilution. Due to the slow kinetics of the tetramer-dimer transition (5), all constructs are assumed to be in the tetrameric form during the single-molecule experiment.

Protein Preparation and Labeling. *Expression and purification of TC.*

P53 Tet + C (293–393) with an N-terminal cysteine was cloned in PET 24-HLTev using BamHI and EcoRI sites. The resulting plasmid encodes a fusion protein with an N-terminal 6xHis tag, followed by a lipoyl domain, a TeV protease cleavage site, and the p53 Tet + C (293–393) sequence of interest. The proteins were expressed in *Escherichia coli* strain BL21 and purified by a Ni-affinity column followed by cleavage with TeV overnight. Subsequent purification by cation exchange chromatography on SP Sepharose and gel filtration on Superdex 75 yielded a purity of >99% (6). To measure the oligomerization state of the TC domain, we measured the lifetime of different TC domains as well as only the C-terminal domain on DNA. The average lifetime of the C-terminal domain on DNA is 0.88 ± 0.05 seconds, whereas the TC domain has the lifetime of 2.41 ± 0.08 seconds on DNA. Both experiments are done in 25 mM total salt concentration. Because the tetramerization domain does not interact with DNA, we conclude than the TC domain in our single-molecule experiment conditions must be a dimer or tetramer.

Labeling of TC.

The labeling was carried out in phosphate buffer (20 mM sodium phosphate, 150 mM NaCl, pH 7.0) with a protein concentration of 100 μM on ice. 10-fold excess Alexa Fluor 555 maleimide was added in the presence of 1 mM of tris(2-carboxyethyl) phosphine (TCEP). The labeling progress was followed by matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). The reaction was quenched with 10 mM β-mercaptoethenol after ∼1 h. The mixture was then loaded onto a G-25 desalting column to separate excess dye.

Purification and labeling of NCT.

The superstable mutant of NCT p53 (N-terminal + core domain + Tetramerization domain, residues 1–363) with mutations M133L, V203A, N239Y, and N268D (7) was used. The protein was expressed in *Escherichia coli* and purified as described (6, 8).

The NCT was labeled with Alexa Fluor 555 carboxylic acid succinimidyl ester (Invitrogen) through the N terminus amine. The labeling was carried out in phosphate buffer (50 mM sodium phosphate, 150 mM NaCl, pH 6.4). Alexa Fluor 555 of equal molarity was added to 1 mL of NCT solution (30 μM). The labeling progress was followed by MALDI-TOF MS. The reaction was quenched after about 1 h with 0.2 mL of 1 M Tris (hydroxymethyl) aminomethane-HCl (pH 7.4) and the labeled protein was separated from the free dye on a G-25 desalting column.

Fluorescence Imaging.

Fluorescence imaging of the movement of the labeled p53 proteins and its different domains along DNA was performed by placing the flow cell on top of an inverted microscope (Olympus IX71) and exciting the AlexaFluor 555 label by the 532-nm line of a frequency doubled, Nd:YAG laser (Crystal laser, GCL-100-M) with 100 mw maximum output. A high-N.A. microscope objective (Olympus, 60×1.6x/1.45 N.A.) was used to illuminate the sample with total-internal reflection. The illuminated area had a diameter of 50 μm at the sample plane. The fluorescence was collected by the same objective and imaged by an EM-CCD digital camera (Hamamatsu C9100-13), after filtering out scattered laser light. The movies were recorded by MetaVue imaging software.

Particle Tracking.

The positions of labeled particles were determined by fitting each single-molecule fluorescence image to a two-dimensional Gaussian distribution. We calculate the standard error of position determination to be $\sigma = 10$–20 nm (3).

The trajectories of the particles were then tracked by custom-written particle-tracking MATLAB code. The trajectories were further analyzed to measure the drift and diffusion coefficient of each protein in the population (3).

Site-Specific Labeling of Biotin-Lambda.

We use a previously published protocol (9, 10) to site-specifically label the biotin-lambda DNA at 14,725 base-pairs from the tethered point. Treatment of biotin-lambda DNA with the sequence-specific Nt.Bst.NBI nicking endonuclease results in single-strand breaks corresponding to the recognition site of the enzyme. Some of the nicks occur closely spaced and on the same strand. The oligonucleotides between closely spaced nicks can be replaced with modified oligonucleotides of the same sequence via a strand-displacement reaction and subsequent ligation. Nicking reactions with Nt.Bst.NBI (NEB, 20 units) were performed on biotin-lambda DNA (2.5 μg) at 50 °C for 2 h in buffer 3 (NEB). The nicked DNA was mixed with 100-fold excess of a 13-base oligonucleotide (IDT) modified with a 5′ digoxigenin (5′-digoxigenin-TTCA-GAGTCTGAC-3′), heated at 50 °C for 10 min and then slowly cooled to room temperature, resulting in the highly efficient replacement of the native sequence. Ligation was performed for 2 h at room temperature using T4 ligase. At the end, the activity of the nicking enzyme was quenched by adding 20 mM EDTA. The digoxigenin was detected by flowing in and then washing out a solution of 1 nM quantum dots (QDot; Invitrogen) which were functionalized with antidigoxigenin Fab fragment (Roche) using the Invitrogen QDot Antibody Conjugation Kit. The DNA was

stretched and the QDot fluctuations characterized using the same flow conditions as described in the p53 experiments.

Determination of Drift Rates. We measure the presence of any directional bias in the protein motions caused by the flow of the buffer by measuring the total displacement of all protein trajectories divided by the total duration of all trajectories. In the absence of any drift, the net displacement per time unit for a population of proteins is a normal distribution around zero. However, in the presence of buffer flow, a drag force will be exerted on the protein and will force it to move in the direction of the flow. Therefore, the net displacement of the population of proteins is shifted in the direction of the flow. To take into account the fact that trajectory lengths are different for every measured protein, we measure the weighted drift of our protein populations as follows:

$$\text{drift}_{\text{mean,weighted}} = \frac{\sum_{i=1}^{N} l_i d_i}{\sum_{i=1}^{N} l_i},$$

where l_i is the length of trajectory I and d_i is the measured drift for that trajectory. The variance of all the trajectories in one biochemical condition can likewise be calculated by:

$$\sigma^2_{\text{weighted}} = \frac{\sum_{i=1}^{N} l_i (d_i - \text{drift}_{\text{mean,weighted}})^2}{\sum_{i=1}^{N} l_i}.$$

The error bars can be calculated as $\frac{\sigma}{\sqrt{N}}$ where N is the total number of trajectories in that biochemical condition. Fig. S1 shows the drift histogram for core and C-terminal domain in 75 mM total salt concentration for 385 and 183 trajectories respectively. The drift for core domain is almost centered around zero, resulting in a small mean of 6.1 ± 7.6 nm/seconds. The weighted mean of drift for C-terminal domain is 355.8 ± 98.4 nm/seconds. A similar distribution of drift is observed for other salt concentrations.

Calculating Diffusion Coefficient by Minimizing the Effect of DNA Fluctuations. To calculate the diffusion coefficients for the different p53 constructs while minimizing the impact from DNA fluctuations, we only fit the portion of the Mean Square Displacement (MSD) plot between 0.25 to 0.5 s, a region in which the MSD of the DNA fluctuations is constant and does not contribute to the diffusion coefficient (3). Furthermore, we exclude trajectories obtained from proteins on the one-third of the DNA farthest from the tether point (between 33,777 and 48,502 bp), as the underlying DNA fluctuations become too large compared to the diffusion coefficient to extract reliable diffusional properties (Fig. S2). The histogram of diffusion coefficient for different constructs of p53 are shown in Fig. S3.

Three-Dimensional Diffusion Coefficient of Different Domains. The diffusion coefficient for a spherical object diffusing by purely translational movement in one dimension can be calculated by the Stokes-Einstein relation:

$$D_{3d} = \frac{k_B T}{6\pi\eta a}.$$

Where η is the solvent viscosity (8.9×10^{-4} Pa·s for water at 25 °C), a is the radius of the diffusing tetrameric p53 protein (4.9 nm) (3), k_B is the Boltzman constant, and T is the temperature. For tetrameric p53 in our experimental conditions, this calculation results in a three-dimensional diffusion coefficient of 4.8×10^7 nm²/seconds. The diffusion coefficient of the TC (Tetramerization + C-terminal domain) and NCT (N-terminal +

Core + Tetramerization domain) can as well be calculated using the molecular weight of these domains, which results in 4.9×10^7 nm²/seconds and 7.5×10^7 nm²/seconds for NCT, and TC domains respectively.

Residence Time of Core Domain on DNA. The salt dependence of the diffusion coefficient of the core domain implies a hopping mechanism for diffusion of the core domain on DNA. The core domain therefore translocates on DNA by undergoing microscopic association and dissociations on and off the DNA. Upon dissociation, the protein will diffuse a short distance in solution, resulting in rebinding to the DNA at a different position. During these brief excursions into solution, the protein will also experience a small hydrodynamic drag from the buffer flow, resulting in the drift of the C-terminal truncated protein. The dependence of the diffusion coefficient of the core domain on salt concentration can be rationalized by the amount of the time that the protein spends in solution. At higher salt concentration, the protein spends longer time in solution, resulting in the increase of diffusion coefficient of core domain. The fraction of time that the protein spends in solution can be determined by comparing our observed experimental diffusion coefficient with that obtained for a protein freely diffusing in solution. For the NCT domain, the diffusion coefficient in solution is 4.9×10^7 nm²/seconds and the experimental diffusion coefficient is measured to be $(2.39 \pm 0.48) \times 10^4$ nm²/seconds (at 75 mM total salt concentration). These values suggest that the core domain spends only 10^{-4}–10^{-3} of the total time in solution and thus is bound to the nonspecific DNA the majority of the time (>99.9%).

Measuring the Time That Core Domain Spends in Solution Based on Drift. The velocity of buffer at the DNA can be estimated using the flow rate of the buffer and the dimensions of the channel. In our flow-stretching assay, the buffer solution was drawn into the channel by a syringe pump with a flow rate of 0.1 mL/min creating a shear flow near the coverslip surface. The flow channel is 100 μm in height and 2 mm in width, which results in the average velocity of 0.83 cm/sec for the buffer in the channel. The flow velocity, however, is not a constant throughout the channel, but is zero at the boundaries, yielding a parabolic flow profile (3). The mean distance of the DNA from the coverslip surface is 0.2 μm (3, 4, 11). With a channel height h, the flow velocity v_y at a distance y from the surface can be expressed as:

$$v_{\text{avg}} = \frac{2}{3} v_{\text{max}} \quad \text{and} \quad v_y = v_{\text{max}} \left[\frac{hy - y^2}{h^2/4} \right] = \frac{3}{2} v_{\text{avg}} \left[\frac{hy - y^2}{h^2/4} \right].$$

The average velocity of the flow at the center of the DNA, 0.2 μm above the surface, can be estimated as 100 μm/sec. Assuming that the core domain only moves when it is dissociated from the DNA, the drift of the core domain can be estimated as the product of the ratio of the time it spends in the solution and the velocity of the buffer at the DNA. Measured from the diffusion coefficient, the core domain spends only 10^{-4}–10^{-3} of the total time in solution and thus is bound to the nonspecific DNA the majority of the time (>99.9%). The small time spent in solution will result in drift of about 10 nm/sec which is within the error bar of the observed experimental drift.

Comparing the Ruggedness of Energy Landscape of Translocating Along DNA for Different p53 Domains and Their Affinity to Nonspecific DNA. Difference in the ruggedness of the sequence-specific energy landscape between the core and C-terminal domains is also consistent with different affinities of these domains to nonspecific DNA. Theory (12) gives the residence time of a domain on DNA as

$$\tau_{1D} \sim \exp(E_{ns}/k_B T + \sigma^2/(2(k_B T)^2)),$$

where $E_{ns} > 0$ is the free energy difference for the protein to be in the solvent and on DNA, and σ is the measure of the landscape ruggedness. This estimate predicts that sequence-specific core domain (large σ) spends more time on nonspecific DNA as compared to the C-terminal domain. Consistent with this argument the Kd for nonspecific DNA for C-terminal (Tet + C-terminal) and core domain (Tet + Core) are several µM and 680 nM respectively (6, 13). These numbers show a higher affinity for the core domain to nonspecific DNA compared to the C-terminal domain. We observed a similar behavior in our single-molecule experiments. While the lifetime of the core domain constructs on DNA is very long (several minutes, limited by photo-bleaching), the TC domain is very short lived on DNA (about 3 s for TC, Fig. S4). This shows that the affinity of the core domain to the nonspecific DNA is higher than that of the C-terminal domain.

Energy Barriers of Translocation for C-Terminal Domain and Core Domain. The energy barriers of translocation for the sliding of core domain and C-terminal domain can be calculated by comparing the measured experimental diffusion coefficient and the maximum rotational diffusion coefficient for TC and NCT domains by (12):

$$D_{exp} = D_{theor, lim}(1 + \sigma^2\beta^2)^{1/2} \exp(-7\sigma^2\beta^2/4).$$

This calculation results in a roughness of the energy of $0.6\ k_B T$ for C-terminal domain and $2.3\ k_B T$ for core domain sliding along DNA. The slow diffusion of core domain, which is the sequence-specific domain of p53, on DNA is consistent with theory where sequence specificity requires high variance ($\sigma > 2\ k_B T$) of the energy landscape. Also the rapid diffusion of C-terminal domain, which nonspecifically binds DNA, is consistent with theory where small roughness of the energy landscape ($\sigma < k_B T$) is required for fast search (12) (Fig. S5). The full-length p53 requires both fast search and sequence-specificity for the target site and its diffusion can be explained in the context of a two state model (See main text).

Estimates for p53 Search Process. *Fraction of time in the search state.* Diffusion of the full-length p53 is slower than that of the C terminus due to two factors: (*i*) difference in the size of the protein; (*ii*) some fraction of time spent in the immobile recognition conformation:

$$D_{fl, exp} = \gamma \cdot f \cdot D_{Cterm, exp},$$

where γ is the ratio of the rotational diffusion coefficient of the two proteins, and f is the fraction of time spent in the search state. Recently the quaternary structure of p53 on the cognate DNA site has been solved by a combination of small-angle X-ray scattering and NMR (14). The structure shows that the core domains are in close contact with the DNA, while the tetramerization domains form the most remote region of the structure on the other side of DNA. C-terminal domains have not been resolved (must be disorder) but are likely to interact with nonspecific DNA on the same side as the tetramerization, i.e., opposite of the core domains. We estimate the difference in the theoretical limit of the rotational diffusion coefficient between the TC (tetramerization and C-terminal domain) construct (293–393 = 100 aa) and the fl p53 (393 aa). The upper limit of the rotational diffusion coefficient can be calculated by:

$$D_{lim} = \frac{k_B T}{\xi}.$$

Where ζ is the total friction of the protein rotating along the DNA helix and can be calculated by (15):

$$\zeta = 6\pi\eta R + \left(\frac{2\pi}{10BP}\right)^2 [8\pi\eta R^3 + 6\pi\eta R(R_{OC})^2]$$

for a protein with radius R and distance R_{OC} from DNA. From Fig. S6, for the full-length p53 the total friction is the sum of the frictions of the tetramerization domain with radius r and distance r_{oc} from DNA and the core domain with radius R and distance R_{OC} from the DNA. The ratio of the upper limit of rotational diffusion for TC domain and full-length p53 can be calculated as:

$$\gamma = \frac{D_{lim, fl}}{D_{lim, TC}} = \frac{\xi_{TC}}{\xi_{fl}},$$

where D_{lim} is the limit of the rotational diffusion for full-length protein or TC domain. By inserting the relevant radia this ratio is measured to be $\gamma = 0.45$.

Therefore the fraction of time in the sliding mode can be calculated as:

$$f = D_{fl, exp}/(\gamma . D_{Cterm, exp}) = 1.4/(6.7 \times 0.45) = 0.46.$$

So p53 spends about half of the time in the sliding state and half in the recognition state. Because our estimate of the diffusion coefficient for the full-length p53 is biased toward higher end, provided estimate is an upper estimate (i.e., protein spends less than 45% in the sliding state)

Sliding length and search time. To measure the mean sliding distance and search time we use (12)

$$n = \sqrt{4D_{1D}/k_{off}},$$

where n is the number of base-pairs visited in each round of one-dimensional diffusion, D_{1D} is the measured diffusion coefficient of full-length p53 on DNA ($D_{1D} = 1.4 \times 10^6\ bp^2$/seconds) (Table 1). The diffusion coefficient in cell nucleus is about five times smaller due to the higher viscosity of the nucleus (16). k_{off} is the dissociation rate of p53 from nonspecific DNA ($k_{off} = 0.4\ s^{-1}$) (17, 18). Therefore the number of base-pairs visited in each round of one-dimensional diffusion can be estimated as

$$n = \sqrt{4D_{1D}/k_{off}} = 2\sqrt{1.4 \times 10^6/5/0.4} = 1,700\ bp.$$

The total search time can be measured by

$$t_s = \frac{M}{n}\left(\frac{1}{k_{on}} + \frac{1}{k_{off}}\right),$$

where k_{on} is the association rate of full-length p53 to nonspecific DNA ($k_{on} = 0.12$–$0.3\ s^{-1}$) (17, 18) and M is the length of the genome ($M = 2 \times 3 \times 10^9$ bps) for human genome. The fraction of accessible DNA due to chromatinization is 1–5% (19). Therefore the total search time can be estimated as:

$$t_s = \frac{M}{n}\left(\frac{1}{k_{on}} + \frac{1}{k_{off}}\right) = \frac{6.10^6(0.01\text{–}0.05)}{1,700}\left(\frac{1}{0.12\text{–}0.3} + \frac{1}{0.4}\right)$$
$$= (20\text{–}200) \times 10^4\ \text{sec}$$

This result is for one copy of p53 per cell. However the copy number of p53 in nucleus is much higher and is estimated to be 500–5,000 (20, 21). The number of active p53 proteins increases at the presence of different stress types. We assume 1,000 as the number of activated p53 molecules in cell nucleus. The total search time

for a single site (e.g., p21 promoter) to be found by any p53 molecule is therefore:

$$t_s = (200\text{–}2{,}000) \ \sec \approx 3\text{–}30 \ \min.$$

This estimate is consistent with p21 initial activation response of about 1 h.

In chromatinized DNA, then presence of nucleosomes can reduce the sliding distance to the size of the nucleosome-free region in the promoter $\sim$500 bp. Then the entire search takes about $1{,}700/500 = 3.1$ fold longer.

The rate of the conformational transition. Efficient search requires that the cognate site is recognized at least once during a round of sliding. This requirement does not mean that the proteins shall undergo the conformational transition each time it visits every site. Each site is visited many ($\sim n$ times). It is imperative that the protein undergoes the conformational transition at least once

during all the visits of a site, i.e., during an interval of time $\sim \tau_{1D}/n$:

$$\tau_{\text{residence}} \approx \frac{\tau_{1D}}{n} = \frac{1}{k_{\text{off}}n} = \frac{1}{0.4 \cdot 1{,}700} = 1.5 \ \text{ms}.$$
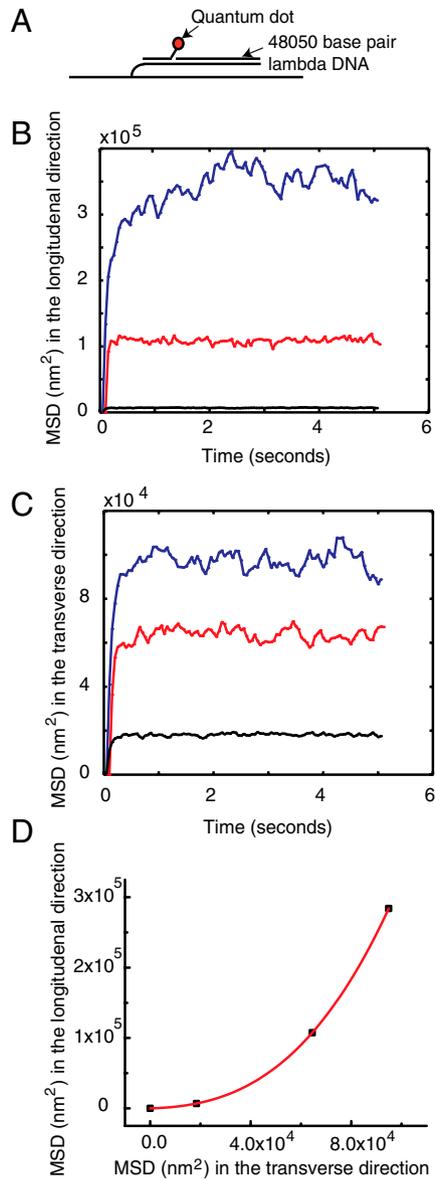
Fast search if conformational transition takes less than residence time:

$$\tau_{\text{conf}} < \tau_{\text{residence}}.$$

The rate of the conformational transition from search state to recognition state k_{SR} (Fig. 4, main text) shall be greater then $700 \ \text{s}^{-1}$, i.e., millisecond range, which is a very reasonable requirement even for such big tetramer as p53 (LacI was shown to undergo a transition in the submilisecond range). The transition from recognition state to search state, k_{RS} is determined by the stability of the complex of p53 and DNA.

1. Lee JB, et al. (2006) DNA primase acts as a molecular brake in DNA replication. *Nature* 439:621–624.
2. van Oijen AM, et al. (2003) Single-molecule kinetics of lambda exonuclease reveal base dependence and dynamic disorder. *Science* 301:1235–1238.
3. Tafvizi A, et al. (2008) Tumor suppressor p53 slides on DNA with low friction and high stability. *Biophys J* 95:L01–03.
4. Blainey PC, van Oijen AM, Banerjee A, Verdine GL, Xie XS (2006) A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. *Proc Natl Acad Sci USA* 103:5752–5757.
5. Natan E, Hirschberg D, Morgner N, Robinson CV, Fersht AR (2009) Ultraslow oligomerization equilibria of p53 and its implications. *Proc Natl Acad Sci USA* 106:14327–14332.
6. Weinberg RL, Freund SM, Veprintsev DB, Bycroft M, Fersht AR (2004) Regulation of DNA binding of p53 by its C-terminal domain. *J Mol Biol* 342:801–811.
7. Nikolova PV, Henckel J, Lane DP, Fersht AR (1998) Semirational design of active tumor suppressor p53 DNA-binding domain with enhanced stability. *Proc Natl Acad Sci USA* 95:14675–14680.
8. Veprintsev DB, et al. (2006) Core domain interactions in full-length p53 in solution. *Proc Natl Acad Sci USA* 103:2115–2119.
9. Kochaniak AB, et al. (2009) Proliferating cell nuclear antigen uses two distinct modes to move along DNA. *J Biol Chem* 284:17700–17710.
10. Kuhn H, Frank-Kamenetskii MD (2008) Labeling of unique sequences in double-stranded DNA at sites of vicinal nicks generated by nicking endonucleases. *Nucleic Acids Res* 36:e40.
11. Doyle PS, Ladoux B, Viovy JL (2000) Dynamics of a tethered polymer in shear flow. *Phys Rev Lett* 84:4769–4772.
12. Slutsky M, Mirny LA (2004) Kinetics of protein-DNA interaction: facilitated target location in sequence-dependent potential. *Biophys J* 87:4021–4035.
13. Ang HC, Joerger AC, Mayer S, Fersht AR (2006) Effects of common cancer mutations on stability and DNA binding of full-length p53 compared with isolated core domains. *J Biol Chem* 281:21934–21941.
14. Tidow H, et al. (2007) Quaternary structures of tumor suppressor p53 and a specific p53 DNA complex. *Proc Natl Acad Sci USA* 104:12324–12329.
15. Bagchi B, Blainey PC, Xie XS (2008) Diffusion constant of a nonspecifically bound protein undergoing curvilinear motion along DNA. *J Phys Chem B* 112:6282–6284.
16. Phair RD, Gorski SA, Misteli T (2004) Measurement of dynamic protein binding to chromatin in vivo, using photo-bleaching microscopy. *Methods Enzymol* 375:393–414.
17. Mueller F, Wach P, McNally JG (2008) Evidence for a common mode of transcription factor interaction with chromatin as revealed by improved quantitative fluorescence recovery after photobleaching. *Biophys J* 94:3323–3339.
18. Hinow P, et al. (2006) The DNA-binding activity of p53 displays reaction-diffusion kinetics. *Biophys J* 91:330–342.
19. Hesselberth JR, et al. (2009) Global mapping of protein-DNA interactions in vivo by digital genomic footprinting. *Nat Methods* 6:283–289.
20. Kuznetsov VA, Orlov YL, Wei CL, Ruan Y (2007) Computational analysis and modeling of genome-scale avidity distribution of transcription factor binding sites in chip-pet experiments. *Genome Inform* 19:83–94.
21. Wang YV, et al. (2007) Quantitative analyses reveal the importance of regulated Hdmx degradation for p53 activation. *Proc Natl Acad Sci USA* 104:12365–12370.

Fig. S1. Histogram of drift for different DNA-binding domains of p53. (*A*) Histogram of drift for 385 individual NCT trajectories. The drift for NCT domain is almost centered around zero, resulting in a small mean of 6.1 ± 7.6 nm/sec. (*B*) Histogram of drift for 183 individual TC domain trajectories. The weighted drift for TC domain is 355.8 ± 98.4 nm/seconds. Similar distribution of drift is observed for other salt concentrations.

Fig. S2. Measurements of the Mean Square Displacements of DNA fluctuations. (*A*) The fluctuation of DNA due to Brownian motion can be measured by attaching a quantum dot to different locations along DNA. (*B*) MSD vs. time in the direction of the flow of quantum dots attached to 14,725 bp (black), 33,777 bp (red) and 48,502 bp (blue) on DNA. As can be seen from the graphs, the MSD of first two quantum dots doesn't increase after time = 0.25 seconds. (*C*) MSD vs. time in the direction transverse to the direction of the flow for the same quantum dots trajectories. (*D*) MSD of bound quantum dots in the longitudinal direction and transverse direction at time $t = 0.5$ seconds. Only trajectories with MSD less than 7×10^4 nm²/seconds in the transverse direction are chosen to assure the minimal effect of DNA fluctuations in the longitudinal direction on measured diffusion coefficient.
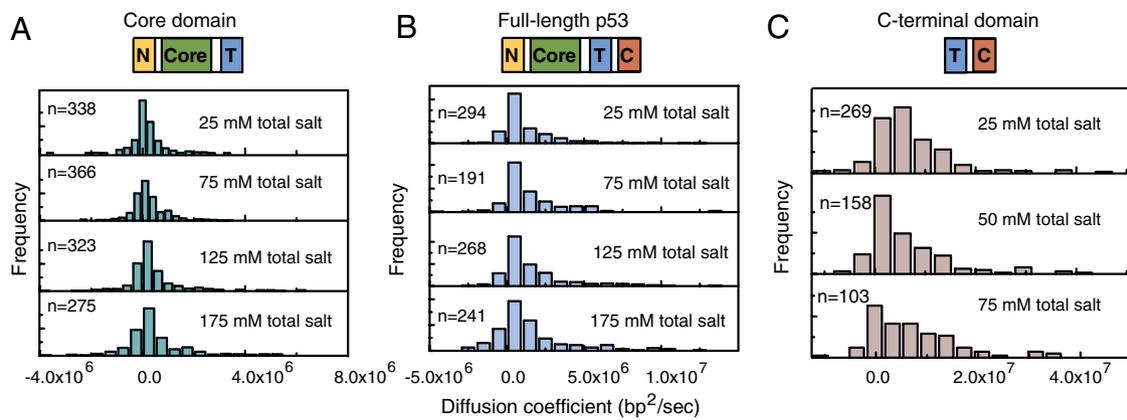
Fig. S3. Histogram of diffusion coefficient on DNA for different p53 constructs. (*A*) Histogram of diffusion coefficient of NCT in different total salt concentration; the number of trajectories *n* are noted in each histogram. (*B*) Histogram of diffusion coefficient of full-length p53 protein in different total salt concentration. (*C*) Histogram of diffusion coefficient of TC domain in different total salt concentration.
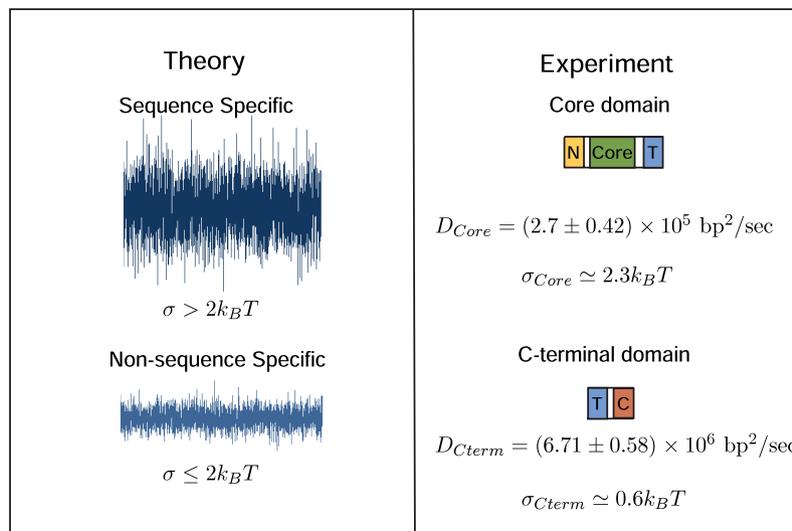
Fig. S4. Effect of the energy landscape on the rate of translocation of different DNA-binding domains: comparison of theoretical predictions and experimental observations. Sequence specificity of the protein-DNA interactions requires rough energy landscapes with large variance σ. The slow diffusion of the core domain, which is the sequence-specific domain of p53, is consistent with this theory and results in $\sigma = 2.3\ k_BT$. C-terminal domain on the other hand is nonsequence dependant and binds all DNA sequences. Theory suggests a small roughness of the energy landscape for such protein-DNA interactions. The fast translocation of the C-terminal domain on DNA is consistent with theory and results in $\sigma = 0.6\ k_BT$.

Fig. S5. Comparison of the lifetime of C-terminal domain and TC domain on DNA. (*A*) Histogram of lifetime of C-terminal domain on DNA for 285 individual trajectories in 25 mM total salt concentration. The average lifetime of C-terminal domain on DNA is 0.88 ± 0.05 seconds. (*B*) Histogram of lifetime of TC domain on DNA for 429 individual trajectories in 25 mM total salt concentration. The average lifetime of TC domain on DNA is 2.41 ± 0.08 seconds. Total acquisition time is 5 s. The asterisk indicates that the frequency shown represents lifetime for all values equal or greater than the indicated lifetime. The average lifetime of the TC domain is much larger than that of C-terminal domain alone. Because Tetramerization domain does not interact with DNA, this suggests that the TC domain in our experimental condition is oligomerized resulting in interactions of more than one C-terminal domain with DNA.

Fig. S6. The structure of p53 used in our model of p53 search on DNA. (*A*) The quaternary structure of p53 on the cognate DNA site by a combination of small-angle X-ray scattering and NMR (14). The structure shows that the core domains are in close contact with the DNA, while the tetramerization domains form the most remote region of the structure on the other side of DNA. The maximum theoretical diffusion coefficient with no energy barrier for TC domain (*B*) and full-length protein (*C*) can be calculated by rotational diffusion of the constructs along DNA.