2017

The more the merrier: high-throughput single-molecule techniques

Flynn R. Hill
University of Wollongong, flynn@uow.edu.au

Enrico Monachino
University of Wollongong, em805@uowmail.edu.au

Antoine M. van Oijen
University of Wollongong, vanoijen@uow.edu.au

The more the merrier: high-throughput single-molecule techniques

Abstract

The single-molecule approach seeks to understand molecular mechanisms by observing biomolecular processes at the level of individual molecules. These methods have led to a developing understanding that for many processes, a diversity of behaviours will be observed, representing a multitude of pathways. This realisation necessitates that an adequate number of observations are recorded to fully characterise this diversity. The requirement for large numbers of observations to adequately sample distributions, subpopulations, and rare events presents a significant challenge for single-molecule techniques, which by their nature do not typically provide very high throughput. This review will discuss many developing techniques which address this issue by combining nanolithographic approaches, such as zero-mode waveguides and DNA curtains, with singlemolecule fluorescence microscopy, and by drastically increasing throughput of force-based approaches such as magnetic tweezers and laminar-flow techniques. These methods not only allow the collection of large volumes of single-molecule data in single experiments, but have also made improvements to ease-of-use, accessibility, and automation of data analysis.

Disciplines

Medicine and Health Sciences | Social and Behavioral Sciences

The more the merrier: High-throughput single-molecule techniques

Flynn R. Hill[1], Enrico Monachino[1,2], Antoine M. van Oijen[1]*

[1]Centre for Medical and Medicinal Bioscience, Illawarra Health and Medical Research Institute and School of Chemistry, University of Wollongong, Wollongong, NSW, Australia

[2]Zernike Institute for Advanced Materials, University of Groningen, Groningen 9747AG, The Netherlands

* Corresponding author: vanoijen@uow.edu.au

The single-molecule approach seeks to understand molecular mechanisms by observing biomolecular processes at the level of individual molecules. These methods have led to a developing understanding that for many processes, a diversity of behaviours will be observed, representing a multitude of pathways. This realisation necessitates that an adequate number of observations are recorded to fully characterise this diversity. The requirement for large numbers of observations to adequately sample distributions, subpopulations and rare events presents a significant challenge for single-molecule techniques, which by their nature do not typically provide very high throughput. This review will discuss a number of developing techniques which address this issue by combining nanolithographic approaches, such as zero-mode waveguides and DNA curtains, with single-molecule fluorescence microscopy, and by drastically increasing throughput of force-based approaches such as magnetic tweezers and laminar-flow techniques. These methods not only allow the collection of large volumes of single-molecule data in single experiments, but have also made improvements to ease-of-use, accessibility and automation of data analysis.

If we consider the molecular processes that are central to life – DNA replication, transcription, translation, cell division *etc.* – intricate, precise and highly evolved mechanisms come to mind. These processes have been subject to tremendous selection pressure over aeons, resulting in highly efficient mechanisms with little room for error. Textbooks often depict these processes moving like clockwork

without detours or off-pathway transitions. However, a picture is now emerging of dynamic molecular pathways which, rather than being shepherded deterministically through a linear series of events, are more free-form than any textbook model (1). On careful consideration, such a picture makes sense – kinetic pathways can only be determined by the thermodynamic boundary conditions of the system, an intricate interplay of interaction strengths, concentrations, avidity effects and catalytic rates of the molecular species involved. These all contribute to a reaction free-energy landscape of peaks and troughs which may result in a preferential pathway, but could allow many different routes (**Figure 1**). Stochasticity thus has a pervading influence and complicated, multi-faceted processes require plasticity to achieve robustness. Experimentally resolving the mechanistic details of such pathways presents a special challenge. In traditional, ensemble-based experiments, which measure the behaviour of a population of molecules as a whole, underlying distributions representing the diversity of kinetic behaviour are hidden. By contrast, single-molecule techniques follow the reaction trajectories of individual molecules and have tremendous power to resolve the fine details of multi-step pathways. Such an approach allows distributions of pathways to be measured, rather than merely the distributions of quantifiable endpoint parameters. For complex biological processes, heterogeneity is typically observed within single molecule trajectories as a function of time, as well as across many single molecules in an ensemble; these two types of heterogeneity are not necessarily equivalent. *Static disorder* describes the heterogeneity present across molecules and complexes that does not change through time, while *dynamic disorder* describes the changing behaviour of single-molecule trajectories through time. Truly high-throughput single-molecule techniques should be capable of observing large numbers of molecules, each for a sufficient duration so that a representative sampling of both static and dynamic disorder can be obtained.

When a process is comprised of a diversity of pathways, obtaining a representative sampling may require hundreds of single-molecule trajectories. Most single-molecule techniques, however, have

only limited throughput, typically following trajectories of at most a handful of molecules at once. The experimental setup for these techniques can be time-consuming and technically demanding, often leading to a woefully low rate of experimental success and overall throughput. Data processing is thus an exercise in separating signal from noise, potentially leading to the introduction of biases in the analysis and interpretation that may not be discernible when curated results are published. High-throughput approaches promise to address these problems by improving the signal-to-noise ratio and making automated methods of analysis viable, and hence will be crucial to the more widespread adoption of single molecule techniques (2, 3). Single-molecule techniques include a variety of methods for following reaction trajectories of individual fluorescently-labelled molecules, such as single-molecule FRET, as well as force-based mechanical techniques including atomic force spectroscopy, and optical/magnetic tweezers (4-7). In general, fluorescence-based techniques are particularly suited to real-time measurement of protein stoichiometries, exchange events, conformational changes and enzymatic catalysis, and many of these techniques can be applied *in vivo* as well as *in vitro*. Although force-based techniques have a somewhat more limited scope, they are particularly suited to the study of protein-nucleic acid interactions and the dynamics of molecular motors – topics of particular interest in the field of single-molecule biophysics.

Nucleic acids provide an ideal stage on which to observe single-molecule processes and nucleic acid-focussed systems have led the way since the inception of the single-molecule field. This focus on nucleic acids continues with the advent of high-throughput techniques, however, improvements to throughput are also being made in other generally applicable and specialised methods, particularly through improvement to optics and sensors (8-10). In this review, we will discuss fluorescence and force-based single-molecule methods with a particular focus on platforms based on nucleic acids and how innovative approaches to increasing their throughput have contributed to significant new findings.

Fluorescence techniques

Single-Molecule Real-Time DNA Sequencing

The reigning champion of high-throughput single-molecule techniques was developed with a commercial motivation. Pacific Bioscience's single-molecule real-time (SMRT) DNA-sequencing platform is capable of simultaneously following nucleotide incorporation events by many thousands of DNA polymerases, allowing rapid whole-genome sequencing (11-14). As a commercial product, the system is standardised and automated, including automated reagent injection, downstream data processing and analysis. Rapid DNA sequencing is achieved by multiplexing the sequencing-by-elongation reaction into discrete zero-mode waveguides (ZMWs), nanostructured reaction wells which are arrayed on a thin metal layer over a glass substrate (15) (**Figure 2**). In the current generation of SMRT sequencers, there are one million ZMWs per reaction cell. Each ZMW has dimensions smaller than the wavelength of the illuminating laser light, and has a volume in the zeptolitre (10^{-21} L) range. Since millimolar concentrations correspond to single molecules within zeptolitre volumes, the very small illumination volume of the ZMW allows high concentrations of fluorophore-labelled substrates to be present in the reaction volume.

The SMRT sequencing reaction is carried out by a highly processive modified phi 29 DNA polymerase enzyme tethered to the surface, with an average of one polymerase per ZMW. In the presence of all four dNTPs, each labelled at the terminal phosphate with a unique fluorophore, the polymerase carries out replication of the target DNA. By using ZMWs, dNTPs can be present at a concentration conducive to replication at rates of ~1-2 nt/s, with minimal background fluorescence, enabling rapid elongation and thus sequencing. Under laser illumination, each nucleotide incorporation event is associated with a burst of fluorescence, the colour of which identifies the incorporated nucleotide. Multi-megapixel image sensors record the sequence of nucleotides incorporated recorded

for each polymerase-containing ZMW. The consensus sequence is then generated by an automated data-processing pipeline.

The single-molecule resolution, the ability to use high concentrations of fluorophore-tagged ligands, and the massive throughput of the Pacific Biosciences sequencing platform have also proven useful as a platform for basic research. Puglisi and colleages have described how the SMRT sequencer can be modified for highly multiplexed studies of biomolecular interactions (16) (17), and they have put this system to use studying protein-RNA interaction dynamics. A recent study used the ZMW platform to study translation initiation at the cricket paralysis virus (CrPV) internal ribosome entry site (IRES), a structured feature of the CrPV mRNA which circumvents the requirement for translation initiation factors (Petrov *et al.*, 2016). As such, this represents a minimal system for studying the initiation of translation, with the complexity of multiple pathways associated with the various eukaryotic initiation factors removed. By fluorescently labelling the IRES, tRNA, and the 40S and 60S subunits with fluorophores of different colour, multicolour experiments using the ZMW array were performed to study the assembly of the 80S ribosome on the surface-immobilised IRES, as well as association of fluorescently labelled tRNA. Similar to the sequencing reaction, the arrival and persistence of any labelled component at the ZMW is measured by a step-wise increase in fluorescence, and multiple colours can be monitored simultaneously. This technique allows association and dissociation rate constants to be determined, and was used to compare the probabilities of various possible modes of assembly. It was found that functional ribosome assembly could occur either with the 40S and 60S ribosomal subunits recruited by the IRES in sequence, or simultaneously with pre-formed 80S ribosomes binding to the IRES, with the Mg^{2+} concentration controlling the balance between these two pathways. By using two different mRNA templates, one with the first codon of the 0 frame encoding phenylalanine, and another doing so with the first codon of the +1 frame, the preference of tRNA^{Phe}-Cy3-Phe for the 0 or +1 frame at the start of elongation was measured. It was found that binding of the tRNA to the 0

frame was faster and more efficient than binding to the +1 frame, and that in both cases elongation factors were required for efficient binding to occur. Although the ribosomal complex used by the authors represents a minimal system for eukaryotic translation, multiple pathways for ribosome assembly and the initiation of elongation were observed. This study would not have been possible without the high-throughput ZMW approach that allowed these different reaction pathways and conditions to be explored.

DNA Curtains

Total Internal Reflection Fluorescence (TIRF) microscopy is one of the most successful and widely used techniques in single-molecule biophysics. The strength of this method lies in the selective illumination of only a thin layer of sample close to the surface of a glass substrate (18, 19). By excluding the fluorescence from above the illuminated volume, significant background reduction is achieved and individual fluorophores can be imaged with high signal-to-background ratios. By tagging proteins of interest with fluorophores, stoichiometries, interactions, exchange events, and many other parameters can be tracked *in vitro* or *in vivo*. A range of imaging techniques and experimental strategies are available to make TIRF microscopy supremely adaptable. In protein-DNA studies, linear DNA molecules are typically tethered to the surface of a microscope cover slip using a binding partner that is randomly distributed on the surface, such as streptavidin in the case of biotin-functionalised DNA. An alternative strategy is the DNA-curtains technique, in which long DNA templates are aligned in rows on the surface (**Figure 3**). This approach has the advantages of providing a high surface density, which increases throughput, and greatly improving the spatial localisation of fluorescent proteins along the DNA sequence. This method has been pioneered by the lab of Eric Greene, relying on nanofabrication approaches such as electron-beam and UV lithography (2, 20-22). In brief, these lithographic techniques are used to create nanofabricated chromium diffusion barriers in rows on a coverslip surface. A flow cell

is then assembled on top of the coverslip, and a lipid bilayer is coated onto the surface with DNA molecules tethered to biotinylated phospholipids in the bilayer. When buffer flow is applied, the DNA tethers diffuse to the diffusion barriers, resulting in the alignment of the DNA in curtains. The spacing between diffusion barriers is only slightly greater than the DNA length, to maximise the number of rows of curtains that can be imaged in a single field of view. DNA curtains are particularly well suited for observing sequence-dependent interactions; as the tether point for each DNA template is known with high accuracy, sequences of interest can be closely matched to the distance from the tether point. This technique has become a mainstay for the Greene and Finkelstein groups, and has been used in a number of studies of protein-DNA interactions (2, 20-43).

DNA curtains are particularly well suited to identifying the means by which DNA-interacting proteins search for their target sequences and sites, balancing the need for speed in the search phase with stability in the target-bound phase (44). A recent study investigated how the eukaryotic mismatch repair complex Msh2-Msh3 finds its target sites on crowded DNA, in the presence of nucleosomes and other protein roadblocks (43). By using doubly-tethered DNA curtains of lambda-phage DNA (with a length of 48.5 kbp) and Msh2-Msh3 complexed with a quantum dot (QD), the diffusion of the complex on DNA could be visualised and tracked. To evaluate whether the protein diffuses on DNA using a sliding (translocation in continuous DNA contact) or hopping (correlated dissociation and reassociation) mechanism, the ionic strength was varied so that the level of electrostatic interaction with the DNA was controlled. The diffusion coefficient was observed to increase with the ionic strength in a manner characteristic of a mechanism of alternating hopping and sliding. In the presence of nucleosomes or other protein roadblocks, the Msh2-Msh3 complex was observed to bypass the roadblock and continue its search. In contrast, another mismatch repair complex, Msh2-Msh6, was shown in a previous DNA curtains study to have no dependence of diffusion coefficient on the ionic strength, indicative of a sliding mechanism (45). Without hopping, this complex was shown to be unable to bypass roadblocks.

The ability to monitor large numbers of DNA molecules simultaneously was crucial in allowing the authors to visualise the interplay between the repair complexes and the nucleosomes, and to dissect how the mode of protein translocation dictates its ability to bypass nucleosomal roadblocks.

One drawback of the DNA curtains technique is the expense and difficulty of the lithographic fabrication methods involved. One way this could possibly be improved is by the use of soft lithography, in which a finely patterned, reusable elastomeric stamp is used to deposit linking molecules such as streptavidin onto the flow cell surface (46). Although soft lithography has been used to deposit patterns on surfaces in single-molecule studies, including for ordered DNA tethering (47-51), this technology has not been employed yet to support the production of DNA curtains.

Force-based methods

Flow-stretching DNA

Force-based single-molecule methods rely on the application of a stretching force to a biopolymer, such as DNA, and the readout of the length of the molecule as a response to that force or to proteins acting on the polymer. Many force-based methods use wide-field microscopy for data collection, and are readily amenable to scaling up by maximising the size of the field of view and the density of species of interest. One such method is the tethered-bead flow-stretching technique (1, 52). Here, antibody-functionalised microbeads are tethered to the functionalised surface of a flow cell via a long double-stranded DNA. A constant flow rate applies a controllable drag force to the beads in the low piconewton range, and they are imaged with a very high signal-to-noise ratio by dark-field microscopy (**Figure 4**). Movement of the bead reports on changes to the DNA length, caused by interconversion between dsDNA and ssDNA, compaction by DNA binding factors, or the formation and release of topological features such as loops and coils. Because the beads are spherically symmetric, their centroid positions

can be easily tracked by sub-pixel fitting, at a precision of ~1-10 nm (52-55). This resolution is possible even when low-magnification optics are used, provided the resolution is sufficient that each bead projects onto multiple pixels on the image sensor. Brownian motion of the beads, mechanical drift and small fluctuations in the flow rate are the main sources of experimental uncertainty.

The mechanism of DNA replication is a well-studied example of how a multi-protein machine supports a complicated series of enzymatic events, and the flow stretching technique is well-suited to observing its dynamics (56-63). Because duplex DNA is antiparallel, one nascent daughter strand, the leading strand, is extended continously in the same direction as replication progression, while the lagging strand is extended discontinuously to form Okazaki fragments in the opposite direction. Each Okazaki fragment is initiated by an RNA primer synthesised by the primase subunit, and is extended by the polymerase. It has been observed that under conditions in which the replisome is first assembled on DNA before being subjected to rapid dilution, leading and lagging-strand replication persists well beyond a single Okazaki-fragment length, implying that in the absence of free polymerase in solution a mechanism exists to recycle the lagging-strand polymerase and use it to synthesize multiple Okazaki fragments (64-66). In the classic 'trombone-loop' model for DNA replication, the lagging-strand polymerase remains physically coupled to the replication fork, forcing DNA to loop out in between the helicase and the lagging-strand polymerase (67). Following the completion of each Okazaki fragment, the loop is released and a new cycle begins with recycling of the polymerase to a newly synthesised primer. While this model was shown to describe the behaviour of the replication complex under conditions of assembly followed by rapid dilution, it had not been thoroughly tested in the physiologically-relevant context of free polymerase in solution.

To understand the coordination between leading and lagging-strand DNA replication, Duderstadt *et al.* created a DNA template for the flow-stretching technique with each end of the DNA attached to a bead, and with a replication fork located at the middle and bound to the flow cell surface

(1). One bead reports on lagging-strand dynamics, while the other reports on a sum of leading and lagging-strand dynamics, which can be deconvoluted. The low frequency of the replication events being studied, and the scarcity of DNA templates with a bead bound at each end necessitated the scale-up of this method to high throughput. Using a low-magnification telecentric lens coupled with a high-megapixel camera, an ultra-wide field of view (5.3 mm x 3.5 mm) was visualised containing tens of thousands of beads that each could be monitored simultaneously. Key steps of the data analysis were automated, including identification of bead pairs, and fitting of line segments to their event trajectories using a statistical technique called change-point detection. This workflow ensured the data could be processed efficiently and sources of bias minimised. In a demonstration of the plastic, multi-pathway nature of DNA replication, Duderstadt *et al.* observed that in the presence of free polymerase in solution, Okazaki-fragment synthesis by the bacteriophage T7 replisome could occur either within replication loops or in their absence, with the majority of lagging-strand replication taking place outside the context of replication loops. Moreover, both modalities of DNA replication were observed within individual trajectories of processive replication reactions. While looping was observed during most cycles of Okazaki-fragment synthesis, the information provided by both the leading and lagging-strand beads indicated that most of these were priming loops – where DNA loops out between the helicase and primase subunits during primer synthesis. The detailed study of the extremely complex dynamic behaviour of the replication machinery relied critically on the high throughput of the flow-stretching method and would have been exceedingly challenging with any other technique.

Magnetic Tweezers

A related technique, magnetic tweezers, has similarly been shown to be amenable to scaling up to high throughput (3, 68-70), and its application to multi-pathway processes has been described in detail (71). In this method, a force is applied to superparamagnetic beads attached to nucleic-acid polymers in a

flow cell by using permanent magnets, and similarly to the flow-stretching technique, these are imaged and tracked by wide-field microscopy. To achieve high throughput, Dekker and colleagues use a relatively large field of view, 400 x 300 μm, with a high-resolution 12 MPix camera. As in the flow-stretching experiments, the power of sub-pixel fitting means that even at low magnification, the beads can be tracked with no sacrifice of resolution. With hundreds of tethered microbeads in the field of view, and a fast imaging frame rate of 58 Hz, tracking of the beads for a long-duration experiment could potentially cause a bottleneck in data analysis. To speed up this step, Cnossen *et al.* developed GPU-based tracking software that makes use of the NVIDIA CUDA platform for parallelisation and tracks all the beads in real-time (68). The high-throughput power of this technique has been used to study error incorporation by a viral RNA-dependent RNA polymerase (RdRP) (69). In this experiment, beads were tethered to double-stranded RNA, which was converted to single-stranded RNA by the RdRP, causing lengthening of the RNA tether and commensurate movement of the bead. An unbiased, automated dwell-time analysis method was applied in which the dwell times for the polymerase in consecutive 10-nt windows were quantified and their distributions in a range of conditions characterised. Distinct peak features were present in the dwell-time distribution, indicating different pathways for elongation with differing error rates. These include brief dwell times associated with elongation with no pausing, intermediate dwell times associated with pausing and error incorporation, and longer dwell times caused by backtracking events. The statistical significance of this analysis critically relied on access to the hundreds of trajectories collected using this high-throughput technique.

Conclusions

The growth of the single-molecule biophysics field and the adoption of its techniques by researchers entering the field is dependent on the development of high-throughput, adaptable and reliable techniques. The nucleic acid-focussed techniques reviewed in this article can certainly provide high

throughput, however, there are still gains to be made in adaptability, expense and ease-of-use. Integrated systems for single-molecule experiments and data analysis are starting to become available as commercial packages, for example the acoustic force spectroscopy system produced by Lumicks (72, 73). Single-molecule methods could also benefit from a greater consensus around standards for statistical methods and standards of reporting. In contrast to many established techniques in biochemistry such as gel electrophoresis and chromatography, single-molecule methods produce data that must be extensively processed, analysed and interpreted, and this is especially true of high-throughput techniques. As the field matures, it will be crucial for conventions in statistical methods and transparency to be developed. An example of this is the interpretation of histograms of measured single-molecule kinetic parameters. It is common practise to fit experimentally obtained distributions with a sum of multiple underlying distributions, each corresponding to an ostensibly distinct sub-population, however, statistical tests are often missing to justify the assumption of multi-modality. While high-throughput techniques can greatly improve the statistical significance required to justify multi-modality, it can also increase the temptation to over-fit. It may also become desirable in the future for raw data sets, or randomly selected subsets of these, to be routinely made available online as supplementary data, along with the tools and instructions necessary to replicate published analyses.

Funding

References

1.   Duderstadt Karl E, Geertsema Hylkje J, Stratmann Sarah A, Punter Christiaan M, Kulczyk Arkadiusz W, Richardson Charles C, et al. (2016) Simultaneous real-time imaging of leading and lagging strand synthesis reveals the coordination dynamics of single replisomes. *Mol Cell* **64**, 1035-47 doi:10.1016/j.molcel.2016.10.028.

2.   Robison AD, Finkelstein IJ (2014) High-throughput single-molecule studies of protein-DNA interactions. *FEBS Lett* **588**, 3539-46 doi:10.1016/j.febslet.2014.05.021.

3.   Berghuis BA, Kober M, van Laar T, Dekker NH (2016) High-throughput, high-force probing of DNA-protein interactions with magnetic tweezers. *Methods* **105**, 90-8 doi:10.1016/j.ymeth.2016.03.025.

4.   Xie XS, Lu HP (1999) Single-molecule enzymology. *J Biol Chem* **274**, 15967-70

5.   Moerner WE, Fromm DP (2003) Methods of Single-Molecule Fluorescence Spectroscopy and Microscopy. *Rev Sci Instrum* **74**, 3597-619 doi:10.1146/annurev.physchem.55.091602.094340.

6.   Mehta AD, Rief M, Spudich JA, Smith DA, Simmons RM (1999) Single-molecule biomechanics with optical methods. *Science* **283**, 1689-95 doi:10.1126/science.283.5408.1689.

7.   Neuman KC, Nagy A (2008) Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat Methods* **5**, 491-505 doi:10.1038/nmeth.1218.

8.   Gulinatti A, Rech I, Maccagnani P, Ghioni M. A 48-pixel array of Single Photon Avalanche Diodes for multispot Single Molecule analysis. In: Razeghi M, Tournie E, Brown GJ, editors. Quantum Sensing and Nanophotonic Devices X. Proceedings of SPIE. 8631. Bellingham: Spie-Int Soc Optical Engineering; 2013.

9.   Zhang Z, Kenny SJ, Hauser M, Li W, Xu K (2015) Ultrahigh-throughput single-molecule spectroscopy and spectrally resolved super-resolution microscopy. *Nat Meth* **12**, 935-8 doi:10.1038/nmeth.3528.

10.   Wachsmuth M, Conrad C, Bulkescher J, Koch B, Mahen R, Isokane M, et al. (2015) High-throughput fluorescence correlation spectroscopy enables analysis of proteome dynamics in living cells. *Nat Biotech* **33**, 384-9 doi:10.1038/nbt.3146

http://www.nature.com/nbt/journal/v33/n4/abs/nbt.3146.html#supplementary-information.

11.   Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, et al. (2009) Real-time DNA sequencing from single polymerase molecules. *Science* **323**, 133-8 doi:10.1126/science.1162986.

12.   Roberts RJ, Carneiro MO, Schatz MC (2013) The advantages of SMRT sequencing. *Genome Biol* **14**, 405 doi:10.1186/gb-2013-14-6-405.

13.   Hestand MS, Van Houdt J, Cristofoli F, Vermeesch JR (2016) Polymerase specific error rates and profiles identified by single molecule sequencing. *Mutat Res* **784-785**, 39-45 doi:10.1016/j.mrfmmm.2016.01.003.

14.   Zhu P, Craighead HG (2012) Zero-mode waveguides for single-molecule analysis. *Annu Rev Biophys* **41**, 269-93 doi:10.1146/annurev-biophys-050511-102338.

15.   Levene MJ, Korlach J, Turner SW, Foquet M, Craighead HG, Webb WW (2003) Zero-mode waveguides for single-molecule analysis at high concentrations. *Science* **299**, 682-6 doi:10.1126/science.1079700.

16.   Chen J, Dalal RV, Petrov AN, Tsai A, O'Leary SE, Chapin K, et al. (2014) High-throughput platform for real-time monitoring of biological processes by multicolor single-molecule fluorescence. *Proc Natl Acad Sci USA* **111**, 664-9 doi:10.1073/pnas.1315735111.

17.   Uemura S, Aitken CE, Korlach J, Flusberg BA, Turner SW, Puglisi JD (2010) Real-time tRNA transit on single translating ribosomes at codon resolution. *Nature* **464**, 1012-7 doi:10.1038/nature08925.

18.   Axelrod D, Burghardt TP, Thompson NL (1984) Total internal reflection fluorescence. *Annu Rev Biophys Bioeng* **13**, 247-68 doi:10.1146/annurev.bb.13.060184.001335.

19.   Axelrod D (2001) Total internal reflection fluorescence microscopy in cell biology. *Traffic* **2**, 764-74 doi:10.1034/j.1600-0854.2001.21104.x.

20.   Gallardo IF, Pasupathy P, Brown M, Manhart CM, Neikirk DP, Alani E, et al. (2015) High-Throughput Universal DNA Curtain Arrays for Single-Molecule Fluorescence Imaging. *Langmuir* **31**, 10310-7 doi:10.1021/acs.langmuir.5b02416.

21.   Silverstein TD, Gibb B, Greene EC (2014) Visualizing protein movement on DNA at the single-molecule level using DNA curtains. *DNA Repair (Amst)* **20**, 94-109 doi:10.1016/j.dnarep.2014.02.004.

22.   Finkelstein IJ, Greene EC (2011) Supported lipid bilayers and DNA curtains for high-throughput single-molecule studies. *Methods Mol Biol* **745**, 447-61 doi:10.1007/978-1-61779-129-1_26.

23.   Visnapuu ML, Greene EC (2009) Single-molecule imaging of DNA curtains reveals intrinsic energy landscapes for nucleosome deposition. *Nat Struct Mol Biol* **16**, 1056-62 doi:10.1038/nsmb.1655.

24.   Visnapuu ML, Fazio T, Wind S, Greene EC (2008) Parallel arrays of geometric nanowells for assembling curtains of DNA with controlled lateral dispersion. *Langmuir* **24**, 11293-9 doi:10.1021/la8017634.

25.   Robison AD, Finkelstein IJ (2014) Rapid prototyping of multichannel microfluidic devices for single-molecule DNA curtain imaging. *Anal Chem* **86**, 4157-63 doi:10.1021/ac500267v.

26.   Qi Z, Greene EC (2016) Visualizing recombination intermediates with single-stranded DNA curtains. *Methods* **105**, 62-74 doi:10.1016/j.ymeth.2016.03.027.

27.   Prasad TK, Yeykal CC, Greene EC (2006) Visualizing the assembly of human Rad51 filaments on double-stranded DNA. *J Mol Biol* **363**, 713-28 doi:10.1016/j.jmb.2006.08.046.

28.   Myler LR, Gallardo IF, Zhou Y, Gong F, Yang SH, Wold MS, et al. (2016) Single-molecule imaging reveals the mechanism of Exo1 regulation by single-stranded DNA binding proteins. *Proc Natl Acad Sci USA* **113**, E1170-9 doi:10.1073/pnas.1516674113.

29.   Lee JY, Wang F, Fazio T, Wind S, Greene EC (2012) Measuring intermolecular rupture forces with a combined TIRF-optical trap microscope and DNA curtains. *Biochem Biophys Res Commun* **426**, 565-70 doi:10.1016/j.bbrc.2012.08.127.

30.   Lee JY, Greene EC (2011) Assembly of recombinant nucleosomes on nanofabricated DNA curtains for single-molecule imaging. *Methods Mol Biol* **778**, 243-58 doi:10.1007/978-1-61779-261-8_16.

31.   Lee JY, Finkelstein IJ, Crozat E, Sherratt DJ, Greene EC (2012) Single-molecule imaging of DNA curtains reveals mechanisms of KOPS sequence targeting by the DNA translocase FtsK. *Proc Natl Acad Sci USA* **109**, 6531-6 doi:10.1073/pnas.1201613109.

32.   Greene EC, Wind S, Fazio T, Gorman J, Visnapuu ML (2010) DNA curtains for high-throughput single-molecule optical imaging. *Methods Enzymol* **472**, 293-315 doi:10.1016/S0076-6879(10)72006-1.

33.   Gorman J, Plys AJ, Visnapuu ML, Alani E, Greene EC (2010) Visualizing one-dimensional diffusion of eukaryotic DNA repair factors along a chromatin lattice. *Nat Struct Mol Biol* **17**, 932-8 doi:10.1038/nsmb.1858.

34.   Gorman J, Greene EC (2013) Target search dynamics during post-replicative mismatch repair. *Cell Cycle* **12**, 537-8 doi:10.4161/cc.23669.

35.   Gorman J, Fazio T, Wang F, Wind S, Greene EC (2010) Nanofabricated racks of aligned and anchored DNA substrates for single-molecule imaging. *Langmuir* **26**, 1372-9 doi:10.1021/la902443e.

36.   Gibb B, Ye LF, Gergoudis SC, Kwon Y, Niu H, Sung P, et al. (2014) Concentration-dependent exchange of replication protein A on single-stranded DNA revealed by single-molecule imaging. *PLoS One* **9**, e87922 doi:10.1371/journal.pone.0087922.

37.   Gibb B, Silverstein TD, Finkelstein IJ, Greene EC (2012) Single-stranded DNA curtains for real-time single-molecule visualization of protein-nucleic acid interactions. *Anal Chem* **84**, 7607-12 doi:10.1021/ac302117z.

38.   Finkelstein IJ, Visnapuu ML, Greene EC (2010) Single-molecule imaging reveals mechanisms of protein disruption by a DNA translocase. *Nature* **468**, 983-7 doi:10.1038/nature09561.

39.   Fazio TA, Visnapuu M, Greene EC, Wind SJ (2009) Fabrication of Nanoscale "Curtain Rods" for DNA Curtains Using Nanoimprint Lithography. *J Vac Sci Technol A* **27**, 3095-8 doi:10.1116/1.3259951.

40.   Fazio TA, Lee JY, Wind SJ, Greene EC (2012) Assembly of DNA curtains using hydrogen silsesquioxane as a barrier to lipid diffusion. *Anal Chem* **84**, 7613-7 doi:10.1021/ac302149g.

41.   Fazio T, Visnapuu ML, Wind S, Greene EC (2008) DNA curtains and nanoscale curtain rods: high-throughput tools for single molecule imaging. *Langmuir* **24**, 10524-31 doi:10.1021/la801762h.

42.   Collins BE, Ye LF, Duzdevich D, Greene EC (2014) DNA curtains: novel tools for imaging protein-nucleic acid interactions at the single-molecule level. *Methods Cell Biol* **123**, 217-34 doi:10.1016/B978-0-12-420138-5.00012-4.

43.   Brown MW, Kim Y, Williams GM, Huck JD, Surtees JA, Finkelstein IJ (2016) Dynamic DNA binding licenses a repair factor to bypass roadblocks in search of DNA lesions. *Nat Commun* **7**, 10607 doi:10.1038/ncomms10607.

44.   Tafvizi A, Mirny LA, van Oijen AM (2011) Dancing on DNA: kinetic aspects of search processes on DNA. *Chemphyschem* **12**, 1481-9 doi:10.1002/cphc.201100112.

45.   Gorman J, Chowdhury A, Surtees JA, Shimada J, Reichman DR, Alani E, et al. (2007) Dynamic basis for one-dimensional DNA scanning by the mismatch repair complex Msh2-Msh6. *Mol Cell* **28**, 359-70 doi:10.1016/j.molcel.2007.09.008.

46.   Xia YN, Whitesides GM (1998) Soft lithography. *Angew Chem Int Ed* **37**, 550-75 doi:10.1002/(SICI)1521-3773(19980316)37:5<550::AID-ANIE550>3.0.CO;2-G.

47.   Renault J, Bernard A, Bietsch A, Michel B, Bosshard H, Delamarche E, et al. (2003) Fabricating arrays of single protein molecules on glass using microcontact printing. *J Phys Chem B* **107**, 703-11 doi:10.1021/jp0263424.

48.   Biyani M, Ichiki T (2015) Microintaglio Printing for Soft Lithography-Based in Situ Microarrays. *Microarrays (Basel)* **4**, 311-23 doi:10.3390/microarrays4030311.

49.   Jo K, Dhingra DM, Odijk T, de Pablo JJ, Graham MD, Runnheim R, et al. (2007) A single-molecule barcoding system using nanoslits for DNA analysis. *Proc Natl Acad Sci USA* **104**, 2673-8 doi:10.1073/pnas.0611151104.

50.   Sperling E, Hohlfeld M, Mertig M (2015) Soft-lithographically fabricated nanofluidic channels for single-DNA measurements. *physica status solidi (a)* **212**, 1229-33 doi:10.1002/pssa.201431915.

51.   Yasaki H, Onoshima D, Yasui T, Yukawa H, Kaji N, Baba Y (2015) Microfluidic transfer of liquid interface for parallel stretching and stamping of terminal-unmodified single DNA molecules in zigzag-shaped microgrooves. *Lab Chip* **15**, 135-40 doi:10.1039/c4lc00990h.

52.   van Oijen AM, Blainey PC, Crampton DJ, Richardson CC, Ellenberger T, Xie XS (2003) Single-molecule kinetics of lambda exonuclease reveal base dependence and dynamic disorder. *Science* **301**, 1235-8 doi:10.1126/science.1084387.

53.   Crocker JC, Grier DG (1996) When Like Charges Attract: The Effects of Geometrical Confinement on Long-Range Colloidal Interactions. *Phys Rev Lett* **77**, 1897-900 doi:10.1103/PhysRevLett.77.1897.

54.   Cheezum MK, Walker WF, Guilford WH (2001) Quantitative comparison of algorithms for tracking single fluorescent particles. *Biophys J* **81**, 2378-88 doi:10.1016/S0006-3495(01)75884-5.

55.   Ober RJ, Ram S, Ward ES (2004) Localization accuracy in single-molecule microscopy. *Biophys J* **86**, 1185-200 doi:10.1016/S0006-3495(04)74193-4.

56.   Kulczyk AW, Tanner NA, Loparo JJ, Richardson CC, van Oijen AM (2010) Direct observation of enzymes replicating DNA using a single-molecule DNA stretching assay. *J Vis Exp*, doi:10.3791/1689.

57.   Schermerhorn KM, Tanner N, Kelman Z, Gardner AF (2016) High-temperature single-molecule kinetic analysis of thermophilic archaeal MCM helicases. *Nucleic Acids Res* **44**, 8764-71 doi:10.1093/nar/gkw612.

58.   Hamdan SM, Loparo JJ, Takahashi M, Richardson CC, van Oijen AM (2009) Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis. *Nature* **457**, 336-9 doi:10.1038/nature07512.

59.   Lee JB, Hite RK, Hamdan SM, Xie XS, Richardson CC, van Oijen AM (2006) DNA primase acts as a molecular brake in DNA replication. *Nature* **439**, 621-4 doi:10.1038/nature04317.

60.   Tanner NA, Hamdan SM, Jergic S, Loscha KV, Schaeffer PM, Dixon NE, et al. (2008) Single-molecule studies of fork dynamics in Escherichia coli DNA replication. *Nat Struct Mol Biol* **15**, 170-6 doi:10.1038/nsmb.1381.

61.   Kim S, Blainey PC, Schroeder CM, Xie XS (2007) Multiplexed single-molecule assay for enzymatic activity on flow-stretched DNA. *Nat Methods* **4**, 397-9 doi:10.1038/nmeth1037.

62.   Elshenawy MM, Jergic S, Xu ZQ, Sobhy MA, Takahashi M, Oakley AJ, et al. (2015) Replisome speed determines the efficiency of the Tus-Ter replication termination barrier. *Nature* **525**, 394-8 doi:10.1038/nature14866.

63.   Jergic S, Horan NP, Elshenawy MM, Mason CE, Urathamakul T, Ozawa K, et al. (2013) A direct proofreader-clamp interaction stabilizes the Pol III replicase in the polymerization mode. *EMBO J* **32**, 1322-33 doi:10.1038/emboj.2012.347.

64.   Hamdan SM, Richardson CC (2009) Motors, switches, and contacts in the replisome. *Annu Rev Biochem* **78**, 205-43 doi:10.1146/annurev.biochem.78.072407.103248.

65.   Johnson A, O'Donnell M (2005) Cellular DNA replicases: components and dynamics at the replication fork. *Annu Rev Biochem* **74**, 283-315 doi:10.1146/annurev.biochem.73.011303.073859.

66.   Yao NY, O'Donnell M (2008) Replisome dynamics and use of DNA trombone loops to bypass replication blocks. *Mol Biosyst* **4**, 1075-84 doi:10.1039/b811097b.

67.   Alberts BM, Barry J, Bedinger P, Formosa T, Jongeneel CV, Kreuzer KN (1983) Studies on DNA replication in the bacteriophage T4 in vitro system. *Cold Spring Harb Symp Quant Biol* **47 Pt 2**, 655-68

68.   Cnossen JP, Dulin D, Dekker NH (2014) An optimized software framework for real-time, high-throughput tracking of spherical beads. *Rev Sci Instrum* **85**, 103712 doi:10.1063/1.4898178.

69.   Dulin D, Vilfan ID, Berghuis BA, Hage S, Bamford DH, Poranen MM, et al. (2015) Elongation-Competent Pauses Govern the Fidelity of a Viral RNA-Dependent RNA Polymerase. *Cell Rep*, doi:10.1016/j.celrep.2015.01.031.

70.   De Vlaminck I, Henighan T, van Loenhout MT, Pfeiffer I, Huijts J, Kerssemakers JW, et al. (2011) Highly parallel magnetic tweezers by targeted DNA tethering. *Nano Lett* **11**, 5489-93 doi:10.1021/nl203299e.

71.   Dulin D, Berghuis BA, Depken M, Dekker NH (2015) Untangling reaction pathways through modern approaches to high-throughput single-molecule force-spectroscopy experiments. *Curr Opin Struct Biol* **34**, 116-22 doi:10.1016/j.sbi.2015.08.007.

72.   Kamsma D, Creyghton R, Sitters G, Wuite GJ, Peterman EJ (2016) Tuning the Music: Acoustic Force Spectroscopy (AFS) 2.0. *Methods* **105**, 26-33 doi:10.1016/j.ymeth.2016.05.002.

73.   Sitters G, Kamsma D, Thalhammer G, Ritsch-Marte M, Peterman EJ, Wuite GJ (2015) Acoustic force spectroscopy. *Nat Methods* **12**, 47-50 doi:10.1038/nmeth.3183.
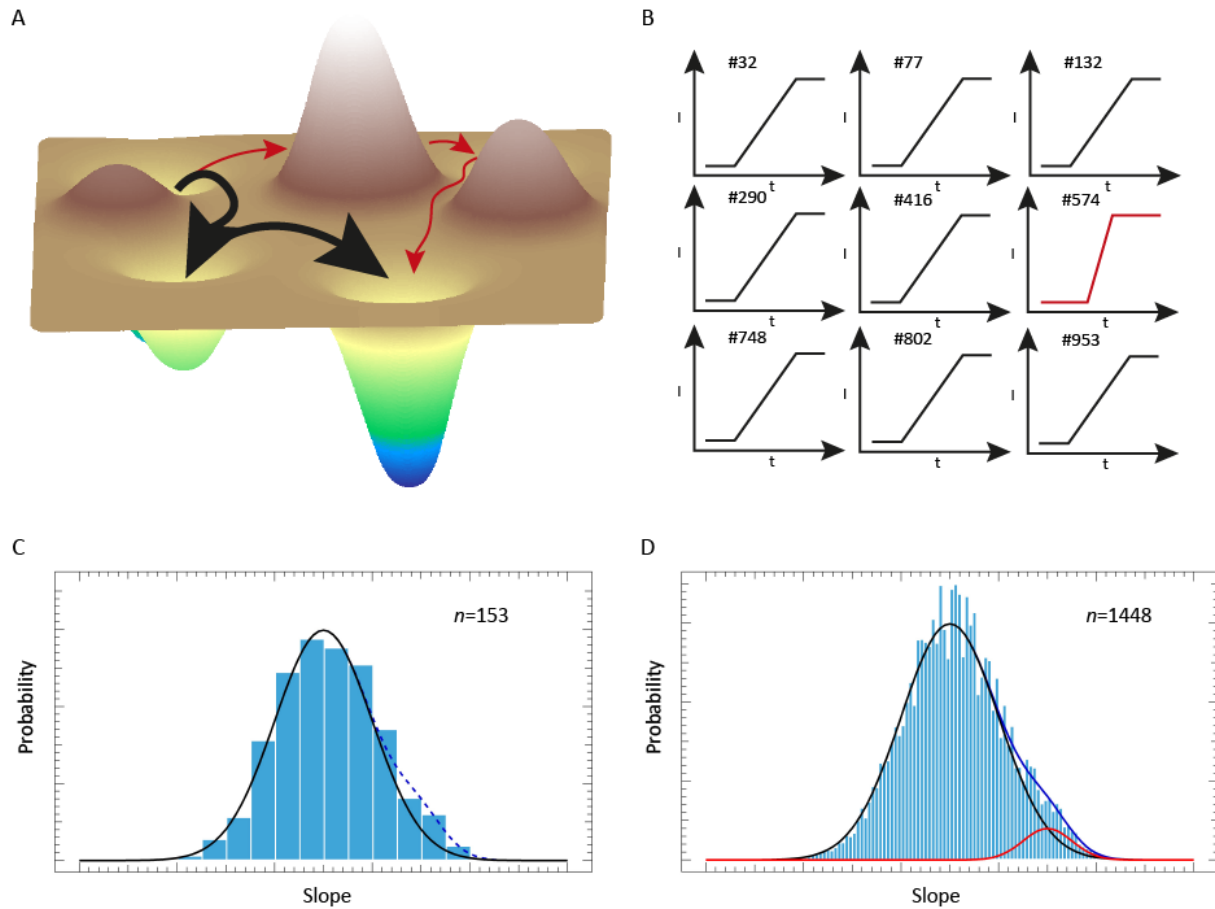
Figure 1: (A) The multi-pathway nature of biochemical processes can be visualised as an energy landscape. While there may be a preferred pathway from one state to another (black), other pathways can also exist (red). (B) Single-molecule methods allow these alternative pathways to be probed. With a large enough throughput, single-molecule trajectories representing alternative pathways (red) are obtained. The kinetic parameters obtained from these trajectories can then be plotted in histograms to visualise their distribution. In a low-throughput experiment (C), alternative pathways may be lost in the noise, whereas high-throughput techniques (D) with their much improved number of observations allow fine-grained binning, which can reveal the presence of small subpopulations.
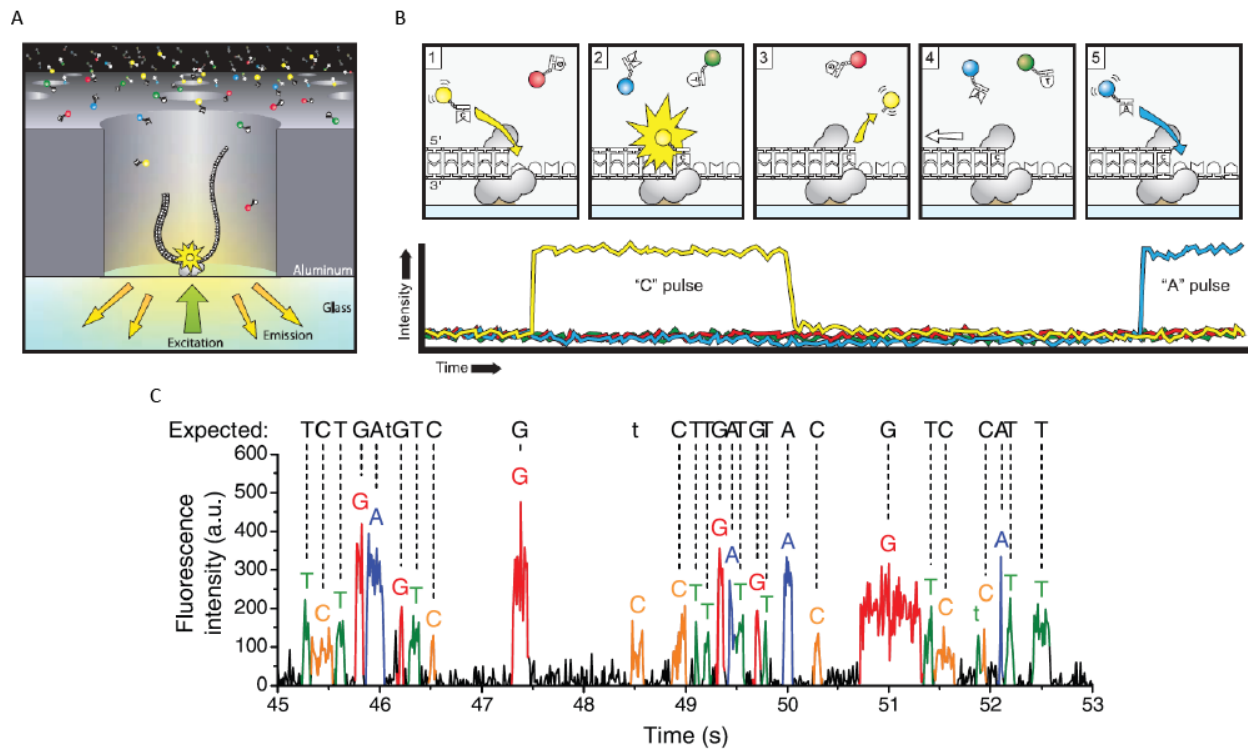
Figure 2: (A) A zero-mode waveguide (ZMW). This nanostructured cavity allows a very high concentration of fluorophores to be used, otherwise creating too much background fluorescence. Fluorophores that remain within the ZMW cavity for an extended period, for example when interacting with a surface-bound molecule, will give a burst of fluorescence well above the background. (B) In the SMRT sequencing reaction, fluorescently-tagged nucleotides give off a burst of fluorescence upon incorporation into the nascent DNA strand by the DNA polymerase that is bound to the bottom surface of the ZMW. From left to right: 1) An incoming cytidine triphosphate has a yellow fluorescent tag. 2) The labelled dCTP remains within the ZMW for an extended period upon capture by the polymerase and incorporation, and a pulse of yellow fluorescence is observed. 3) The next nucleotide to be incorporated is an adenosine, and prior to incorporation no fluorescence is observed. 4) Upon incorporation, a burst of blue fluorescence is associated with the prolonged dwell period of the blue-labelled dATP in the ZMW. (C) A single molecule sequencing trace is obtained, with distinct colours for each of the four nucleotides.
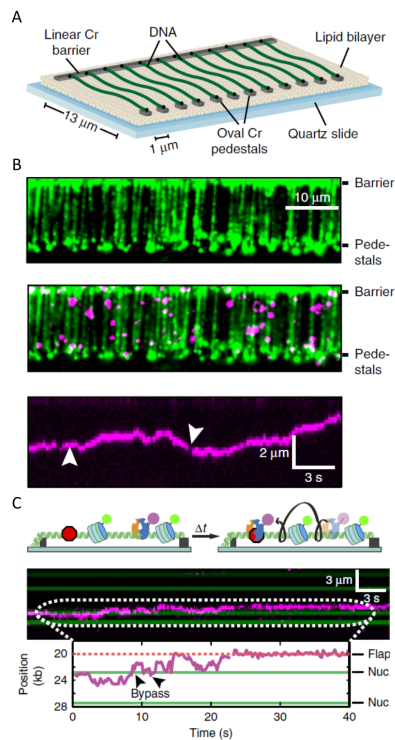
Figure 3: (**A**) DNA curtains are created by anchoring long DNA molecules such as phage λ DNA (48.5 kb) in a lipid bilayer, on top of a lithographically patterned surface. By applying a laminar flow, the DNA becomes aligned into a curtain at the chromium diffusion barrier. In this example, the downstream ends are tethered to chromium pedestals. (**B**) Imaging of doubly-tethered λ DNA curtains stained with the green fluorescent intercalator YOYO-1 (top panel), and with the addition of the quantum-dot conjugated mismatch repair complex Msh2-Msh3 (magenta; middle panel). From the movement of individual labelled complexes, kymographs can be generated (bottom panel). (**C**) Schematic cartoon, kymograph

and single-molecule trajectory of an Msh2-Msh3 complex (magenta) bypassing a nucleosome (green) by hopping, before binding to a 3' ssDNA flap lesion (red). This hopping is captured in the kymograph and its derived single molecule trajectory plot, in which individual hopping events to bypass the nucleosome (green line) are discernible at ~10 seconds, prior to a stable association with the lesion (dashed red line) from ~25 seconds onwards.
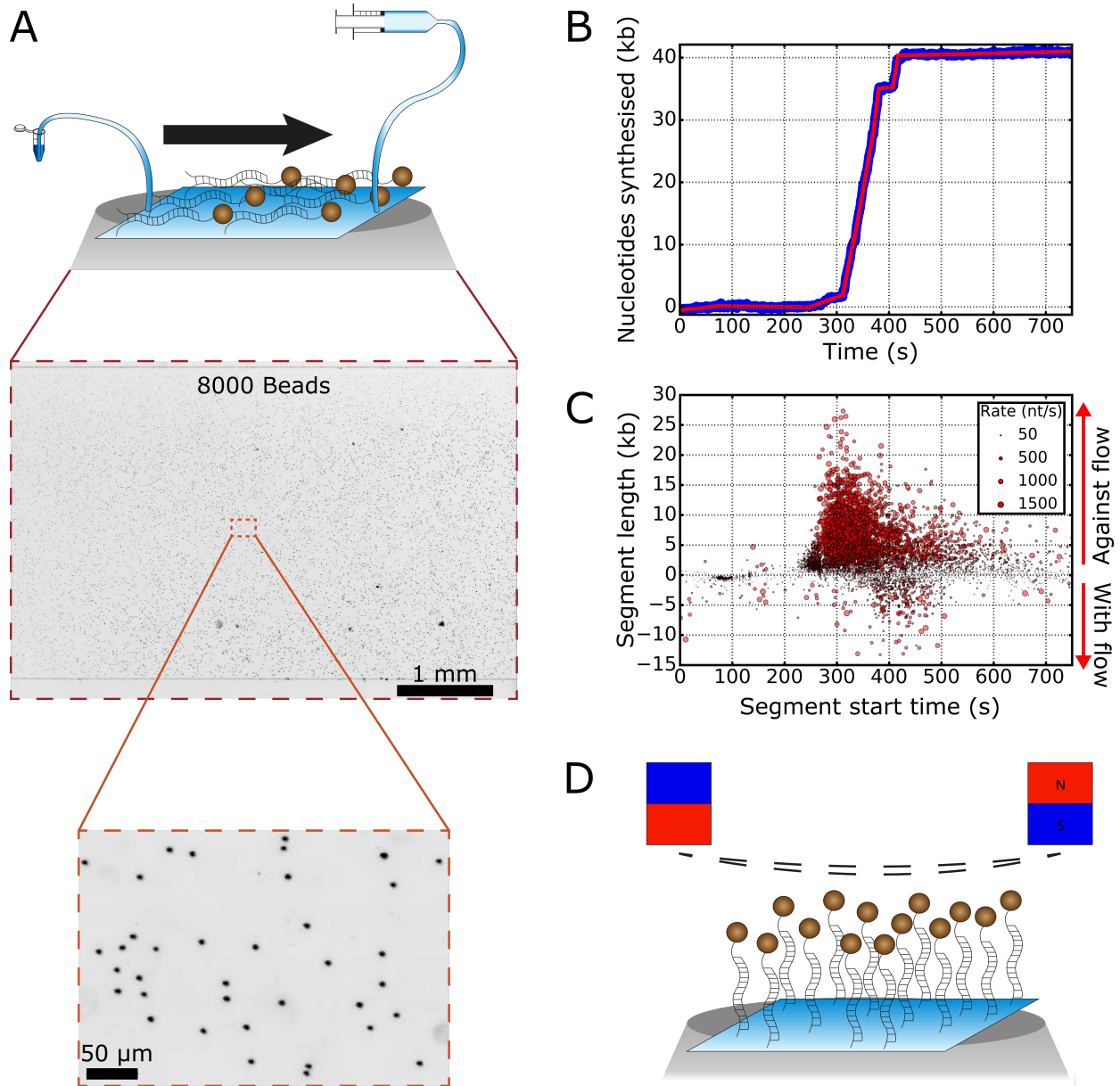
Figure 4

Figure 4: High-throughput force-based single-molecule methods. (A) Flow stretching, showing the very large field of view that can be achieved by coupling a low-magnification telecentric lens with a large, high-resolution image sensor (colour inverted). (B) An example of a single-molecule DNA replication trace obtained in the flow-stretching technique (blue), fit with straight line segments obtained by

change-point analysis (red). (C) A scatter plot depicting all of the change-point line segments from all of the bead trajectories in a single experiment. Each point represents an individual segment in a trajectory, plotted as a function of the segment length in nucleotides, versus the time at which it occurs in the experiment. Enzymes are flowed in at time zero, and reach the flow cell at approximately 250 s, when a large number of segments showing movement of beads against the flow are visible. This type of analysis can be completely automated and helps to separate signal from noise, providing an informative overview of the result of an experiment. (D) Magnetic tweezers using wide-field microscopy, which allows dozens of beads to be imaged simultaneously.