A single-molecule characterization of p53 search on DNA

Anahita Tafvizi,a,b Fang Huang,c1 Alan R. Fersht,d,e Leonid A. Mirny,e,f A.M. van Oijen,b,e, and Antoine M. van Oijenb,a

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

Contributed by Alan R. Fersht, October 27, 2010 (sent for review August 9, 2010)

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.

A single-molecule characterization of p53 search on DNA

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.
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 in intrinsic DNA fluctuations. The drift was determined from individual trajectories 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 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. 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.
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 contacts in contact with DNA while translocating along it. On the other hand, a hopping protein is suggested to make microscopic contacts in contact with DNA while translocating along it. On the other hand, a hopping protein is suggested to make microscopic contacts with DNA while translocating along it. 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**

<table>
<thead>
<tr>
<th>p53 Constructs</th>
<th>C-terminal domain (TC)</th>
<th>Core domain (NCT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full-length p53</td>
<td>(1.62 ± 0.17) × 10^5 nm^2/sec</td>
<td>(7.76 ± 0.98) × 10^5 nm^2/sec</td>
</tr>
<tr>
<td></td>
<td>(1.40 ± 0.15) × 10^5 bp^2/sec</td>
<td>(6.71 ± 0.58) × 10^5 bp^2/sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.39 ± 0.48) × 10^6 nm^2/sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.07 ± 0.42) × 10^6 bp^2/sec</td>
</tr>
</tbody>
</table>

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.
The C-Terminal Domain of p53 Translocates on DNA via a Sliding Mechanism. Fig. S3C shows distributions of the diffusion coefficient of the C-terminal domain diffusing along DNA in different 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 (~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 agreement 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 domaindiffuses 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 multidiomain 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 unbound, 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. 4D). 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 $10^7 \text{s}^{-1}$ (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 $\sim 200$ 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 a positive regulator for p53 by facilitating the search process. An optimal affinity is required for fast search and a stable cognate complex. This 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/recognized (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′ AGTTGCGCCGCCC-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 $\sim 100$ surface-tethered DNA molecules per field of view ($\sim 50 \times 50$ 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 (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 p53. TC 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 6×His 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 $\sim 95\%$ (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. 53). Both experiments are done in 25 mM total salt concentration. Because the tetramerization domain does not interact with DNA, we conclude that 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 mercaptoethanol 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 (as described (33, 35)).

The NCT was labeled with Alexa Fluor 555 carboxylic acid succinimidyl ester (Invitrogen) through the N terminal 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) amino- methane-HCl (pH 7.4) and the labeled protein was separated from the free dye on a G-25 desalting column.

Acknowledgments. A.T. thanks Dr. J. L. Loparo for his help with the quantum-dot experiments, and J.S. Leith for helpful discussions. A.M. and O. acknowledges support from the National Science Foundation and the National Institutes of Health. L.A.M. acknowledges support from the National Cancer Institute through Physical Sciences in Oncology Center at MIT.


