High-throughput single-molecule studies of protein–DNA interactions

Aaron D. Robison a, Ilya J. Finkelstein a,b,c,*

a Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX 78712, United States
b Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, TX 78712, United States
c Center for Systems and Synthetic Biology, The University of Texas at Austin, Austin, TX 78712, United States

ABSTRACT

Fluorescence and force-based single-molecule studies of protein–nucleic acid interactions continue to shed critical insights into many aspects of DNA and RNA processing. As single-molecule assays are inherently low-throughput, obtaining statistically relevant datasets remains a major challenge. Additionally, most fluorescence-based single-molecule particle-tracking assays are limited to observing fluorescent proteins that are in the low-nanomolar range, as spurious background signals predominate at higher fluorophore concentrations. These technical limitations have traditionally limited the types of questions that could be addressed via single-molecule methods. In this review, we describe new approaches for high-throughput and high-concentration single-molecule biochemical studies. We conclude with a discussion of outstanding challenges for the single-molecule biologist and how these challenges can be tackled to further approach the biochemical complexity of the cell.

Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.

1. Introduction

The advent of single-molecule methods—ultrasensitive tools that are capable of imaging and manipulating individual biochemical reactions—has revolutionized our understanding of biology. Single-molecule studies can directly interrogate transient biochemical steps that are obscured by ensemble averaging. These approaches are particularly useful for elucidating complex multi-step biochemical mechanisms and have proven especially amenable for studying protein–nucleic acid interactions. For example, single-molecule enzymology has shed critical insights into our understanding of DNA replication [1–3], transcription [4–6], chromatin remodeling [7], and DNA damage repair [8,9]. The development of single-molecule experiments in cell-free extracts [10–12] and within living cells [13,14] will continue to shed critical insights into all aspects of genome maintenance.

In the last two decades, the single-molecule methods toolkit has continued to expand at a dizzying pace. The choice of an appropriate method is dictated largely by the biochemical details and relevant length-scales of the desired biological process. For example, single-molecule Förster Resonance Energy Transfer (smFRET) can be used to monitor protein–nucleic acid interactions on the ~5 nm length-scale. For a complete discussion of smFRET-based approaches, we direct the reader towards several comprehensive reviews [15–19]. As DNA replication, transcription, and repair frequently involve highly processive molecular motors, these reactions must be studied on kilobase-length DNA substrates. These reactions can be indirectly visualized via tethered particle motion (TPM), where a long DNA molecule is used to tether a micron-size bead to the surface of a flowcell. Changes in the DNA length are observed as changes in the Brownian motion of the tethered beads [20–22]. To directly visualize biochemical reactions with high spatio-temporal resolution, DNA molecules are immobilized and extended on the surface of a microscope flow cell. The biochemical reaction is then followed by fluorescently tracking the enzyme or by monitoring a change in the length of the substrate DNA molecule.

These single-molecule particle-tracking experiments are hampered by two fundamental limitations. First, obtaining statistically relevant datasets is a challenge for experiments that are designed to observe individual molecules. This challenge is compounded by the fact that biochemical reconstitution of multi-subunit enzymatic machines (e.g. the replisome, chromatin remodelers, or DNA repair complexes) rarely approaches 100%. This biochemical heterogeneity further reduces the throughput of single-molecule data acquisition. Second, single-molecule fluorescence imaging must reliably discriminate weak signals from spurious background fluorescence. Most imaging experiments are carried out at extremely dilute (~1–10 nM) fluorescent protein concentrations. However, most transient biological interactions have evolved to be reversible...
## Table 1

High-throughput single-molecule methods.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Applications</th>
<th>Comments</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>smFRET</td>
<td>Observing protein conformations, protein-nucleic acid binding and short-distance translocation</td>
<td>Used for short-distance (1–5 nm) interactions</td>
<td>[15–19]</td>
</tr>
<tr>
<td>Microfluidic DNA curtains</td>
<td>Observing micron-length protein–DNA interactions</td>
<td>1000's of molecules in single field-of-view, defined DNA orientation</td>
<td>[26,27,32]</td>
</tr>
<tr>
<td>Random surface tethering DNA</td>
<td>Tens of molecules in single field-of-view. DNA orientation unknown</td>
<td>[1–3]</td>
<td></td>
</tr>
<tr>
<td>Tethered particle motion DNA</td>
<td>Monitors changes in DNA length, which can be used as an indirect probe of enzyme activity</td>
<td>100's of molecules in single field-of-view</td>
<td>[20]</td>
</tr>
<tr>
<td>Multiplexed magnetic tweezers</td>
<td>Force spectroscopy of protein–DNA interactions. Torsional control possible</td>
<td>10's to 100's of molecules in single field-of-view</td>
<td>[70,71,93]</td>
</tr>
<tr>
<td>Multiplexed optical tweezers</td>
<td>Observing mechanical properties of DNA and protein–DNA interactions</td>
<td>Offers 3D control of captured particles</td>
<td>[78,79]</td>
</tr>
<tr>
<td>Centrifugal force microscopy</td>
<td>Applies uniform centrifugal force on all molecules within an orbiting flowcell</td>
<td>Imaging must occur on a rotating stage</td>
<td>[65]</td>
</tr>
</tbody>
</table>

![Fig. 1](image-url)  

(a) An illustration of a DNA molecule organized at a lipid diffusion barrier (side view). (b) Fluorescent image of a λ-DNA curtain in the presence (top) and absence (bottom) of a 50 μl min⁻¹ buffer flow. In the absence of buffer flow (bottom panel), the DNA collapses and begins to diffuse away from the mechanical barrier. The DNA was stained with the intercalating dye YOYO-1. (c) Schematic of the objective-TIRF microscope used for imaging DNA curtains.
in the 1–100 μM range, precluding their analysis by conventional single-molecule methods [23].

In this review, we summarize emerging experimental approaches for interrogating protein–nucleic acid interactions at the single-molecule level. We focus on methods that permit the organization, manipulation, and imaging of long DNA molecules. In addition, we highlight a general strategy to break the concentration barrier in single-molecule fluorescence imaging studies. We conclude with a summary of next-generation single-molecule methods that combine fluorescence imaging with force spectroscopy to probe protein–nucleic acid interactions with unprecedented resolution.

2. Discussion

2.1. Tracking enzymes on long DNA substrates

To visualize enzymes that traverse a long DNA substrate, the DNA molecule is immobilized on a microscope flow cell surface in an extended conformation. This is typically achieved by tethering DNA with streptavidin–biotin linkages on poly(ethylene glycol)-coated surfaces and extending the tethered DNA with hydrodynamic flow (see Table 1) [24,25]. Alternatively, long DNA molecules can be suspended between two poly-L-lysine-coated silica beads adsorbed to the surface of a flow cell, forming DNA “tightropes.” As the DNA is tethered randomly on the flow cell surface, the number of DNA molecules per field-of-view must remain low. To avoid overlapping DNA molecules, only tens of DNA molecules are imaged within a single field-of-view. Moreover, because the DNA is randomly attached to the flow cell surface, individual DNA molecules have different tensions. Finally, for double-tethered DNA molecules, the orientation of the DNA sequence relative to its tether points is not known.

2.1.1. Microfluidic DNA curtains

To overcome these limitations, we have developed microfluidic “DNA curtains,” a high-throughput experimental platform for organizing and imaging hundreds of individual DNA molecules in a single field-of-view. To assemble DNA curtains, a supported lipid bilayer is first deposited on the surface of a microfluidic flow cell. The supported lipid bilayer provides excellent biomimetic surface passivation and DNA or proteins can be directly tethered to the lipid head groups via a biotin–streptavidin linkage [26,27]. As the bilayer forms a two-dimensional fluid, hydrodynamic flow can be used to organize hundreds of lipid-tethered DNA molecules at patterned barriers to lipid diffusion [Fig. 1]. Importantly, all DNA molecules have identical orientation with respect to their DNA sequences and buffer flow maintains all molecules at the same tension [28,29]. DNA curtains have been used to investigate how enzymes find specific DNA sites amidst a vast pool of non-specific DNA [30,31], to observe how motor proteins translocate on crowded DNA [32], and to visualize the roles of accessory polymerases during DNA replication [3]. Finally, the development of single-stranded DNA (ssDNA) curtains has facilitated single-molecule studies of homologous DNA recombination and other biochemical processes that occur on long tracks of ssDNA [33–35].

Although DNA curtains increase the number of molecules that can be imaged in a single field-of-view, numerous time-consuming and repetitive experiments are required to completely characterize a biochemical reaction (e.g. by changing the protein composition, or a nucleotide or salt concentration). To further increase the high-throughput capabilities of this experimental platform, we integrated DNA curtains with lab-on-chip poly(dimethylsiloxane) (PDMS) microfluidics [36,37]. As a proof-of-principle, we designed a chaotic gradient mixer which enables simultaneous analysis of DNA curtains in discrete microfluidic channels [Fig. 2a]. For example, we observed the rate of DNA resection by RecBCD, a heterotrimeric helicase and nuclease that uses the energy from ATP hydrolysis to translocate along DNA [38,39]. In Fig. 2b, RecBCD is not fluorescently labeled. Rather, its helicase/nuclease activity is observed as a shortening of the fluorescently-stained duplex DNA. RecBCD activity was imaged at five different ATP concentrations (in five analysis channels), allowing for rapid enzyme characterization in a single microfluidic chip [Fig. 2b and c].

![Fig. 2.](image)
2.1.2. Breaking the single-molecule concentration barrier

To fluorescently observe individual molecules, all single-molecule approaches must minimize spurious background fluorescent signals. Wide-field illumination via total internal reflection fluorescence (TIRF) microscopy reduces the laser excitation volume to a $\approx 100 \text{ nm}$ region near the surface of a coverslip, thereby substantially reducing background signals [40]. However, most TIRF-based experiments must still maintain the fluorophore concentration below $\approx 10 \text{ nM}$ to discriminate signal from background. Most methods that seek to image individual molecules at higher fluorophore concentrations either increase the local protein concentration, or further confine the laser illumination volume. For example, the reaction volume can be reduced by encapsulating the biochemical reaction of interest in a porous lipid vesicle [41–45], within a PDMS nanochannel [46], or within a confined volume induced by a convex lens and a coverslip [47]. Alternatively, the laser excitation can be confined to an attoliter volume within a zero-mode waveguide [48–50], or near plasmonic nano-structures that locally enhance the light excitation [51,52]. These methods are applicable for monitoring reactions that occur on short ($\approx 100$ bp) nucleic acid substrates and are listed in Table 2. We direct the reader to several recent reviews summarizing these approaches [15,16,53].

As extended DNA molecules cannot be encapsulated within small vesicles or circular zero-mode-waveguides, these methods cannot be used to image proteins on long DNA molecules. To
overcome this limitation, Loveland and co-workers developed a general approach to break the ‘concentration barrier’ for single-molecule experiments on both short and long DNA substrates [54]. In this approach, a protein of interest is labeled with mKikGR, a photo-convertible fluorescent protein that emits green fluorescence in its un-activated state (mKikG). Upon photo-activation with 405 nm light, an isomerization in the fluorophore active site shifts the emission spectrum to the red (mKikR), permitting spectral discrimination between the photoactivated and un-activated states [55]. Total internal reflection excitation is used to selectively photoconvert mKikG to mKikR near the surface [Fig. 3a]. Because unbound mKikR rapidly diffuses out of the illumination volume, only the DNA-bound mKikR is imaged. Using this method, termed PhADE (PhotoActivation, Diffusion, Excitation), Loveland et al. observed DNA replication in cell-free Xenopus laevis egg extracts [54]. By imaging mKikGR-labeled flap endonuclease 1 (Fen1mKikGR), the authors could dynamically visualize the Okazaki fragments of replicating λ-DNA molecules [Fig. 3c].

Two caveats must be considered when selecting this approach for single-molecule imaging at high fluorophore concentrations. First, as only a fraction of the mKikGR proteins are photoactivated by the 405 nm laser, the mKikGR-labeled protein must be present at a high density on the DNA molecule. Second, the mKikGR-labeled protein must not dissociate from the DNA molecule, as rapid exchange with un-activated protein still present in solution could rapidly ablate the mKikR signal. Despite these two caveats, PhADE provides the first general method to circumvent the concentration barrier in single-molecule studies on extended nucleic acid substrates and will greatly benefit from the continuing development of new photo-switchable fluorophores [56,57].

2.2. High-throughput force spectroscopy

Single-molecule force spectroscopy is a powerful tool for interrogating the mechanical properties of protein–nucleic acid interactions. Early force spectroscopy studies elucidated the mechanical properties of DNA and RNA [58–61]. These pioneering early experiments paved the way for mechanistic studies of protein–DNA interactions, such as those that probe the mechanical unzipping of DNA strands by helicases [62], the unwinding of
nucleosomes [63], or relaxation of supercoiled DNA strands by topoisomerases [64].

Most force spectroscopy methods, such as optical and magnetic tweezers, require the manipulation of DNA molecules on a one-by-one basis. To address this challenge, several groups have developed high-throughput force spectroscopy approaches. For example, Wong and colleagues developed a massively parallel centrifugal force microscope, where uniform piconewton forces are applied on thousands of molecules within an orbiting sample [65]. However, this method requires that both the sample chamber and the imaging optics must be within the same rotating frame, precluding the integration of modern microscopes and ultrasensitive CCD detectors. In addition, several groups have developed novel approaches for high-throughput optical and magnetic tweezers. Below, we highlight two of these approaches.

### 2.2.1. Magnetic tweezers

In a magnetic tweezers experiment, a DNA molecule is tethered between the surface of a flow cell and a paramagnetic bead. To extend or supercoil the DNA, an external magnetic field is used to manipulate the paramagnetic bead [Fig. 4a and b]. Protein-dependent activities are inferred from the bead movement [64,66–69].

To simultaneously manipulate hundreds of trapped DNA molecules, De Vlaminck et al. developed a strategy for depositing precisely controlled arrays of DNA-tethered beads [Fig. 4]. Repeating micron-scale arrays of anti-digoxigenin antibodies were printed onto a glass coverslip and the rest of the surface was passivated with a supported lipid bilayer [Fig. 4c]. DNA molecules were affixed to these pads via a digoxigenin-antibody linkage. The density of DNA molecules was tuned to minimize the nearest-neighbor paramagnetic bead croststalk probabilities [Fig. 4c and d] [70].

This approach offers a high-throughput strategy for single-molecule force spectroscopy. However, the number of beads that can be observed simultaneously is limited by non-uniformity of the applied magnetic field. To overcome this limitation, the authors analyzed the motion of the beads in a rotating magnetic field. Under a rotating magnet, the bi-circular rotational pattern of the paramagnetic beads is sensitive to both the angle of the applied magnetic force and the orientation of the bead-DNA attachment [71]. Systematic analysis of these rotational patterns allows accurate calculation of the magnetic force experienced by each bead and the bead-DNA attachment orientations, thereby compensating for inhomogeneity in the magnetic field. This calibration technique provides accurate analysis of protein-DNA interactions over large fields-of-view. Thus, by integrating micron-scale surface patterning with a sophisticated magnetic field calibration scheme, hundreds of surface-tethered molecules can be imaged within a single field-of-view.

### 2.2.2. Optical tweezers

Unlike magnetic tweezers, optical tweezers use highly focused laser beams to trap and manipulate polystyrene beads. To increase the throughput of optical tweezers experiments, a single beam can be time-shared via acousto-optical deflectors [72]. Alternatively, a single beam can be split into an array of optical traps through the use of computer-generated holograms [73,74], refractive microlenses [75,76], or mechanical gratings [77]. For example, Noom et al. were able to simultaneously trap four polystyrene beads by splitting a laser into two orthogonally polarized beams and keeping one of these beams as a stationary trap. The other beam was temporally shared between three independent trapping positions [Fig. 5] [78]. By time-sharing the laser between several traps, Noom et al. physically wrapped one DNA strand around a second independent DNA molecule, providing a means of ‘scanning’ one DNA molecule force spectroscopy assays [80–82].

### 3. Concluding remarks

Single-molecule studies continue to add tremendous insights into our understanding of protein–nucleic acid interactions. In this review, we discussed emerging high-throughput single-molecule methods for observing and manipulating long-range protein–DNA interactions [Table 1]. In addition, we discussed strategies for imaging individual molecules at high (μM) fluorophore concentrations [Table 2]. Further integration with highly multiplexed and temperature-controlled microfluidic-based systems will expand the throughput of single-molecule biophysical studies.

Complimentary aspects of a biochemical reaction can simultaneously be probed by a combination of single-molecule imaging and force spectroscopy modalities. For example, a combined fluorescence and optical tweezers microscope has been used to investigate protein–DNA interactions as a function of the DNA tension [83–85]. Magnetic tweezers have also been used in...
conjunction with fluorescence techniques such as FRET [86] and TIRF [87] and to visualize individual proteins bound to DNA [88]. Developing high-throughput versions of these methods will further enable single-molecule biophysical studies of multi-protein systems. Finally, the integration of new particle-manipulation modalities such as standing surface acoustic waves (SSAW) [89], hydrodynamic focusing [90], and electrokinetic traps [91] with existing fluorescence and force-manipulation techniques will further increase the information content of in vitro single-molecule approaches.

Acknowledgements

This research was supported in part by the Welch Foundation (F-1808) and by the NIH (GM097177 to I.J.F.). Dr. Ilya Finkelstein is a CPRIT Scholar in Cancer Research.

References


