Single-Molecule Studies of Unlabelled Full-Length p53 Protein Binding to DNA

Philippa Nuttall,¹ Kidan Lee,² Pietro Ciccarella,³ Marco Carminati,³ Giorgio Ferrari,³ Ki-Bum Kim,² Tim Albrecht¹*

¹Imperial College London, Department of Chemistry, Exhibition Road, London SW7 2AZ, UK

²Department of Materials Science and Engineering, Seoul National University, Seoul 151-742, Korea

³Politecnico di Milano, Dipartimento di Elettronica, Informazione e Bioingegneria, P.za Leonardo da Vinci 32, Milano, Italy

*t.albrecht@imperial.ac.uk
ABSTRACT: p53 is an anti-tumor protein that plays an important role in apoptosis, preserving genomic stability and preventing angiogenesis, and it has been implicated in a large number of human cancers. For this reason it is an interesting target for both fundamental studies, such as the mechanism of interaction with DNA, and applications in biosensing. Here, we report a comprehensive study of label-free, full length p53 (flp53) and its interaction with engineered double-stranded DNA in vitro, at the single-molecule level, using Atomic Force Microscopy (AFM) imaging and solid-state nanopore sensing. AFM data show that dimeric and tetrameric p53 bind to the DNA in a sequence-specific manner, confirming previously reported relative binding affinities. The statistical significance is tested using both the Grubbs test and stochastic simulations. For the first time, ultra-low noise solid-state nanopore sensors are employed for the successful differentiation between bare DNA and p53/DNA complexes. Furthermore, translocation statistics reflect the binding affinities of different DNA sequences, in accordance with AFM data. Our results thus highlight the potential of solid-state nanopore sensors for single-molecule biosensing, especially when labelling is either not possible or at least not a viable option.

INTRODUCTION

The tumor suppressor protein p53 is a transcription factor which regulates the cells’ response to stress, such as DNA damage, UV-induced damage or elevated oncogene expression. By binding to specific response elements (REs) on DNA in its many gene targets, p53 mediates transcriptional activation, leading to cell-cycle arrest, DNA-repair, or apoptosis. More than 50% of human cancers have been found to have mutations of full length p53 (flp53) protein, with the majority found in the sequence-specific DNA-binding domain, which mediates recognition and binding to specific REs. This highlights the importance of the p53 interaction with DNA. The flp53 protein consists of four distinct
domains: (1) The unstructured N-terminal trans-activation domain, (2) the central regions known as the DNA-binding domain (DBD), (3) the tetramerization domain, and (4) the C-terminal domain which has been shown to bind non-specifically to DNA. p53 binds to dsDNA sequences known as 'consensus sites', which consist of two copies of a 10 bp motif, or 'half site', 5'-RRRCWWGYYY-3' separated by 0-13 bp. Here, R = A,G; W = A,T; and Y = C,T. flp53 REs in vivo have fewer spacer sequences between half-sites,\(^5,6\) while transcriptional activation has been found for half-sites and three quarter-sites, although this requires tetrameric p53.\(^7\) p53 forms dimers in solution, with a dissociation constant of 20 nM, and is predominantly dimeric at low concentrations. At high concentrations, tetramerization occurs naturally.\(^8\) In cells, there exists an equilibrium between dimers and tetramers. A step-wise dimer binding mechanism occurs in the instance of specific binding of p53 to consensus DNA sequences such that consensus DNA appears to induce the formation of a tetramer of p53 on DNA.\(^8\) A complex regulatory network therefore exists where the preference for dimeric or tetrameric p53 may be linked to function, and consensus sequences may induce binding with only a half-site. Additionally, sliding of p53 has been shown to be mediated by the C-terminus, with a switching mechanism between non-specific and specific binding.

Single molecule detection methods, such as fluorescence microscopy,\(^9\) total internal reflection fluorescence (TIRF) microscopy\(^10\) and single molecule FRET\(^11\) have recently been utilised to probe labelled p53-DNA complexes. These studies examined the sliding dynamics of p53 along DNA, where the C-terminal slides along DNA whilst the core domain interacts frequently by association and dissociation. This two-state interaction allows p53 to both read the DNA sequence quickly as it moves along the strand and bind strongly, once it has found the binding sequence. However, protein labelling is not always possible or desirable, and it is therefore of interest to probe the p53-DNA interaction without any labelling or staining. To this end, Long et al. studied complexes of a p53 polypeptide and a 40 bp DNA and their
capture in the vestibule of the biological pore α-hemolysin. The authors show that complex formation can enhance the association rate constant with the pore, which was in turn used to study the complex itself (e.g. the interaction strength). However, α-hemolysin is too narrow to allow for translocation of the intact protein-dsDNA complex and in this context a greater flexibility, in terms of the pore dimensions would be desirable.

Here, we therefore report a comprehensive study of label-free, flp53 and its interaction with engineered double-stranded DNA in vitro, at the single-molecule level, using Atomic Force Microscopy (AFM) imaging and for the first time solid-state nanopore sensing. Briefly, nanopore sensors are composed of a thin membrane, usually consisting of silicon nitride (SiN) or other highly-insulating materials, which have a small, nanometre-sized aperture connecting two electrolyte-filled compartments. Upon application of a bias voltage across the pore, an ion current is established that depends on the pore dimensions and surface properties, the conductivity of the electrolyte solution and the magnitude of the applied bias. The nanopore typically represents the largest resistance in the system, and as such the majority of the voltage drops across the membrane. Due to the small membrane thickness (say, < 100 nm), this results in a large electric field across the pore, driving ions and other charged species, such as proteins and DNA, through the nanopore. This ‘translocation’ of analytes results in an ion current modulation, details of which can reflect the dimensions, charge and conformation of the translocating species.

We use four DNA samples, namely pET24a DNA with no consensus sequence as control, and three consensus DNA samples with specific fragments inserted into the pET24a DNA: ‘26DNA’ with a half-site consensus sequence, ‘36DNA’ with a full-site consensus sequence, and ‘47DNA’ with a full-site consensus sequence where the two half-sites are separated by an 11 bp gap (the numbers refer to the length of the fragment).
METHODS:

Preparation of linear cfPET dsDNA: Three sets of oligonucleotides with 26-bases, (5'-AATTCAATTGGGCAAGTCTAGGA-3’, 5'-AGCTTTCTTAGACTTGGCCCAATTATG-3’), 36-bases (5'-AATTCAATTGGGCAAGTCTGGAATTGGCAAGTCTTAGGA-3’, 5'-AGCTTTCTAGACTTGGCCCAATTATG-3’) and 47-bases (5’-AATTCAATTGGGCAAGTCTAGGAATTGGCAAGTCTTAGGA-3’, 5’-AGCTTTCTAGACTTGGCCCAATTATG-3’) were obtained from Sigma Aldrich (Dorset, UK) and hybridised to obtain double-stranded DNA consensus fragments. Circular pET-24a-d(+) double-stranded DNA (69752-3, Novagen) was digested with EcoRI and HindIII restriction enzymes, resulting in a 19 basepair DNA fragment and linear pET24a DNA. The digestion reaction contained 14 µL of hybridized DNA, 2 µL of 10X Buffer E (Promega, Southampton, U.K.), 2 µL of 10 mg mL⁻¹ Bovine serum albumin (BSA), 1 µL of 50 U µL⁻¹ EcoRI enzyme and 1 µL of 10 U µL⁻¹ HindIII enzyme. After incubating at 37 °C for 4 h, the mixture was purified using QIAquick PCR Purification kit (Qiagen) and the purified digested DNA was analyzed on 0.8% agarose gel. DNA concentrations were measured by UV Vis spectroscopy (NanoDrop ND2000m Labtech International Ltd., East Sussex, U.K.) and stored at -20 °C until use. The linear pET24a dsDNA was used with the three DNA consensus fragments in ligation experiments. The ligation reaction contained 1 µL T4 DNA ligase, 2 µL T4 ligase buffer (New England BioLabs Inc.), 1 µL 10 mM ATP, 5 µL pET24a dsDNA and 1 µL consensus dsDNA fragment to a total volume of 10 µL. The ligation reaction was incubated at 4 °C overnight, then transformed in XL10-Gold E.coli competent cells (Agilent Technologies) and cloned on LB agar plates (Merck) with kanamycin at 37 °C. Colonies were picked and grown overnight in LB Broth (Merck) and kanamycin at 37 °C. Centrifugation allowed for removal of excess liquid and the pelleted bacteria was lysed. The supernatant was purified using Spin Mini Prep
kit (Qiagen). Purified ligated DNA samples were eluted in 10 mM Tris-HCl, 1 mM EDTA, pH 8.0. Sequencing confirmed successful ligation and samples were named 26DNA, 36DNA and 47DNA corresponding to the length of inserted consensus fragments respectively. These three DNA samples and circular pET24a were digested with BamH1 restriction enzyme in a 20 µL total volume mixture that contained 15 µL of DNA, 2 µL of 10X Buffer SB (Promega, Southampton, U.K.), 2 µL of 10 mg mL⁻¹ Bovine serum albumin (BSA), 1 µL of 10 U µL⁻¹ BamH1 enzyme.

Binding Assay for flp53 and DNA: flp53 protein was obtained from Sigma-Aldrich (recombinant, expressed in baculovirus infected Sf21 cells). Incubation of flp53 with short DNA consensus fragments was performed in a DNA binding buffer of 100 mM KCl, 20 mM Tris-HCl, 5 mM MgCl₂ and 10% glycerol at pH 7.0 at 22 °C for 45 min, then at 4 °C for 30 min. Molar ratios of flp53 to DNA were 1.2:1 for 26 bp DNA, 1:1 for 36 bp DNA and 1.9:1 for 47 bp DNA. The mixtures were incubated at 22 °C for 45 min, then at 4 °C for 30 min. After incubation all samples were analysed on 5% native polyacrylamide gel in 1X TBE buffer, stained with 50 mL for 3X Gel Red (Biotium) for 30 minutes and destained with sterile water for 10 minutes. DNA bands were analysed using Gel Doc XR System (Bio-Rad, Hertfordshire, U.K.).

Further binding assays were performed with the longer dsDNA samples 26DNA, 36DNA, 47DNA and pET24a, where flp53 protein was incubated with DNA at a molar ratio of 80:1. These binding assays were incubated in 100 mM KCl, 30 mM Tris-HCl, 1 mM DTT, 0.2 mM EDTA, 20% glycerol at pH 8.5 for 45 min at 22 °C, then at 4 °C for 30 min. After incubation the samples were analysed with AFM or nanopore devices.

**AFM.** Samples were prepared for AFM in 10 mM HEPES (pH 7.6), 4 mM MgCl₂ and sterile, 0.2 µm filtered water was used to bring the mixture to a total volume of 10 µL. 5 µL of this mixture was applied to freshly cleaved mica (Agar Scientific) and incubated at room
temperature for 5 min. The mica was rinsed with 2 mL nuclease free water and dried under a flow of dry N₂. Images were obtained in tapping mode in air at 295 K from at least three independently prepared samples using an Agilent Technologies 5500 AFM/SPM microscope and commercial “Super Sharp Silicon” AFM probes (Windsor Scientific). Imaging parameters: 512 x 512, 1024 x 1024 or 2048 x 2048 pixels (0.3 - 1.0 lines/s) with scan areas of 3 x 3, 5 x 5 or 10 x 10 \( \mu \text{m}^2 \). Images were processed with a second order “flatten filter” (WSxM 5.0 Develop 6.5, Nanotec Electronica S.L.) and ImageJ.

**Nanopore Fabrication.** The nanopore devices used in this work were fabricated using protocols described elsewhere. Briefly, the devices consisted of 2 parts, a Pyrex®-based chip with a micrometer scale hole formed by HF etching, and a SiNx membrane. 100 nm-thick low-stress SiNx was initially deposited on double-side polished Si wafer (thickness 500 \( \mu \text{m} \)) by low pressure chemical vapor deposition (LPCVD). A SiNx window of 2 mm x 2 mm in size was fabricated using photolithography and KOH wet etching of the patterned Si. The SiNx membrane on the Si chip (10 mm x 10 mm) was then transferred onto a Pyrex® chip of 9 mm x 9 mm in size. Before nanopore drilling, the transferred membrane thickness was reduced down to \(~20\ \text{nm}\) by CF\(_4\) plasma dry etching. A transmission electron microscope (JEOL JEM-2010F) was used to drill a single nanopore on each Pyrex® chip, with nominal diameters from 15 nm~20 nm.

**Nanopore Preparation and Analysis.** Membrane devices with nominal pore diameters of 15 – 20 nm were assembled in a Teflon cell and sealed with PDMS rings. 800 \( \mu \text{L} \) of 1 M KCl, 10 mM Tris-HCl, pH 8.0 was added to each compartment of the cell. 0.25 mm diameter Ag/AgCl electrodes were connected to the cell and IV curves were measured between -0.5 V and +0.5 V with a Gamry Reference 600 potentiostat to determine the pore conductance \( G \) and hence estimate the effective pore diameter, according to eq (1) (assuming cylindrical pore geometry and membrane thickness \( L \gg \text{diameter d} \):
\[ G = \frac{\pi d_{\text{pore}}^3}{4 l_{\text{pore}}} \left( \mu_K + \mu_{\text{Cl}} \right) n_{K\text{Cl}} e + \mu_K \frac{\sigma_*}{d_{\text{pore}}} \]  

where \( \sigma_* \) is the surface charge density, \( \mu_K \) and \( \mu_{\text{Cl}} \) are the electrophoretic mobilities of K\(^+\) and Cl\(^-\) ions respectively, \( n_{K\text{Cl}} \) is the charge density of the electrolyte and \( e \) is the elementary charge.\textbf{Error! Bookmark not defined.}

The cell was connected to a low-noise current amplifier, either an Axopatch model 200B (Molecular Devices, Sunnyvale, USA) or custom-built instrument. The analyte, either bare DNA or DNA complexed with flp53, was added to the cell to a final concentration of 300 nM of dsDNA. Negative voltages (0.05 – 0.2 V) were applied to drive the molecules through the nanopores. With the custom-built amplifier, the ionic current is formally split into a slowly changing ("DC") channel, which shows the open-pore current, and an "AC" channel, which records fast current modulations such as the translocation events. The AC channel was filtered at 100 kHz and sampled at 250 kHz or 1 MHz. Ionic current blockades and translocation (or "dwell") times were further analysed with a custom Matlab code and Origin 9.0. Events were defined as current transients that exceed a 5\( \sigma \) threshold, based on the AC channel noise; the event start and end were taken to be the time points adjacent to each event, where the current passes the respective 1\( \sigma \) current threshold. This was done to find a balance between recording the largest possible part of each translocation event and being robust towards any local baseline fluctuations.

Three different nanopore devices with calculated diameters of 15, 16 and 12 nm were used for the translocation of 36DNA, 36DNA-flp53 complex and pET24a-flp53 complex.

RESULTS AND DISCUSSION

Gel Electrophoresis Confirms flp53 Binds Consensus DNA Fragments. Prior to performing AFM and translocation experiments, the binding of flp53 protein to DNA was
investigated. We initially used short DNA consensus fragments, 26 bp, 36 bp and 47 bp in length (see Methods section for sequences) to confirm binding to the consensus sequences of interest. The DNA fragments were incubated with flp53 protein in DNA binding buffer (see Methods). This was followed by gel electrophoresis using a 5% native polyacrylamide gel. Shifted bands of flp53-DNA complex were observed, as seen in Figure 1(c), confirming that flp53 protein indeed binds to the DNA consensus fragments under the conditions used here.

**AFM Study of flp53-DNA Binding.** We examined the flp53-DNA binding with longer, roughly 5.3 kbp DNA molecules, with and without the consensus sequences used for the gel electrophoresis experiments. The non-consensus DNA, pET24a DNA, and the three consensus DNAs with inserted specific fragments, 26DNA, 36DNA and 47DNA, were used to examine the specific or non-specific nature of flp53 binding to different DNA consensus sequences using AFM. Due to digestion of the DNA samples with BamH1 restriction enzyme, the inserted consensus fragments were positioned 6 bp from the end of the DNA, while the pET24a DNA lacked an inserted consensus fragment. Therefore, we expected that positional analysis of the flp53 binding to the DNA samples would provide insights into specific or non-specific binding.
**Figure 1.** flp53 protein-DNA complexes translocated through a nanopore (a) Crystal structure of *E.coli* multi-domain p53 polypeptide as a dimer (PDB: 3Q01) showing estimated dimensions (~6.6 nm x 9.0 nm x 4.0 nm). (b) Illustration of bare dsDNA (top) and flp53-bound dsDNA (bottom) at the nanopore. A suitable potential is used to drive the analyte through the solid state nanopores. (c) 5% native polyacrylamide gel showing shifted bands of free dsDNA and flp53-DNA complexes for consensus fragments 26bp, 36bp and 47bp DNAs. (d) Representative ionic current traces of 1s duration and typical translocation events for 36DNA, 36DNA-flp53 complex and pET24a-flp53 complex at 200 mV.
We first performed AFM studies on all DNA samples and the flp53 protein individually (see Figs. S1 and S2). The bare DNA lengths for pET24a, 26DNA, 36DNA, and 47DNA, respectively, were in excellent agreement with the expected lengths of between 1.798 µm and 1.815 µm, as shown in Table 1. The height of bare DNA of 0.9 ± 0.2 nm was significantly smaller (56%) than the diameter of the molecule as estimated from the crystal structure (2.0 ± 0.1 nm; RCSB protein data bank, 1BNA), which is presumably due to the effect of tip forces acting on the sample. The lateral diameter of the DNA was measured as 12.0 ± 3.2 nm, which is larger than the physical dimension of the DNA, due to tip/sample convolution. Based crystallographic diameter of 2.0 nm, we used eq 1 in the SI to reconstruct the AFM tip shape and estimated the effective tip radius to be 9.0 ± 0.7 nm (see SI for a more detailed discussion).

Similarly, we analyzed the height and two cross-sectional diameters of bare flp53 protein. It was found that two populations were present for both the height and diameter measurements (see Fig. S3). As the height measurements for the flp53 protein were obtained in similar imaging conditions as the DNA samples, it is likely that the flp53 protein would also exhibit a reduced measured height. Assuming the same distortion factor as determined from the DNA samples, the height of the flp53 protein was 6.6 ± 1.9 nm and 12.5 ± 4.1 nm for populations 1 and 2 respectively. The corresponding populations for the diameter measurements, corrected for tip effects, gave population 1 and 2, as shown in Table 2.

<table>
<thead>
<tr>
<th>DNA sample</th>
<th>pET24a</th>
<th>26DNA</th>
<th>36DNA</th>
<th>47DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected length of molecule (µm)</td>
<td>1.798</td>
<td>1.808</td>
<td>1.811</td>
<td>1.815</td>
</tr>
<tr>
<td>Measured length (µm)</td>
<td>1.71 ± 0.03</td>
<td>1.76 ± 0.09</td>
<td>1.79 ± 0.04</td>
<td>1.85 ± 0.03</td>
</tr>
</tbody>
</table>

a. Assuming a length of 0.34 nm per bp
We compared the dimensions obtained from our measurements with those in the literature (see Figure 1a.),\textsuperscript{25,26,27} and found that most likely population 1 corresponds to dimeric flp53 protein, whilst population 2 corresponds to tetrameric flp53 protein. Approximately 56 – 66\% of the flp53 protein was found to be dimeric at the concentration range used in these experiments of 5.4 – 28.7 nM (in dimers), which is again in agreement with previous work under comparable conditions.\textsuperscript{8,28}

Next, we examined the flp53-DNA interaction with AFM. The DNA samples were incubated with flp53 protein (see Methods) and imaged in tapping mode in air. Most species on the surface appear to be bare DNA, however the formation of flp53-DNA complexes is visible in the representative AFM images of the flp53-DNA incubated mixture in Figure 2. Interestingly, there appears to be on average one flp53 protein bound along the full length of the DNA molecules. DNA lengths were comparable to those for bare DNA, and both dimers and tetramers of flp53 were found to bind DNA (see Figure S4, Table S1). We calculated the percentage of DNA molecules with flp53 bound for each DNA sample. For the same molar ratios of DNA and flp53, 36DNA formed the highest percentage of flp53-DNA complexes (15.6\%), whilst 26DNA formed the lowest (2.5\%), cf. Table S2. 47DNA and pET24a DNA were found to have similar percentages of bound flp53 protein, namely 6.5\% and 5.8\%, respectively. This is in agreement with previous findings where the binding of p53CT (residues 94-360) bound to a full length consensus element with high affinity but little binding was observed to the half-site DNA. In our case, the 36DNA is the only full length consensus sequence without a gap between half sites.

\begin{table}[h]
\centering
\caption{Height and Diameter Calculated Values for flp53 Protein}
\begin{tabular}{lccc}
\hline
flp53 population & Height (nm) & Diameter 1 (nm) & Diameter 2 (nm) \\
\hline
1 & 6.6 ± 1.9 & 4.8 ± 0.8 & 2.6 ± 0.4 \\
2 & 12.5 ± 4.1 & 17.7 ± 3.7 & 9.7 ± 2.6 \\
\hline
\end{tabular}
\end{table}
Next, we measured the position of the flp53 protein bound along the DNA molecules for each DNA sample. Since the DNA was unlabelled, the direction of the strand on the surface is unknown. Therefore, the position of the bound flp53 protein along the DNA was

**Figure 2.** Representative AFM images of flp53-DNA complexes (a) pET24a-flp53, (b) 26DNA-flp53, (c) 36DNA-flp53, and (d) 47DNA-flp53 (Imaging conditions: 1024 points/line, except (a) which is 512 points/line; scan speed: ~0.3 Hz). Scale bar is 600 nm, Z scale bar shows 0 nm to 3 nm left to right. The black arrows indicate the locations of flp53 bound to DNA. Insets show individual flp53 proteins bound to DNA, the scale bar is 50 nm, the Z scale is the same as in the large images.
measured with respect to the closest DNA end. Figure 3 shows four position histograms for the position of flp53 along each DNA sample, where each bin represents roughly 130 bp or 1/40\textsuperscript{th} of the total length of the molecule.

**Figure 3.** Position histograms of flp53-DNA complexes along the DNA. The position of the bound flp53 protein along the DNA was measured with respect to the closest DNA end and given as a fraction of the total length. The relative frequency is the observed probability of flp53 binding in the given range. The number of flp53-DNA complexes measured for each sample are indicated in these plots. The position of flp53 binding on pET24a (blue) appears to be most prominent around 0.2-0.3 along the DNA from the end (but statistically insignificant, cf. main text), whereas the three consensus DNA sequences (26DNA, black; 36DNA, green; 47DNA, red) show a significant peak at the end of the DNA, most likely due to sequence-specific DNA binding.

It can be seen that in the case of pET24a-flp53 complexes, there are two local maxima at positions 0.24 and 0.29. On the other hand, for the three consensus DNA samples, the most probable position for flp53 binding to DNA is observed at the end of the molecules. We
apply the Grubb’s test to the data in Figure 3, with a variation of the test for testing for two simultaneous outliers in the case for pET24a. The latter two were however not found to be significant, even at the 10% confidence level (α = 0.1), thus suggesting that specific binding did not take place in these sequence regions (as expected). On the other hand, the probabilities observed for flp53 protein binding to the end of the consensus DNA samples were found to be significant (α=0.05, cut-off 2.557), with normal Z-scores for the maxima of 2.855, 3.063 and 3.174 for 36-, 26- and 47DNA, respectively. This would suggest that binding to these regions is indeed not random, as expected based on the sequence.

We further investigated these findings by simulating random distributions of the same form as the histograms in Figure 3 (of which examples can be seen in Figure S5). Specifically, we numerically simulated 10,000 normalized histograms for each sample using a custom-built MATLAB script (cf. SI), assuming a uniform probability distribution of protein binding along the DNA, i.e. that there is no preferential binding of the protein to any part of the DNA. We then determined the fraction of histograms, which show binding probabilities P higher than or equal to those observed experimentally, at the same position along the DNA. This provides an estimate of the probability that the experimentally observed histograms are in fact due to random fluctuations in the protein position, at a given sample size. Convergence of the simulations was checked by multiple runs of different sets of random numbers as well as plots of P vs. simulation number (typically up to 10,000, see above, occasionally up to 100,000). Using this procedure, we obtained P values of 0.33%, 0.15% and 0.12% for 26DNA, 36DNA and 47DNA respectively, indicating that the observed binding patterns are unlikely to be the result of random fluctuations, in accordance with the Grubbs test described above.

**Nanopore Studies of flp53-DNA Binding.** Translocation studies were performed with the bare 36DNA, as well as 36DNA and pET24a DNA incubated with flp53 protein,
respectively. The 36DNA sample was chosen for its high percentage of flp53 binding and comparatively high sequence specificity, according to AFM data, whereas the pET24a DNA was chosen as the non-specific control sample. Translocation events were detected for all three samples at negative applied potentials (see Figure 1d) such that the species move towards the positively biased compartment. This is in accordance with the expected negative molecular charge of dsDNA, and suggests that the translocation characteristics of the complex at low density of bound flp53 are similar to the corresponding dsDNA (see Figure S6).

![Graphs showing translocation events](image)

**Figure 4.** (a) Scatter plot of ΔG versus dwell time for 36DNA translocation events only (b) Scatter plot of ΔG versus dwell time for 36DNA (black squares), 36DNA-flp53 complex (magenta diamonds) and pET24a-flp53 complex (blue circles). (c) Histogram of ΔG for 36DNA (black), 36DNA-flp53 complex (magenta) and pET24a-flp53 complex (blue). Gaussian distributions were fitted to the data
to obtain most probable $\Delta G$ values. Both DNA-flp53 complexes exhibit a peak at less than 1.7 nS, whilst 36DNA translocation events exhibit values of greater than 1.7 nS. (d) Histograms of dwell time for 36DNA (black), 36DNA-flp53 complex (magenta), and pET24a-flp53 complex (blue). It is notable that 36DNA has a comparatively short translocation time and a narrow time distribution, whereas for the DNA-flp53 complexes dwell times are longer and more widely distributed (see main text for details).

A total number of 273, 304 and 154 events were recorded at 200 mV for 36DNA, 36DNA-flp53 and pET24a-flp53 experiments respectively. Figure 4a shows the conductance blockade versus the dwell time (cf. Methods section) data obtained from a sample containing only 36DNA, whereas Figure 4b contains the corresponding results for 36DNA and pET24a after incubation with p53 (magenta diamonds and blue circles, respectively). The data from panel (a) is included again, for ease of comparison (black squares). We formally separate the data into two groups or clusters, namely ‘Group 1’ with $\Delta G > 1.7 - 20$ nS and dwell times $< 2 \cdot 10^{-4}$ s), which is the only group observed in translocation experiment with bare DNA, and ‘Group 2’ with $\Delta G < 1.7$ nS and dwell times between $2 \cdot 10^{-5}$ to $1 \cdot 10^{-2}$ s. Note that the three samples were characterised with three different pore devices, which most likely explains subtle differences in the translocation data, despite the with very similar pore diameters, *vide infra*.

The translocation of the pure 36DNA sample (pore diameter $d = 15$ nm) yielded events formally associated with Group 1 and two sub-populations with most probable conductance blockade of $2.1 \pm 0.2$ nS and $3.1 \pm 0.6$ nS, and a most probable duration of $50 \pm 23$ µs. We tentatively assign those sub-populations to linear and partially folded DNA.$^{30}$ While the ratio of $\Delta G$ should be perhaps be closer to 2 in these circumstances, taking into account the experimental error, it is possible that the short duration of the translocation events
prevented the actual shape of the events to be fully resolved and hence the observed $\Delta G$ is a convolution of the linear and the folded part.

Similar observations were made for the translocation of pET-24a-flp53 (d = 12 nm), with $\Delta G = 1.4 \pm 0.8$ nS, and $3.2 \pm 0.1$ nS and a dwell time of $49 \pm 14$ µs (presumably linear and partially folded again), and third population with $\Delta G = 0.6 \pm 0.1$ nS and a dwell time of $351 \pm 171$ µs (‘Group 2’). The latter is absent in the pure DNA sample and we therefore assign it to the pET-24-flp53 complex. In the case of the 36DNA-flp53 sample (16 nm pore), we again observe free DNA ($\Delta G$ between 2 and 4 nS with a duration of $50 \pm 10$ µs, Group 1) as well as the complex ($\Delta G = 0.9 \pm 0.1$ nS with a duration $270 \pm 236$ µs, Group 2). The conductance blockade is slightly larger than for the pET24a-flp53 sample, which we rationalise based on the somewhat smaller effective pore diameter. Thus, the above results demonstrate that solid-state nanopore sensors can be used to differentiate free and flp53-bound DNA.

Furthermore, we calculated the number of ‘bare DNA’ events in each data set, relative to the total number of observed events, and found values of 100% for 36DNA, 9.9% for 36DNA-flp53 and 28% for pET24a-flp53 translocation events were of the DNA alone. This again confirms that p53 interacts more strongly with 36DNA than with pET24a DNA, in agreement with our AFM data (see Table S2 for individual values).

Unfortunately, for the present datasets we were unable to identify the protein itself as a I(t) feature during the translocation event (e.g. as a short pulse superimposed on the main DNA translocation event), at least beyond reasonable doubt compared to control experiments. This may be due to a lack of time resolution and the relatively large variation of observed protein positions along the DNA (see AFM results above).
Figure 5: The bias voltage dependence of the translocation time for bare DNA (squares, ‘36DNA’, as well as Group 1 pET24a-p53 and 36DNA-p53) and for DNA/protein complexes (triangles, Group 2). As described in the main text, the translocation time decreases exponentially with bias voltage for bare DNA. For the complex, no translocation events were observed for bias voltages (numerically) smaller than -150 mV. Devices became unstable for voltages below -200 mV.

Finally, we investigated the bias dependence of the translocation process, for both free DNA and respective complexes, Figure 5, to gain further insight into the mechanism of the translocation process. A $V_{\text{bias}}^{-1}$-dependence of $\tau$ is typically observed, if $\tau$ is governed by an interplay of electrophoretic and hydrodynamic forces, and potentially surface friction, whereas an inverse exponential dependence on $V_{\text{bias}}$ points towards an activated translocation process.\textsuperscript{32} Error! Bookmark not defined.\textsuperscript{35}

For all Group 1 species, we found that the latter scenario fits the data best, even though given the experimental error further study is desirable in order to clearly delineate the two scenarios in the present case, see Figure 5. For comparison, we plot the $\tau$ vs. $V_{\text{bias}}^{-1}$-dependence for 36DNA-flp53 (group 1) in Fig. S8 in the SI; deviations become apparent particularly at low voltages.
We thus fitted the relation $\tau = \tau_0 \exp(-V/V_0)$ to the dwell times of Group 1 (free DNA events only), where $\tau_0$ is the translocation time in the absence of an external electric field (e.g. diffusion only) and $V_0$ is a parameter that characterizes the voltage-dependence of $\tau$. We obtained values of $\tau_0 = 0.21 \pm 0.03$ ms and $V_0 = 134.0 \pm 18.1$ mV for bare 36DNA, $\tau_0 = 0.28 \pm 0.04$ ms and $V_0 = 112.5 \pm 11.5$ mV for the 36DNA sample incubated with flp53, and $\tau_0 = 0.22 \pm 0.03$ ms and $V_0 = 116.9 \pm 16.4$ mV for the pET24a-sample after incubation with flp53; the same within experimental error, as expected.

Translocation of the flp53/DNA complexes was only observed for $V_{bias} = -150$ mV and -200 mV, while the devices tended to become unstable for $V_{bias}$ larger than that. Interestingly, the translocation times in the latter case appear to be somewhat independent of $V_{bias}$, within experimental error, even though more comprehensive data on this aspect is clearly desirable. It may be that adsorption of the protein to the nanopore surface plays an important part during the translocation process, even while bound to DNA, as we have been to show for Single-Stranded DNA Binding (SSB) protein/ssDNA complexes. Protein adsorption during the translocation process may also offer an explanation as to why $\Delta G$ is in fact smaller for the complex, compared to the bare DNA. If indeed the protein adsorbs to the surface, then the enhanced friction, in combination with the electrophoretic pulling force, would tend to stretch the DNA during translocation, resulting smaller $\Delta G$ and longer translocation times.

**CONCLUSION**

We have presented a study of flp53 binding to different DNA motives at the single-molecule level, using AFM and solid-state nanopore sensing. Although AFM has been used before to examine the p53-DNA interaction, the AFM studies here provide the first detailed information on the binding efficiency and spatial distribution of flp53 along the DNA. We
found that a full consensus sequence (for 36DNA) resulted in a higher percentage of DNA-flp53 complexes than for half-sites, full consensus sequences with a gap between half sites, and non-specific DNA. We also measured individual and DNA-bound flp53 proteins to determine their dimensions, suggesting the presence of both dimeric and tetrameric flp53 protein. Furthermore, we demonstrated, for the first time, that solid-state nanopore sensors can successfully distinguish between bare DNA and DNA-flp53 complexes in a pre-incubated mixture of flp53 and DNA. The presence of bound flp53 protein appeared to have a significant effect on the magnitude of conductance blockade and the duration of translocation, presumably due to protein adsorption to the pore walls or the membrane surface. Nanopore translocation data further indicate that the 36DNA formed more flp53 complexes than the control pET24a DNA, in agreement with our AFM experiments. Hence, solid-state nanopore sensing can provide some detailed and unique insight into the interaction between proteins and DNA (here fp53/dsDNA), even though the method is not without challenges.

Acknowledgements

Supporting Information. AFM results and further analysis; simulation of “random” protein binding to DNA and exemplary results from the script; further nanopore translocation results and analysis; and example translocation events for various analytes used in this study. This material is available free of charge via the Internet at http://pubs.acs.org.
REFERENCES


TOC IMAGE