Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

Single-molecule manipulation quantification of site-specific DNA binding

Xiaodan Zhao[1], Shiwen Guo[2], Chen Lu[2], Jin Chen[2], Shimin Le[1], Hongxia Fu[3,4] and Jie Yan[1,2]

The execution of functions on DNA relies on complex interactions between DNA and proteins in a sequence and structure dependent manner. Accurate quantification of the affinity and kinetics of these interactions is critical for understanding the molecular mechanisms underlying their corresponding biological functions. The development of single-molecule manipulation technologies in the last two decades has made it possible to apply a mechanical constraint to a single DNA molecule and measure the end-to-end extension changes with nanometer resolution in realtime. While it has been shown that such technologies can be used to investigate binding of ligands, which can be proteins or other molecules, to DNA in a fluorescence-label free manner, a systematic review on such applications has been lacking. Here, we provide a review on some of recently developed methods for fluorescence-label free single-molecule quantification of site-specific DNA binding by ligands and demonstrate their wide scope of applications using several examples of binding of ligands to dsDNA and ssDNA binding sites.

Addresses
[1] Department of Physics, National University of Singapore, Singapore 117542, Singapore
[2] Mechanobiology Institute, National University of Singapore, Singapore 117411, Singapore
[3] Division of Hematology, Department of Medicine, University of Washington, Seattle, WA 98109, USA
[4] Institute for Stem Cell and Regenerative Medicine, University of Washington School of Medicine, Seattle, WA 98109, USA

Corresponding author: Yan, Jie (phyyj@nus.edu.sg)

Current Opinion in Chemical Biology 2019, **53**:xx–yy

This review comes from a themed issue on **Chemical biophysics**

Edited by **Yan Jie** and **Terence Strick**

https://doi.org/10.1016/j.cbpa.2019.08.006

Introduction

A variety of critical biological processes, such as transcription, DNA replication and DNA damage repair, depend on highly complex interactions between DNA and many proteins, which are often in a DNA sequence and DNA structure dependent manner [1–5]. A deep understanding of how proteins interact with their specific DNA binding sites is critical to gain a better understanding of their biological functions.

In traditional biochemical assays, site-specific binding of ligands is often investigated using a number of different biochemical technologies in bulk, such as electrophoretic mobility shift assays (EMSA) [6], footprinting [7], fluorescence resonance energy transfer (FRET), isothermal titration calorimetry (ITC), stop-flow fluorimetry, surface plasmonic resonance [8], protein binding microarrays (PBM) [9] and so on. Bulk measurements can only provide ensemble-average of the information of ligand-DNA interaction and they do not provide information on the conformation of the ligand-DNA complex. In addition, increasing evidences suggest that DNA is subject to tensile force in live cells [10,11•], while the information of force-dependent binding of proteins to DNA cannot be obtained from traditional bulk assays since they do not apply mechanical constraints to DNA.

Recent rapid development of single-molecule techniques has made it possible to detect molecular interactions at a single-molecule level with very high sensitivity. These techniques can be broadly divided into two groups. One is based on fluorescence detection of the vicinity between the ligand molecules and target binding molecules, which requires labeling the molecules with fluorescence tags and detecting using fluorescence microscopy. The fluorescence based single-molecule detection methods have been extensively developed, applied and reviewed elsewhere [12,13•]. The other is based on detection of the binding-induced change in the micro-mechanical properties of the DNA binding sites, which can be detected by applying a mechanical constraint to single target molecules using single-molecule manipulation technologies [14–16,17•] without fluorescence labeling. Such fluorescence-label free single-molecule detection has a number of advantages, such as providing information of the conformation of the ligand-DNA complexes, offering longer measurement time scale without photobleaching, eliminating the potential influence of the fluorescence label on the interaction, and being able to quantify interactions under mechanical constraints.

In single-DNA manipulation experiment, the DNA molecule is tethered between surface (e.g., a coverslip surface) at one end and a probe (e.g., a bead in optical

tweezers or magnetic tweezers experiments) at the other end [14–16,17•,18–21]. A mechanical constraint such as tensile force is applied to the DNA through the probe. The tension and the end-to-end extension of DNA along the direction of the mechanical constraint can be measured with resolutions in piconewton (pN) and nanometer (nm) ranges, respectively. The mechanical responses of DNA to an applied mechanical constraint sensitively depend on the micro-mechanical and structural properties of the tethered DNA such as the stiffness, size, and structure [22]. Since ligand-binding often alters the shape and the stiffness at the binding site, it in turn changes the tension and the end-to-end extension of the DNA under mechanical constraint. Thus, it is possible to detect ligand-binding to a single tethered DNA molecule based on the binding-induced changes in micro-mechanical and structural properties of the DNA binding sites.

In spite of many potential advantages of using single-molecule manipulation technologies to quantify ligand-binding to DNA in the fluorescence-label free manner, this approach has not been well exposed to many scientists in a wide range of research areas. Aiming to encourage more scientists to employ the power of single-molecule manipulation based fluorescence label-free qualifications of site-specific ligand-DNA interactions in their research, here we provide a review on the recent development of methods that enable convenient construction of specific DNA binding sites for single-molecule manipulation studies, and methods to quantify the affinity and kinetics between the ligands and their specific DNA binding sites.

In situ building specific DNA binding sites

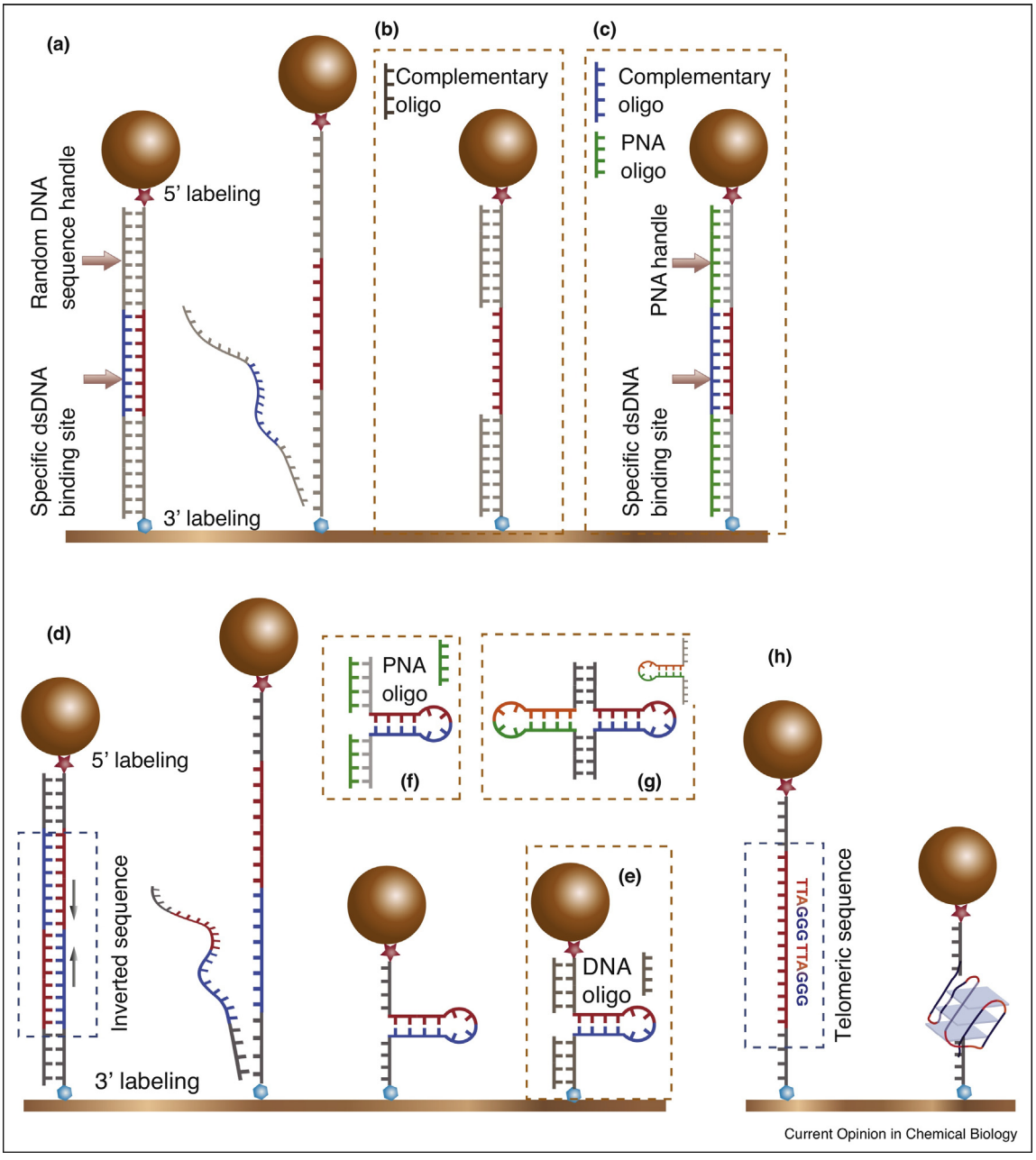
To study ligand-binding to a single or a few specific binding sites, it is necessary to make a single-molecule tether including these sites. Various methods have been developed to make DNA constructs specific to certain binding sites; however, many of the current methods involve tedious procedures to make simple constructs such as a ssDNA site [23–27], a gapped DNA that has a ssDNA region spanned between two dsDNA regions [25], or a double-stranded hairpin [28–31,32••,33••,34••,35••]. These procedures typically contain several steps of biochemical and enzymatic reactions.

Is it possible to build different DNA constructs that contain different binding sites from a convenient scaffolding DNA structure? An ssDNA tether is an ideal candidate because the exposed bases on the ssDNA provide numerous ways for construction based on the highly-specific Watson-Crick basepairing interaction [36]. Starting from a ssDNA scaffold, a gapped DNA can be made by flowing in oligos that can hybridise with the tethered ssDNA at two distal regions (Figure 1a,b). A specific dsDNA binding site of certain size and sequence can be formed by annealing complementary ssDNA oligos to the region of interest on the tethered ssDNA, and blocking the remaining ssDNA using complementary peptide nucleic acid (PNA) oligoes (Figure 1c). If the ssDNA contains two inverted repeats, a hairpin can form at forces below the threshold DNA unzipping force (Figure 1d). Hybridising the spanning ssDNA handles near the fork with complementary ssDNA oligos, the hairpin can be made to be spanned between two dsDNA handles (Figure 1e). The ssDNA handles can also be coated with complementary PNA oligos, leaving the DNA hairpin the only binding site (Figure 1f). A four-way junction can be produced by hybridising a small DNA hairpin to the forked DNA tether (Figure 1g). If the target region on the tethered ssDNA contains G-quadruplex forming sequences, G-quadruplex structures can also be formed (Figure 1h). Thus, most of biologically important nucleic acids structures can be built from a tethered ssDNA scaffold by simple sequence design and annealing procedure.

To build DNA binding sites using ssDNA as a scaffold, an efficient way of generating a ssDNA tether is required. One approach is to start with an original dsDNA tethered between surface and a probe at the two ends of the same strand, and then remove the untethered strand biochemically or using enzymatic reactions [23–27]. The untethered strand can also be removed biophysically by applying tensile force slightly above 65 pN (Figure 1a). At this force, a well-known DNA over-stretching transition takes place [23,37], leading to two distinct structural reorganisations of DNA. One is the strand-separation transition in solution conditions of low salt concentration, high temperature or high AT contents in the sequence, which destabilise the Watson–Crick basepairing interaction [38–46]. The other is the B-to-S transition in conditions (high salt concentration, low temperature or GC rich sequence) that increase the stability of the Watson–Crick basepairing interaction. The B-to-S transition leads to an elongated structure where the two strands are still associated with each other [40,42,43,47,44–46].

Thus, a scaffolding ssDNA tether can be made from the original dsDNA tether through the force-induced strand-separation transition in appropriate solution conditions. This method has been utilised to produce a single ssDNA tether from an original dsDNA tether [43,48,49,50•]. Owing to the convenience of forming dsDNA tethers and the instant force-induced strand-separation transition, this method provides high efficiency of forming ssDNA tethers. By designing the sequence on the tethered ssDNA strand, various specific DNA binding sites can be assembled by annealing interactions. Comparing with previous methods that remove the untethered strand biochemically or using exonucleases [23–27], this approach only involves mechanical force and highly-specific annealing reactions in all steps.

Figure 1



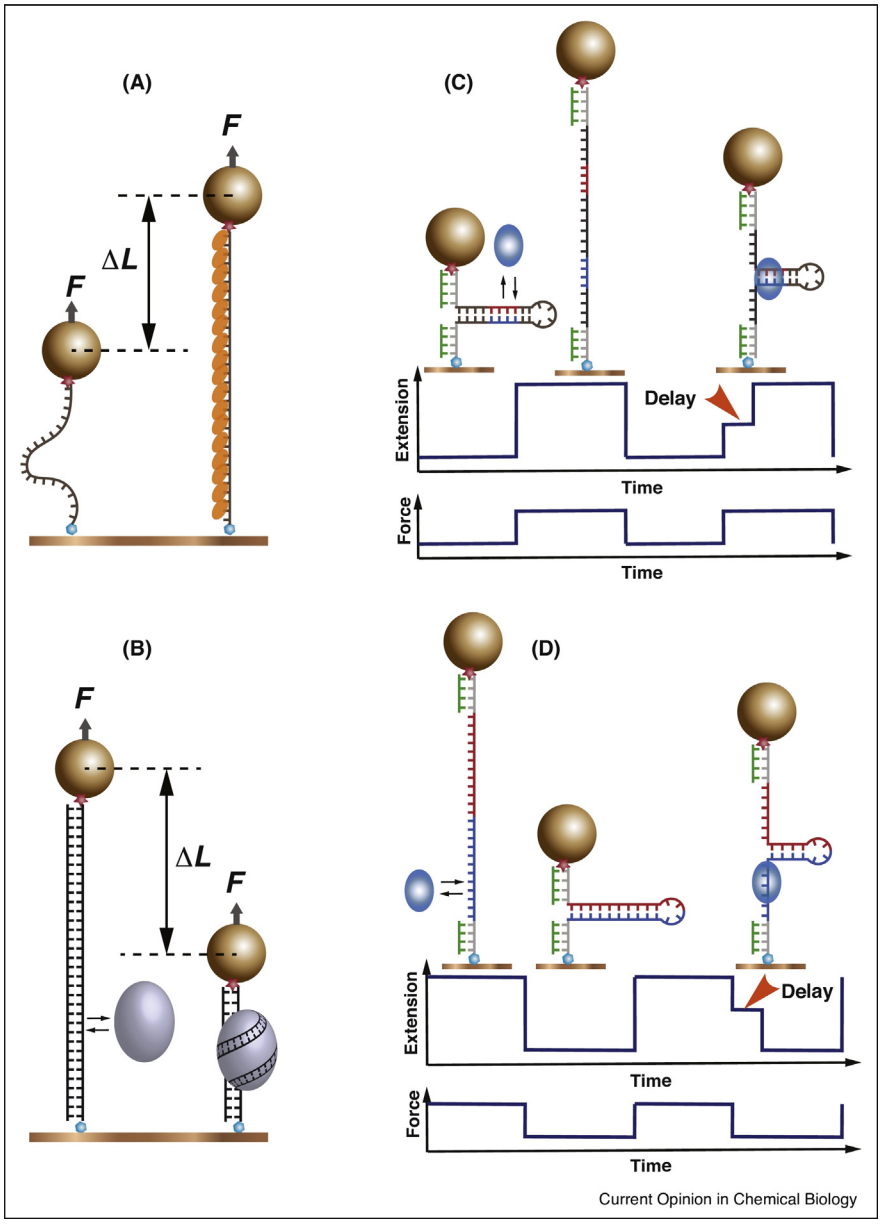
Schematics of building different DNA binding sites from a scaffolding ssDNA. **(a)** A scaffolding ssDNA tether is produced from an original dsDNA tether through force-induced strand-separation transition. The region marked in red indicates the target region to build binding sites. **(b)** A gapped DNA can be made by coating the two ssDNA regions flanking the target region with complementary ssDNA oligos. **(c)** A specific binding site for dsDNA-binding ligands is built by annealing a complementary ssDNA oligo to the target region and blocking the flanking ssDNA regions with complementary PNA oligos. **(d)** Incorporating inverted repeats indicated by blue and red colours on the tethered strand, a forked DNA flanked by two ssDNA handles is formed at forces below the unzipping force. **(e)** The forked DNA can be further processed into a three-way dsDNA junction by coating the ssDNA overhangs with complementary ssDNA oligos. **(f)** The ssDNA overhangs of the forked DNA can also be blocked with complementary PNA oligos, leaving the hairpin the only binding site for dsDNA-binding ligands. **(g)** A four-way junction can be formed by flowing in small hairpins with short overhangs that are complementary to the overhangs of the forked DNA. **(h)** Other structural motifs of DNA, such as G-quadruplex, can be made from the ssDNA by incorporating corresponding sequences into the tethered strand.

Principles of fluorescence-label free detection of site-specific ligand-DNA interactions

Detection based on binding site deformation — Ligand-binding can be probed based on the changes in the end-to-end extension as a result from binding-induced deformation of the binding site (Figure 2a,b). This approach is suitable for ligands that deform DNA significantly, such as proteins that result in DNA wrapping around the protein surface [51,52], sharp DNA bending [53,54], increasing the bending stiffness [55–57,49], mediating DNA loop formation [58,59], or DNA contour elongation by DNA intercalators [60,61]. In particular, this method has been extensively applied to study the dynamics of nucleosomes using DNA containing high affinity nucleosome positioning sequences such as the 601 sequence [62] or sea urchin 5S positioning sequence [63].

Detection based on delayed structural transitions — A bound ligand on dsDNA can be detected using a DNA hairpin

Figure 2



Binding detection principles. (a,b) Schematics of detection based on binding-induced extension changes for ssDNA-stiffening proteins that increase the extension of ssDNA **(a)** and for dsDNA-wrapping proteins that decrease the extension of dsDNA **(b)**). (c,d) Schematics of the detection of binding using a hairpin detector based on delayed unzipping after increasing force to an above-threshold value for dsDNA-binding ligands **(c)** and delayed rezipping after decreasing force to a below-threshold value for ssDNA-binding ligands **(d)**.

construct as a detector, as ligands bound on the dsDNA hairpin may resist to mechanical unzipping of the hairpin. Mechanical DNA unzipping can be realised by gradually increasing the distance between the 3′ and 5′ termini of the dsDNA at the same end using instruments such as optical tweezers and AFM. This results in splitting the two strands of the dsDNA by sequentially breaking the base pairs at a force range (10-20 pN) depending on DNA sequence and ionic strength [64,65]. When the progressing fork reaches a site bound with a ligand, due to the binding-induced resistance to unzipping of the hairpin, tension increases until the ligand is displaced. Thus, the abrupt increased tension indicates the binding of a ligand and the fork-paused position specifies the location of the bound ligand on the DNA [28,29]. A bound ligand can also be detected based on delayed unzipping of the DNA hairpin after jumping from a lower force at which the hairpin is stable to a force slightly above the threshold unzipping force (Figure 2c). A stably bound ligand can cause a detectable delay in unzipping after the force jump. This delay can serve as a readout to determine whether the hairpin is bound with a ligand or not before the force jump. Such delayed hairpin unzipping has been recently applied to detect sequence-specific dsDNA binding ligands [30,31,32••,33••,35••]. As the hairpin is excluded from the force-transmission pathway, this method can be applied to ligands that do not introduce significant DNA deformation at the binding site.

Binding to a ssDNA binding site can also be quantified using a hairpin as a detector. In this case, the hairpin is held at a force above a threshold to expose the ssDNA binding site for ligand-binding. When force is jumped to a value slightly below the threshold, a bound ligand may cause detectable delay during hairpin rezipping (Figure 2d). On the basis of such delay, one can determine whether the ssDNA binding site is bound with a ligand or not before the force jump. In principle, binding to other nucleic acids structures such as three-way junctions, four-way junctions, and G-quadruplexes could also be detected based on such binding-induced delay of force-dependent unfolding transitions of these structures.

Detect ligand-binding based on binding site deformation

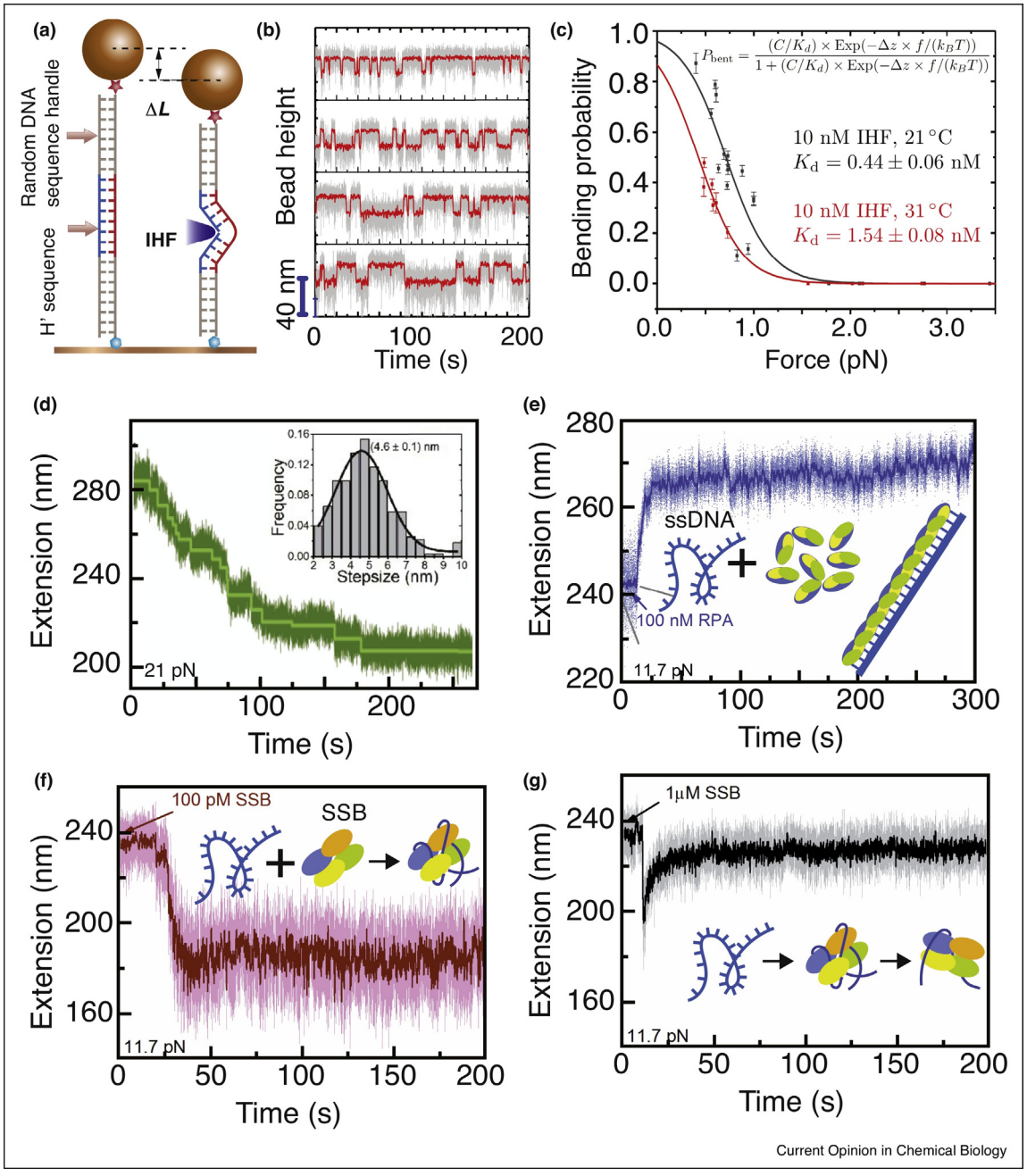
Detect ligand-binding to dsDNA — As aforementioned, ligand-binding to a DNA site can be detected based on the binding induced deformation of DNA under mechanical force. This method was previously applied by our group to quantify the binding of IHF to a single specific site, the H′-sequence [66], inserted into a 445 bp dsDNA tether under constant forces [54] (Figure 3a–c). This example reveals that IHF binding-induced bending of the H′-sequence results in a detectable extension change at sub-pN forces (Figure 3b). The stepwise change between two distinct extensions indicates that the binding and unbinding can be described with a simple two-state model. Recording a time trace of the two-state fluctuation at a given force, the time fraction of the bound state can be interpreted as the binding probability at the force.

Since the DNA is subject to a force F, the measured binding probability $p_{on}(F)$ depends on the applied force. Indeed, the experimental data reveal that a tiny increase of force from 0.5 to 0.9 pN results in significantly decreased binding probability (Figure 3b,c). Such sensitive dependence of the binding probability on force on one hand indicates that tension in the DNA may play a critical role in regulating ligand-binding to DNA, while on the other hand it results in a difficulty to compare the quantified ligand-DNA interactions with results obtained in traditional assays where DNA is not subject to forces. For simple binding and unbinding that can be modelled by two-state transitions, it can be shown that $p_{on}(F) = \frac{c/K_d \times e^{-(\Phi_{on}(F)-\Phi_{off}(F))/k_BT}}{1+c/K_d \times e^{-(\Phi_{on}(F)-\Phi_{off}(F))/k_BT}}$, where K_d is the dissociation constant of the interaction, and $\Phi_{on}(F)$ and $\Phi_{off}(F)$ are force-induced conformational free energy that can be calculated based on the force-extension curves of the tether that bears the binding site in the bound and unbound states [67,68]. Fitting the experimental data of the binding probability with this model, one could determine the value of K_d of the interaction (Figure 3c). Using this approach, the estimated values of K_d for the IHF-H′ interaction at different temperatures [54] are consistent with those determined in previous bulk assays [66,69].

Detect ligand-binding to ssDNA — Ligand-binding to a ssDNA site can also be detected based on the binding-induced deformation of the ssDNA. Being a highly flexible polymer, any binding may result in increased bending stiffness at the binding site that tends to increase the extension of ssDNA. If the ligand-binding involves bending, wrapping or forming a local structure with a shorter contour length, it will tend to decrease the ssDNA extension. These factors may compete with each other that collectively determine the observed extension change of the ssDNA under a mechanical constraint. For example, the activity of DNA polymerisation results in progressive conversion from ssDNA to dsDNA. Utilising the drastic difference in the force-extension response between ssDNA and dsDNA, the dynamic actions of DNA polymerase were directly monitored based on the dynamic extension change of a tethered ssDNA [26]. Similar approach was also used to detect the dynamic polymerisation of RecA on ssDNA, binding of ssDNA-binding proteins and homologous interaction between ssDNA and dsDNA [49,50•,57,70].

Here we provide additional examples to demonstrate the detection of binding of ssDNA-binding ligands based on binding-induced conformational change of ssDNA. In the first example, we probed the binding of ssDNA oligos, each with a size of 30 nt, to a mechanically stretched

Figure 3



Detect ligand-binding to dsDNA and ssDNA based on extension change. **(a)** Schematic of the IHF unbound (longer extension) and bound (shorter extension) states of a single H′-sequence inserted into a 445 bp dsDNA tether. **((b)** IHF binding to and unbinding from H′ site result in dynamic fluctuations between two distinct extension levels at constant forces varied from 0.9, 0.76, 0.6 to 0.5 pN corresponding from top to bottom panels, in solution containing 10 nM IHF (50 mM KCl, 2.5 mM $MgCl_2$, 10 mM Tris (pH 7.4) at 21°C). **(c)** The binding probabilities at different forces at the temperatures of 21°C (dark gray symbols) and 31°C (red symbols), respectively, fitted with a two-state model of force-dependent binding probability. **(d)** At a constant force of ~21 pN, annealing of 19 pieces of 30 nt ssDNA oligos (5.3 nM each) to a tethered 572 nt ssDNA resulted in stepwise extension decreases corresponding to individual annealing events. Inset shows the histogram of the step sizes. **(e)** At a constant force of ~12 pN, introducing 100 nM RPA results in gradual extension increase of a 572 nt ssDNA. **(f)** At a constant force of ~12 pN, introducing 100 pM *Escherichia coli* SSB results in rapid extension decrease of a 572 nt ssDNA. **(g)** At a constant force of ~12 pN, introducing 1 μM SSB results in initial rapid extension decrease of a 572 nt ssDNA, followed by partial recovery of the extension indicating a kinetic switch of the binding modes.

572 nt ssDNA. The 572 nt ssDNA contains 19 specific non-overlapping annealing sites for the complementary oligos. Figure 3d shows the extension change of the 572 nt ssDNA tether at 21 ± 2.0 pN after the introduction of 19 different oligos, each at a concentration of 5.3 nM. The extension of the ssDNA decreases in a stepwise manner after the introduction of the oligos. Each step has a stepwise extension decrease of ~4.6 nm (Figure 3a inset), consistent with the extension difference between 30-nt ssDNA and 30-bp dsDNA at the force of 21 ± 2.0 pN based on the differential force-extension curves of ssDNA and dsDNA [71,72], indicating that each step corresponds to a single annealing event. While the data in this demonstration were obtained at ~21 pN, the experiment can be done at other forces as long as there is a significant extension difference between the ssDNA and dsDNA.

In the second example, we flowed in 100 nM of full-length human replication protein A (RPA) at a force of 12 ± 1.5 pN applied to the ssDNA. After RPA was introduced, gradual elongation of ssDNA extension was observed (Figure 3e), indicating an ssDNA-stiffening effect of RPA. In contrast, introduction of 100 pM *E. coli* single-stranded DNA binding (SSB) protein to the 572 nt ssDNA resulted in significant shortening of the ssDNA at the same force (Figure 3f), while introduction of 1 μM SSB led to an initial rapid extension decreasing phase followed by a second stage during which the extension is partially recovered (Figure 3g). Previous structural studies have revealed that *E. coli* SSB forms a tetramer as the basic ssDNA binding unit, and upon binding it can cause ssDNA wrapping around the tetramer at two different levels depending on the concentrations of SSB. Overall, besides of being consistent with results from previous structural studies [73], this example demonstrates that it is possible to provide the dynamic aspects of the interaction and mode of binding by detecting the binding induced conformational change of ssDNA.

Overall, the examples described in this section demonstrate that one can detect ligands binding to dsDNA and ssDNA binding sites based on binding-induced deformation of the binding site in a fluorescence-label free manner. Such measurements can inform us about important information on the conformation of the nucleoprotein complexes. With the help of theoretical models, it is possible to compare results obtained at different mechanical constraints.

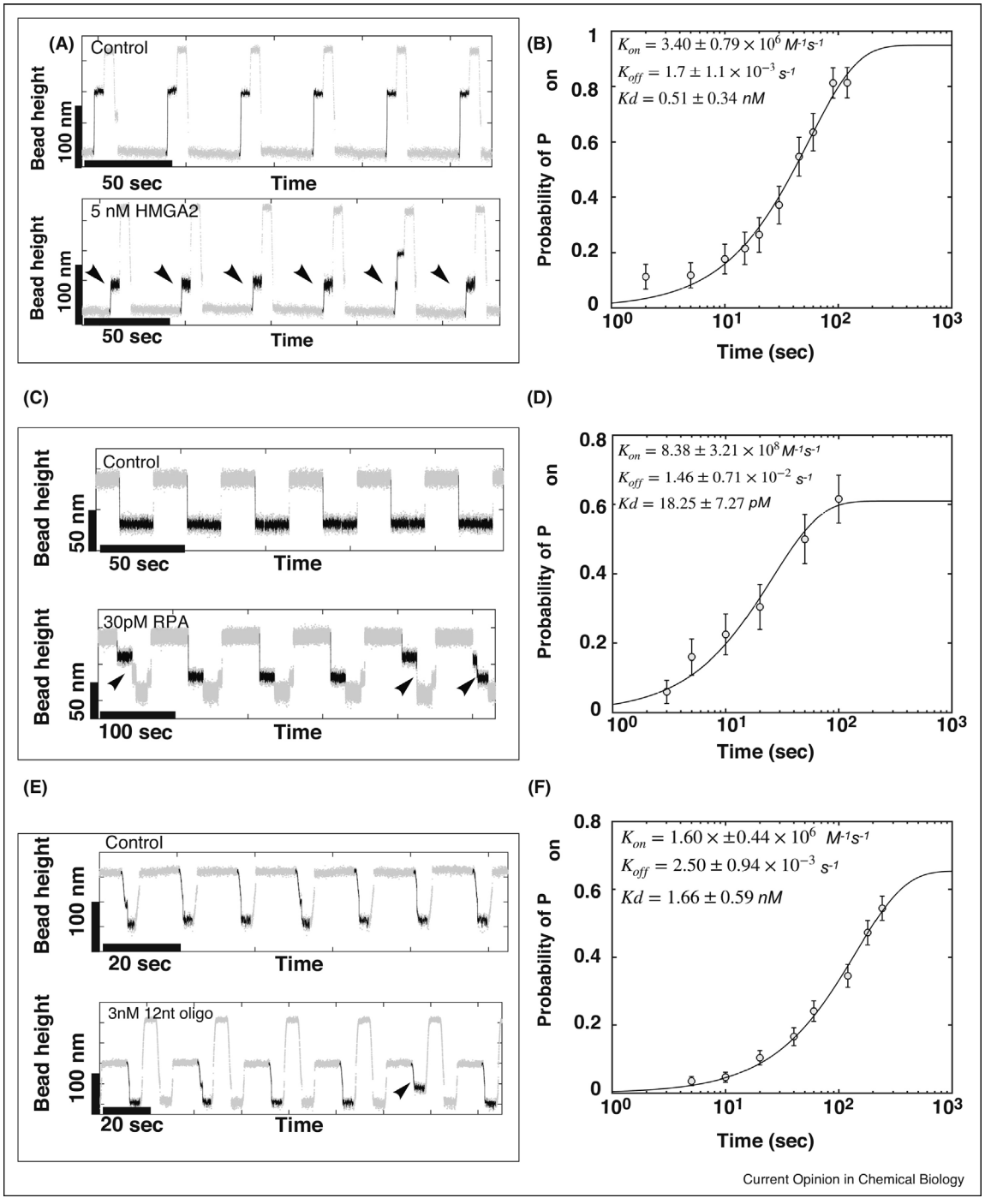
Detect ligand-binding based on delayed structural transitions

Detect binding to dsDNA based on delayed DNA unzipping — For ligands that do not cause detectable extension change to DNA, the ligand-DNA interaction can still be probed with the help of using a short DNA hairpin that contains one or multiple sites for ligand-binding. As illustrated in Figure 2c, the binding of dsDNA-binding ligands can be detected based on delayed unzipping of the DNA hairpin, which has been demosntrated in several previous studies [28,30,31,32••,33••,35••]. Here we demonstrate how to further improve this approach to enable quantification of the sequence-specific binding affinity and binding kinetics of ligands, using HMGA2 binding to a 52-bp long hairpin that contains a specific sequence (5′-ATATTCGCGAWWATT-3′) [74] for an example.

The binding of HMGA2 was investigated in a force cycle procedure. In each cycle, the force applied to DNA was sequentially changed between three levels: a protein-binding force (7.7 ± 0.8) pN at which the hairpin is stable for 30 s to allow HMGA2 binding, a binding-detection force (12.1 ± 1.2 pN) for 5 s to detect whether the DNA hairpin is bound with HMGA2, and a protein-displacing force (30.0 ± 3.0 pN) for 5 s causing immediate unzipping regardless whether the hairpin is bound with protein, to remove any bound HMGA2. At the binding-detection force, the hairpin unfolds within 0.3 s without HMGA2. In the presence of HMGA2, after force-jump from the protein-binding force to the binding-detection force, if the dwell time (i.e., the time taken to unfold) is longer than one second, we interpret that the hairpin is bound with at least one HMGA2 before the force jump. After each cycle, the protein was displaced at protein-displacing force, ensuring that at the beginning of the next cycle the DNA hairpin is unbound with any protein. Repeating the cycles for a large total number of cycles obtained from multiple tethers, the binding probability was calculated as the number of cycles when the binding was observed divided by the total number of cycles. Representative time traces during such force-cycle procedures for a hairpin in the absence and presence of 5 nM HMGA2 obtained from the same DNA hairpin tether are shown in Figure 4a. Indeed, delays in the unzipping of the hairpin (dwell time >1 s) were observed.

The duration Δt when the DNA hairpin is exposed for binding at the protein-binding force can be controlled in experiments, and the binding probability $p(\Delta t)$ over this duration was quantified. At the limit of $\Delta t \to 0$, no binding can occur, thus $p(\Delta t) \to 0$. In contrast, at the limit of $\Delta t \to \infty$, many binding and unbinding events may take place, $p(\Delta t) \to p_{eq}$ where p_{eq} is the equilibrium binding probability. Therefore, measuring the $p(\Delta t)$ at different durations will provide important information on the binding and unbinding kinetics of the protein. More specifically, modelling protein binding and unbinding as a two-state transition, it can be shown that $p(\Delta t) = p_{eq}(1 - e^{-r\Delta t})$. Here r is the rate of relaxation, which depends on the association rate ck_{on} and the dissociation rate k_{off} through a simple relation $r = ck_{on} + k_{off}$, where $c = 5$ nM is the concentration of HMGA2. Furthermore, the equilibrium binding

Figure 4



Quantify ligand-binding to dsDNA and ssDNA based on delayed hairpin unzipping and rezipping. **(a)** HMGA2 binding to dsDNA specific sequence (5′-ATATTCGCGAWWATT-3′) inserted to the near-middle region of a 52-bp hairpin spanned between two PNA-coated 200 nt ssDNA handles. Representative time traces show the height change of a bead attached to the DNA construct during cycles of sequential force jumps among 7.7 ± 0.8, 12.1 ± 1.2 and 30.0 ± 3.0 pN, in the absence (top panel) and presence of 5 nM HMGA2 (bottom panel). Arrows indicate delayed unzipping. **(b)** $\alpha(\Delta t)$ obtained from $N = 50$ force cycles of HMGA2 binding at each Δt on three independent tethers. **(c)** Representative time traces of the height change of a bead attached to the 30-bp hairpin, during cycles of sequential force jumps among 18.7 ± 1.9 and 7.6 ± 0.8 pN in the absence (top panel) and among 18.7 ± 1.9, 7.6 ± 0.8, 3.4 ± 0.3 and 7.6 ± 0.8 pN in the presence of 30 pM RPA (bottom panel). The lowest force of 3.4 ± 0.3 pN is used to facilitate removal of the bound RPA. **(d)** $\alpha(\Delta t)$ calculated from $N = 50$ force cycles at each Δt on three independent tethers. **(e)** Representative time traces of the height change of a bead attached to DNA construct containing a 52-bp hairpin, during cycles of

probability is related to the rates by equation $p_{eq} = \frac{ck_{on}}{ck_{on}+k_{off}}$. Solving the two equations based on the best-fitting values of r and p_{eq}, one can obtain estimations for the values of k_{on}, k_{off} and $K_d = k_{off}/k_{on}$.

In our experiments, at each Δt, we estimated the sample proportion $\alpha = m/N$, where N is the number of force-jump cycles and m is number of detected binding in these force-jump cycles. Repeating this procedure on three independent tethers at different values of Δt, we obtained the average values of the sample proportion, $\overline{\alpha}$ (i.e., the data points in Figure 4b). The number of detected binding (m) among the number of cycles ($N = 50$) follows binomial distribution, $P_N^m = \frac{N!}{(N-m)!m!}p^m(1-p)^{N-m}$, where p is the probability of binding occurring to the hairpin during Δt. Approximating the value of p with the best estimate of $\overline{\alpha}$, at each Δt, we regenerated 100 replica of the force-jump cycles based on the binomial distribution. In each replica, a new estimated value of $\alpha(\Delta t) = m/N$ was calculated at each Δt. Fitting $\alpha(\Delta t)$ with $\alpha(\Delta t) = p_{eq}(1 - e^{-r\Delta t})$ with a condition $p_{eq} \leq 1$, each replica generated estimations for the values of k_{on}, k_{off} and $K_d = k_{off}/k_{on}$. The mean values and the standard deviations of these quantities were then obtained from these 100 replica as $k_{on} = 3.40 \pm 0.79 \times 10^6\ \mathrm{M}^{-1}\ \mathrm{s}^{-1}$, $k_{off} = 1.7 \pm 1.1 \times 10^{-3}\ \mathrm{s}^{-1}$, and $K_d = 0.51 \pm 0.34$ nM (Figure 4b). The measured value of K_d is comparable with previously reported ones measured using traditional biochemical methods [75].

Detect binding to ssDNA based on delayed DNA rezipping — A hairpin can also be used to quantify the kinetics and affinity of site-specific binding of ligands to ssDNA. At a force below the threshold force of a hairpin at which rapid complete rezipping takes place, a bound ssDNA-binding ligand may cause detectable delay in the rezipping. The probability of ligands binding can be quantified in a reverse force cycle procedure used for dsDNA binding. In each cycle, the force applied to DNA hairpin was also sequentially changed among three levels: a factor-binding force above the threshold at which the hairpin is unzipped, a binding-detection force slightly below the threshold, and a ligand-displacing force which is well below the threshold for displacement of the bound ligands. We demonstrate this approach by investigating the binding of RPA to a mechanically unzipped 30-bp hairpin and annealing of 12-nt ssDNA oligos to the complementary region in a mechanically unzipped 52-bp DNA hairpin.

The tether was held at the protein-binding force for certain time interval Δt for RPA or oligos binding to the mechanically exposed ssDNA substrate. Repeating the force cycles many times at a given concentration of the ligand c, one can obtain the binding probability $p(\Delta t)$ based on the delayed rezipping (Figure 4c and e). Fitting the experimental data with the two-state model described in the above section, we have quantified the binding kinetics and affinity for full-length RPA to be $k_{on} = 8.38 \pm 3.21 \times 10^8\ \mathrm{M}^{-1}\ \mathrm{s}^{-1}$, $k_{off} = 1.46 \pm 0.71 \times 10^{-2}\ \mathrm{s}^{-1}$, and $K_d = 18.25 \pm 7.27$ pM (Figure 4d). Similar experiment performed for the annealing of a 12 nt ssDNA oligo to a complementary region in the 52 bp hairpin determined the association rate, dissociation rate and the dissociation constant to be $k_{on} = 1.6 \pm 0.44 \times 10^6\ \mathrm{M}^{-1}\ \mathrm{s}^{-1}$, $k_{off} = 2.5 \pm 0.94 \times 10^{-3}\ \mathrm{s}^{-1}$, and $K_d = 1.66 \pm 0.59$ nM (Figure 4f).

The results show that ligand-binding to dsDNA and ssDNA can be quantified in a fluorescence label-free manner based on binding-induced delay in unzipping for dsDNA-binding ligands and rezipping for ssDNA-binding ligands. For dsDNA binding ligands, because the hairpin is excluded from the force transmission pathway, the quantified interaction is not affected by the force applied to the tether. In contrast, since the mechanically exposed ssDNA substrate is under forces typically greater than 10 pN, the results are likely influenced by the tension in the ssDNA. Thus, theoretical models of force-dependent ligand-ssDNA interactions will be needed to compare the results with those obtained in traditional assays where ssDNA is not placed under a mechanical constraint.

The approach to quantify the binding kinetics and affinity demonstrated in the above examples can be extended to studies of ligand-binding to many other DNA structures such as G-quadruplex and four-way junctions based on binding induced delay in the transition of the respective structures. In a recent study from our group, we applied this approach to probe the binding of the helicase RHAU to a G-quadruplex structure based on the binding-induced stabilisation of the G-quadruplex [76].

Future perspective

In this review, we have provided an overview on the recent development of fluorescence-label free detection and quantification of site-specific ligand-DNA interactions based on mechanical manipulation of single DNA tethers that contain the binding sites. We have introduced a highly efficient method to build various DNA binding sites on a scaffolding ssDNA that can be conveniently produced using force-induced strand-separation of an original dsDNA tether. Two different methods of detection, one based on binding-induced extension change of DNA under a mechanical constraint and the

(Figure 4 Legend Continued) sequential force jumps among 18.0 ± 1.8 and 10.0 ± 1.0 pN in the absence (top panel) and among 18.0 ± 1.8, 10.0 ± 1.0, 50.0 ± 5.0 and 10.0 ± 1.0 pN in the presence of 12-nt complementary oligos (bottom panel). The highest force of 50.0 ± 5.0 pN ensures rapid removal of the bound complementary oligos. **(d)** $\alpha(\Delta t)$ obtained from $N = 200$ force cycles on three independent tethers. In panels (b), (d) and (f), the error bar is the standard error of the binomial distribution based on the $\sigma_{error} = \sqrt{\frac{p(1-p)}{n}}$.

other based on the binding-induced delay in the structural transition of the binding site, are explained with details and supplemented with demonstrating examples.

The detection based on binding-induced extension change of DNA are applicable to ligands that can cause significant deformation or significant change of the flexibility at the binding sites. Many proteins such as histones in eukaryotic cells and nucleoid associated proteins in prokaryotic cells that play a critical role in chromosome packaging as well as many transcription factors can cause a variety of large deformations such as wrapping, bending, looping, stiffening and contour elongating of dsDNA [77,78]. Similarly, ssDNA binding proteins can also cause wrapping, bending or stiffening at the binding sites [27,25,57,70]. Therefore, this detection method can be applied to a very large class of proteins.

As the detection requires applying a tensile force to a DNA tether that contains the binding site, the quantified interaction is dependent on the force applied to the DNA tether if the binding site is within the force-transmission pathway. As demonstrated by the example of IHF-DNA interaction, it is possible to develop theoretical models on force-dependent ligand-DNA interactions [79]. Such models on one hand allow comparison of the results obtained under different mechanical constraints, on the other hand can provide insights into how mechanical constraints may play a biological role in regulation of ligand-DNA interactions. The latter point is important since mounting evidences have suggested that, due to the existence of actomyosin contractility, cytoskeleton-nuclear linkage and the nuclear attachment of chromosome, many regions of chromosomal DNA are likely subject to tensile forces [11•]. Thus, the chromosomal packaging and transcription activities in such regions may be regulated by force-dependent ligand-DNA interactions.

The detection based on binding-induced delay in structural transition of the binding site is in principle applicable to any ligands as long as the binding can cause a delay that is long enough to be detected by the instruments. As demonstrated using the delayed DNA unzipping to quantify dsDNA-binding proteins and the delayed DNA hairpin rezipping to quantify ssDNA-binding proteins using very simple force-jump procedures, this approach can be used to quantify not only the affinity but also the kinetics of the interactions. The simple procedure as well as the directness of the readout make it promising to further develop this approach for automated screening of DNA binding drugs or compounds such as cisplatin, doxorubicin, 5-fluorouracil, etoposide, and gemcitabine [80] and drugs that influence site-specific binding of target proteins [81,82]. In addition, based on the molecular extension at the delayed unzipping or rezipping of a DNA hairpin, it is capable of inferring position of the bound ligand with nanometer (a few basepair) resolution [28,29,35••,83••]. Thus, this method is also promising to be used as a footprinting technology. Besides of being used as a molecular detector for single-molecule quantification of binding kinetics and affinity, DNA hairpins have also been extensively utilised to probe DNA translocation activity by helicases, polymerases and nucleosome remodelers [84–86].

The fluorescence-label free detection discussed in this review is based on extension change of the DNA tether under external force. Since protein binding to a DNA site often causes DNA twist, in principle, fluorescence-label free detection could also be based on probing the rotation of the rotationally unconstrained DNA using free-orbiting magnetic tweezers [87,88]. For rotationally constrained DNA, the detection can be based on probing the change in the number of supercoils using magnetic tweezers [89], or the change in torque in the DNA using torque-wrench magnetic tweezers or torque-wrench optical tweezers [90,91].

We argue that the general concept proposed in this study for quantification of site-specific ligand-DNA interaction can be further extended to study site-specific protein-protein interactions. A ligand bound on a protein domain may alter the stability and the rate of mechanical unfolding of the protein domain, which could be detected based on binding-induced change in the unfolding force or based on detection of delayed unfolding after force-jumping to an above-threshold force value [92–94]. Likewise, ligands binding to an unfolded protein conformation could be detected based on detection of delayed refolding after force-jumping to a below-threshold force value [95]. Finally, linking two interacting protein domains that can form a complex using a flexible peptide linker, it is possible to detect and quantify ligand-binding to either of the domains based on binding-induced delay in complex formation. Together, the fluorescene label-free single-molecule binding assay can potentially be applied to a very wide scope of biomolecular interactions from nucleic acids to proteins.

Authors' contributions

XZ and JY wrote the review. XZ, CL, SG, JC, SL, and HF provided and analysed the demonstrating experimental data.

Conflict of interest statement

None declared.

Acknowledgements

The authors acknowledge funding support by the Ministry of Education Academic Research Fund Tier 1; the National Research Foundation (NRF), Prime Minister's Office, Singapore under its NRF Investigatorship Programme (NRF Investigatorship Award No. NRF-NRFI2016-03); and the National Research Foundation, Prime Minister's Office, Singapore and the Ministry of Education under the Research Centres of Excellence programme.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Falkenberg M, Larsson N-G, Gustafsson CM: *Annu Rev Biochem* 2007, **76**:679.
2. Jackson SP, Bartek J: *Nature* 2009, **461**:1071EP.
3. Hall RM, Collis CM: *Mol Microbiol* 1995, **15**:593.
4. Dunham MA, Neumann AA, Fasching CL, Reddel RR: *Nat Genet* 2000, **26**:447EP.
5. Tse WC, Boger DL: *Chem Biol* 2004, **11**:1607.
6. Fried M, Crothers DM: *Nucleic Acids Res* 1981, **9**:6505.
7. Galas DJ, Schmitz A: *Nucleic Acids Res* 1978, **5**:3157.
8. Fägerstam LG, Frostell-Karlsson Å, Karlsson R, Persson B, Rönnberg I: *J Chromatogr A* 1992, **597**:397.
9. Berger MF, Philippakis AA, Qureshi AM, He FS, Estep PW III, Bulyk ML: *Nat Biotechnol* 2006, **24**:1429.
10. Iskratsch T, Wolfenson H, Sheetz MP: *Nat Rev Mol Cell Biol* 2014, **15**:825.
11. Shivashankar G: *Curr Opin Cell Biol* 2019, **56**:115.
• This is an extensive review on mechanical regulation of genome organisation and mechanical constraints applied to DNA in living cells.
12. Joo C, Balci H, Ishitsuka Y, Buranachai C, Ha T: *Annu Rev Biochem* 2008, **77**:51.
13. Shabestari MH, Meijering A, Roos W, Wuite G, Peterman E:
• *Methods in Enzymology, vol 582*. Elsevier; 2017:85-119.
This is a review on the integration of optical tweezers based single-molecule manipulation with fluorescence microscopy imaging to study DNA-protein interactions.
14. Noom MC, Van Den Broek B, Van Mameren J, Wuite GJ: *Nat Methods* 2007, **4**:1031.
15. Neuman KC, Nagy A: *Nat Methods* 2008, **5**:491EP.
16. De Vlaminck I, Dekker C: *Annu Rev Biophys* 2012, **41**:453.
17. Zhao X, Zeng X, Lu C, Yan J: *Nanotechnology* 2017, **28**:414002.
• This is a recent review on magnetic tweezers technology and its current application in studies of the force-dependent stability and interactions of biomolecules.
18. Strick TR, Allemand J-F, Bensimon D, Bensimon A, Croquette V: *Science* 1996, **271**:1835.
19. Halvorsen K, Wong WP: *Biophys J* 2010, **98**:L53.
20. Sitters G, Kamsma D, Thalhammer G, Ritsch-Marte M, Peterman EJ, Wuite GJ: *Nat Methods* 2014, **12**:47.
21. Yang D, Ward A, Halvorsen K, Wong WP: *Nat Commun* 2016, **7**:11026.
22. Yan J, Marko JF: *Phys Rev E* 2003, **68**:011905.
23. Smith SB, Cui Y, Bustamante C: *Science* 1996, **271**:795.
24. Maier B, Bensimon D, Croquette V: *Proc Natl Acad Sci U S A* 2000, **97**:12002.
25. Bell JC, Plank JL, Dombrowski CC, Kowalczykowski SC: *Nature* 2012, **491**:274.
26. Wuite GJ, Smith SB, Young M, Keller D, Bustamante C: *Nature* 2000, **404**:103.
27. Van Der Heijden T, Seidel R, Modesti M, Kanaar R, Wyman C, Dekker C: *Nucleic Acids Res* 2007, **35**:5646.
28. Koch SJ, Wang MD: *Phys Rev Lett* 2003, **91**:028103.
29. Hall MA, Shundrovsky A, Bai L, Fulbright RM, Lis JT, Wang MD: *Nat Struct Mol Biol* 2009, **16**:124.
30. Dittmore A, Landy J, Molzon AA, Saleh OA: *J Am Chem Soc* 2014, **136**:5974.
31. Mandal S, Koirala D, Selvam S, Ghimire C, Mao H: *Angew Chem* 2015, **127**:7717.
32. Camunas-Soler J, Alemany A, Ritort F: *Science* 2017, **355**:412.
•• This paper introduces a label-free assay to detect ligand binding to a DNA hairpin based on binding induced change in the unzipping kinetics using optical tweezers, and to quantify the binding energy using non-equilibrium fluctuation theorem.
33. Zhao X, Peter S, Dröge P, Yan J: *Sci Rep* 2017, **7**:8440.
•• This paper demonstrates label-free single-molecule quantification of the binding affinity of human HMGA2 protein to DNA based on delayed unzipping of DNA hairpin using magnetic tweezers.
34. Okoniewski SR, Uyetake L, Perkins TT: *Nucleic Acids Res* 2017,
•• **45**:10775.
This paper introduces the method of in-situ building of ssDNA and DNA hairpin by force-induced strand-separation transition.
35. Manosas M, Camunas-Soler J, Croquette V, Ritort F: *Nat Commun*
•• 2017, **8**:304.
This paper reports a label-free single-molecule approach to quantify the sequence-specific kinetics and affinity of ligand-DNA interactions based on delayed hairpin unzipping using magnetic tweezers.
36. Watson JD, Crick FH *et al.*: *Nature* 1953, **171**:737.
37. Cluzel P, Lebrun A, Heller C, Lavery R, Viovy J-L, Chatenay D, Caron F: *Science* 1996, **271**:792.
38. Williams MC, Wenner JR, Rouzina I, Bloomfield VA: *Biophys J* 2001, **80**:874.
39. Williams MC, Wenner JR, Rouzina I, Bloomfield VA: *Biophys J* 2001, **80**:1932.
40. Cocco S, Yan J, Léger J-F, Chatenay D, Marko JF: *Phys Rev E* 2004, **70**:011910.
41. van Mameren J, Gross P, Farge G, Hooijman P, Modesti M, Falkenberg M, Wuite GJ, Peterman EJ: *Proc Natl Acad Sci U S A* 2009, **106**:18231.
42. Fu H, Chen H, Marko JF, Yan J: *Nucleic Acids Res* 2010, **38**:5594.
43. Fu H, Chen H, Zhang X, Qu Y, Marko JF, Yan J: *Nucleic Acids Res* 2011, **39**:3473.
44. Zhang X, Chen H, Fu H, Doyle PS, Yan J: *Proc Natl Acad Sci U S A* 2012, **109**:8103.
45. Zhang X, Chen H, Le S, Rouzina I, Doyle PS, Yan J: *Proc Natl Acad Sci U S A* 2013, **110**:3865.
46. King GA, Gross P, Bockelmann U, Modesti M, Wuite GJ, Peterman EJ: *Proc Natl Acad Sci U S A* 2013, **110**:3859.
47. Paik DH, Perkins TT: *J Am Chem Soc* 2011, **133**:3219.
48. Candelli A, Hoekstra TP, Farge G, Gross P, Peterman EJ, Wuite GJ: *Biopolymers: Orig Res Biomol* 2013, **99**:611.
49. Fu H, Le S, Chen H, Muniyappa K, Yan J: *Nucleic Acids Res* 2012, **41**:924.
50. Chen J, Tang Q, Guo S, Lu C, Le S, Yan J: *Nucleic Acids Res* 2017,
• **45**:10032.
This paper demonstrates the application of force-induced strand-separation transition to generate a ssDNA and study its force-dependent interaction with homologous dsDNA.
51. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ: *Nature* 1997, **389**:251.
52. Raghunathan S, Ricard CS, Lohman TM, Waksman G: *Proc Natl Acad Sci U S A* 1997, **94**:6652.
53. Rice PA, wei Yang S, Mizuuchi K, Nash HA: *Cell* 1996, **87**:1295.
54. Le S, Chen H, Cong P, Lin J, Dröge P, Yan J: *Sci Rep* 2013, **3**:3508.
55. Amit R, Oppenheim AB, Stavans J: *Biophys J* 2003, **84**:2467.
56. Liu Y, Chen H, Kenney LJ, Yan J: *Genes Dev* 2010, **24**:339.
57. Chen J, Le S, Basu A, Chazin WJ, Yan J: *Sci Rep* 2015, **5**:9296.

58. Skoko D, Yan J, Johnson RC, Marko JF: *Phys Rev Lett* 2005, **95**:208101.

59. Yan J, Skoko D, Marko JF: *Phys Rev E* 2004, **70**:011905.

60. Vladescu ID, McCauley MJ, Nu nez ME, Rouzina I, Williams MC: *Nat Methods* 2007, **4**:517.

61. Biebricher AS, Heller I, Roijmans RF, Hoekstra TP, Peterman EJ, Wuite GJ: *Nat Commun* 2015, **6**:7304.

62. Meng H, Andresen K, Van Noort J: *Nucleic Acids Res* 2015, **43**:3578.

63. Brower-Toland BD, Smith CL, Yeh RC, Lis JT, Peterson CL, Wang MD: *Proc Natl Acad Sci U S A* 2002, **99**:1960.

64. Woodside MT, Behnke-Parks WM, Larizadeh K, Travers K, Herschlag D, Block SM: *Proc Natl Acad Sci U S A* 2006, **103**:6190.

65. Huguet JM, Bizarro CV, Forns N, Smith SB, Bustamante C, Ritort F: *Proc Natl Acad Sci U S A* 2010, **107**:15431.

66. Yang SW, Nash HA: *EMBO J* 1995, **14**:6292.

67. Marko JF: *Physica A: Stat Mech Appl* 2015, **418**:126.

68. Yao M, Chen H, Yan J: *Integr Biol* 2015, **7**:1154.

69. Sugimura S, Crothers DM: *Proc Natl Acad Sci U S A* 2006, **103**:18510.

70. Le S, Serrano E, Kawamura R, Carrasco B, Yan J, Alonso JC: *Nucleic Acids Res* 2017, **45**:8873.

71. Marko JF, Siggia ED: *Macromolecules* 1995, **28**:8759.

72. Bosco A, Camunas-Soler J, Ritort F: *Nucleic Acids Res* 2013, **42**:2064.

73. Lohman TM, Overman LB: *J Biol Chem* 1985, **260**:3594.

74. Cui T, Leng F: *Biochemistry* 2007, **46**:13059.

75. Borrmann L, Schwanbeck R, Heyduk T, Seebeck B, Rogalla P, Bullerdiek J, Wisniewski JR: *Nucleic Acids Res* 2003, **31**:6841.

76. You H, Lattmann S, Rhodes D, Yan J: *Nucleic Acids Res* 2016, **45**:206.

77. Luijsterburg MS, White MF, van Driel R, Dame RT: *Crit Rev Biochem Mol Biol* 2008, **43**:393.

78. Heller I, Hoekstra TP, King GA, Peterman EJ, Wuite GJ: *Chem Rev* 2014, **114**:3087.

79. Wang Y, Yan J, Goult BT: *Biochemistry* 2019 http://dx.doi.org/10.1021/acs.biochem.9b00453.

80. Cheung-Ong K, Giaever G, Nislow C: *Chem Biol* 2013, **20**:648.

81. Ehley JA, Melander C, Herman D, Baird EE, Ferguson HA, Goodrich JA, Dervan PB, Gottesfeld JM: *Mol Cell Biol* 2002, **22**:1723.

82. Lambert M, Jambon S, Depauw S, David-Cordonnier M-H: *Molecules* 2018, **23**:1479.

83. •• Gulvady R, Gao Y, Kenney LJ, Yan J: *Nucleic Acids Res* 2018, **46**:10216.
This paper demonstrates label-free single-molecule quantification of sequence-specific cooperative binding of the *E. coli* H-NS protein based on delayed unzipping of DNA hairpin.

84. Li M, Wang MD: *Methods in Enzymology, vol 513*. Elsevier; 2012:29-58.

85. Hodeib S, Raj S, Manosas M, Zhang W, Bagchi D, Ducos B, Fiorini F, Kanaan J, Le Hir H, Allemand J-F *et al.*: *Protein Sci* 2017, **26**:1314.

86. Lu C, Le S, Chen J, Byrd AK, Rhodes D, Raney KD, Yan J: *Nucleic Acids Res* 2019.

87. Lipfert J, Wiggin M, Kerssemakers JW, Pedaci F, Dekker NH: *Nat Commun* 2011, **2**:439.

88. Basu A, Schoeffler AJ, Berger JM, Bryant Z: *Nat Struct Mol Biol* 2012, **19**:538.

89. Revyakin A, Ebright RH, Strick TR: *Proc Natl Acad Sci U S A* 2004, **101**:4776.

90. Lipfert J, Kerssemakers JW, Jager T, Dekker NH: *Nat Methods* 2010, **7**:977.

91. La Porta A, Wang MD: *Phys Rev Lett* 2004, **92**:190801.

92. Yao M, Goult BT, Chen H, Cong P, Sheetz MP, Yan J: *Sci Rep* 2014, **4**:4610EP.

93. Yao M, Qiu W, Liu R, Efremov AK, Cong P, Seddiki R, Payre M, Lim CT, Ladoux B, Mège R-M, Yan J: *Nat Commun* 2014, **5**:4525EP.

94. Le S, Hu X, Yao M, Chen H, Yu M, Xu X, Nakazawa N, Margadant FM, Sheetz MP, Yan J: *Cell Rep* 2017, **21**:2714.

95. Pang SM, Le S, Yan J: *Biol Cell* 2018, **110**:65.