Introduction

Joachim Frank

The term "Machine" invokes familiar images from the macroscopic world, of gears, springs, and levers engaging each other sequentially in a deterministic chain of events. Gravity and inertia are major forces to be reckoned with in the design, or are in fact utilized in the very function, of a device such as a crane, a locomotive, or a centrifuge. The world of molecules this book is concerned with is quite different in many respects. In this water-drenched nanodimension world, mobile parts are in constant jittering motion, powered by random thermal bombardment from the molecules of the aqueous solvent. The forces of gravity and inertia are dwarfed, by orders of magnitude, by those produced by non-covalent interactions, collisions with water molecules, and drag in the solvent. Complicating matters further, biological molecules and the "levers" and "actuators" within them lack the rigidity of materials like steel. Perhaps most difficult to conceive by a mind trained on the experience of the macroscopic world, however, is the disintegration of causal connections into a series of sporadic irreversible chemical events that impart directed motion, separated by stretches of "time" where the molecule has no apparent directionality.

Molecular Machines as a concept existed well before Bruce Alberts' (1998) programmatic essay in the journal *Cell*, but his article certainly helped in popularizing the term, and in firing up the imagination of students and young scientists equipped with new tools that aim to probe and depict the dynamic nature of the events that constitute life at the most fundamental level. "Machine" is useful as a concept because the molecular assemblies in this category share important properties with their macroscopic counterparts, such as processivity, localized interactions, and the fact that they perform work toward making a defined product. The concept stands in sharp contrast to the long-held view of the cell as a sack, or compendium of sacks, in which molecules engage and disengage one another more or less randomly. In invoking mechanistic imagery, it is also, finally, a view that invites physicists to employ their tools and creative minds in developing models that take into account the reality of the nanoworld in a quantitative way. Thus Bruce Albert's is, as invitations go, the third in the past century, with Werner Schroedinger and Richard Feynman having been the first and second to usher their colleagues into tackling the grand challenges posed by biology.

To approach the subject, we must foremost know the structure of the static molecular machine at the atomic level, as a precondition for making sense of its behavior and going beyond mere phenomenological description. X-ray crystallography, a well-matured technique by now, has provided us with an ever-increasing inventory of structures – either complete structures with and without their functional ligands or, if unavailable, structures of isolated components. One of the crowning achievements of X-ray crystallography, the solution of the bacterial ribosome structure (Ban et al., 2000; Schluenzen et al., 2000; Wimberly et al., 2000; Yusupov et al., 2001), has given us unparalleled insights into the architecture of one of the most complex molecular machines in Nature.

Next, we must gather three-dimensional images showing how the structure of the molecular machine changes as a function of performing its work. Such *3-D snapshots* can be obtained by cryogenic electron microscopy (cryo-EM) of samples in which the molecular machine is imaged as an ensemble of single particles in the act of performing their work precisely as they do in solution. (In vitro systems now exist that mimic a variety of fundamental processes in the cell, such as transcription and translation). However, statistical requirements for low-dose three-dimensional imaging, which call for tens if not hundreds of thousands of projections, have the consequence that the snapshots can only be obtained for well-populated conformational states. Thus molecules that are in transition between such states defy attempts at imaging. It is possible, of course, to "trap" or

stabilize the complex in additional short-lived intermediate states by some kind of intervention. Interventions that have been successfully employed in such studies include the addition of small-molecule inhibitors, use of non-hydrolyzable Guanosine triphosphate (GTP) or Adenosine triphosphate (ATP) analogs, transition state analogs, and introduction of targeted mutations. Nevertheless, of a virtual continuum of conformational states, only a relatively small number can be imaged in three dimensions. A movie generated from such a small number of samples would be filled with jarring jumps between frames of successive conformations and binding states.

Fortunately, a new powerful technique that has been developed in the past decade is rapidly filling these knowledge gaps: single-molecule fluorescence resonance energy transfer (smFRET). A donor-acceptor pair of fluorophores, placed strategically on molecular components implicated in dynamic changes, allows monitoring of distance changes in individual molecules in real time, as they perform their work.

In organizing this book, I have made the distinction between "Theory/Methods" and "Results/Biology." The distinction is not sharp because the explanation of methods entails the presentation of illustrative examples. Conversely, the description and interpretation of results is often intertwined with methodological narrative, as each system may require methods to be tweaked or refined in specific ways.

Single-molecule FRET (Chapter 1; Xinghua Shi and Takjep Ha) and Cryo-EM (Chapter 2; my own contribution) are relatively new as complementary techniques for the investigation of molecular machines, and this is why two chapters in the Methods section are devoted to them. A chapter on the theoretical description of molecular machines treated from the point of view of Statistical Mechanics (Chapter 3; Debashish Chowdhury) seemed in order, as well. Furthermore, a computational technique with predictive ambitions, Normal-Mode Analysis, often successfully employed in predicting domain motions, is featured in a contribution by prominent experts (Chapter 4; Karunesh Arora and Charles Brooks III). The degree to which a mechanistic model of a biological system lends itself to accurate predictions of the observed behavior of the system, using the principles of physics, is evidently a gold standard by which the viability of the model can be judged. (Molecular Dynamics simulations, the most prominent technique for predicting dynamics of a molecular machine, is featured in the Results section in its application to the highly complex machinery of translation.)

In the Results part of the book, I have strived to cover those molecular machines that are currently best understood through applications of a battery of biophysical probing and imaging techniques. If a preference is expressed in the selection, then it is in the limitation to globular systems and the exclusion of linear motors, which could be well covered on their own in a separate venture under a defined thematic umbrella.

The contributions in this volume toward the theme of molecular machines follow a protein from its conception as an abstract string of code; over its birth in translation as a polypeptide; its maturation, upon folding, to a fully functioning "adult"; and its function in the cell as part of a complex assembly of fellow proteins. Finally, we see the protein to its final end, as it is dissolves into its amino acid components, the universal building blocks for the next cycle.

In this logical order, we begin with the molecule instrumental for transcription of the genetic information residing on the DNA into a messenger RNA (Chapter 5: RNA polymerase; Finn Werner). The ribosome as the site of protein synthesis serves as an illustration for several lines of multidisciplinary investigation and follows with four contributions: on its dynamics as investigated either by smFRET (Chapter 6: Daniel MacDougall et al.), by cryo-EM (Chapter 7: Xabier Agirrezabala and Mikel Valle), or by molecular dynamics simulations (Chapter 8: James Gumbart et al.), and on its interpretation as a stochastic Brownian machine (Chapter 9: Alexander Spirin and Alexey Finkelstein).

Next we deal with a prominent example for a whole family of complexes involved in protein folding or refolding (Chapter 10: GroEL; Helen Saibil and Arthur Horwich). We have an instructive example for the jobs proteins are made to do once they are assembled into molecular machines performing essential functions in the cell: a protein that packages energy into the currency (ATP) universally used in the cell (Chapter 11: ATP-synthase; Thomas Meier et al.). Having said this, I should note that mature proteins are also incorporated as building blocks and catalytic agents in every single molecular machine discussed in this book. The final chapter concerns the end of a protein's life cycle and deals with the machines involved in protein degradation and recycling (Chapter 12: Andreas Matouschek).

On the whole, I believe the systems selected, while displaying the innate similarities of molecular machines I have outlined at the beginning, are different enough to let the reader appreciate the range of solutions Nature has found to achieve processivity, high fidelity, regulation, and control in a world so far removed from our own macroscopic universe.

ACKNOWLEDGMENTS

I would like to thank Ruben Gonzalez for helpful comments and suggestions. This work was supported by HHMI and NIH R01 GM29169.

REFERENCES

Alberts, B. (1998). The Cell as a Collection of Protein Machines: Preparing the next generation of molecular biologists. *Cell* 92, 291–294.

Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289, 905–920.

Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., et al. (2000). Structure of functionally activated small ribosomal subunit at 3.3 Å resolution. *Cell* 102, 615–623.

Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Jr, Morgan-Warren, R.J., Carter, A.P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. *Nature* 407, 327–339.

Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H.D., and Noller, H.F. (2001). Crystal structure of the ribosome at 5.5 Å resolution. *Science* 292, 883–896.

CHAPTER 1

Single-Molecule FRET: Technique and Applications to the Studies of Molecular Machines

Xinghua Shi
Taekjip Ha

I. INTRODUCTION

I.1 Properties of Molecular Machines

Molecular machines are molecule-based devices, typically on the nanometer scale, that are capable of generating physical motions, for example, translocation, in response to certain inputs from the outside such as a chemical, electrical, or light stimulus. A large number of such sophisticated small devices are found in Nature, including the many biological motors discussed in this chapter, such as helicases and polymerases. These tiny nanomachines work in many ways just like an automobile on the highway, and many consume fuels on a molecular level, for instance, through the hydrolysis of adenosine-5′-triphosphate (ATP) molecules, to power their motions on their tracks. As a result, when lacking the required fuel, these nanomachines tend to slow down and even stop, same as a motor vehicle would. In addition, these biological motors often move in a directional manner with variable speeds, and their processivity characteristics can be described by how far they move on their track of a molecular highway, often formed by a biopolymer such as a nucleic acid or actin filament, before taking off at a later time. Motions of individual components within these protein machines, for example, the ribosome which is discussed in great detail throughout this book, are often nicely coordinated like in any sophisticated, larger-scaled mechanical machines. In recent years, details of the composition, stoichiometry, and three-dimensional arrangement of components within many nanomachines have become available, thanks to the ever-increasing number of high-resolution crystal structures that have been solved, which have provided valuable insights into the mechanisms of how these biological motors accomplish their tasks. In the past two decades, researchers have also brought these machines under scrutiny by a number of novel and powerful methods with ultra-high sensitivity, watching their motions one molecule at a time, and have learned a great deal of previously hidden mechanistic details about their action and dynamics, such as the size of the fundamental steps taken by these motorized nanodevices. In a simplified view of the mechanism of action of biological motors, their strokes of physical translocation are powered by processes such as ATP hydrolysis through a modulation of their conformation, thus converting the chemical energy stored in the molecular fuel, in a stepwise fashion, into directed motions.

This chapter focuses on a particular type of single-molecule method termed Förster (fluorescence) resonance energy transfer (FRET). We start with a discussion of the technique and its developments to date and then feature a number of recent notable studies concerning the application of this technique to the investigation of a series of important molecular machines, including helicases and polymerases.

I.2 Single-Molecule Methods

By definition, single-molecule methods are based on the concept of examining the behavior of individual molecules, one at a time, therefore allowing for a direct detection of molecular heterogeneity, unsynchronizable dynamics, rare but key events, and so on. These approaches have been very successful in uncovering important mechanistic details that are otherwise hidden by ensemble averaging. As the last two decades have witnessed, single-molecule techniques and their applications to biological problems have experienced significant growth. In a survey published in 2007 of single-molecule studies (Cornish and Ha, 2007), the authors found that the number of research articles containing the words "single molecule" in the title, as obtained from a search of the PubMed database, had grown almost exponentially leading up to 2007, with a doubling time of 2.2 years. A rather dramatic prediction from this survey is that all the papers published in the biological sciences in thirty years will contain some single-molecule aspect. Although a bit bold, this prediction illustrates the ever-increasing

importance and popularity of single-molecule methods in modern biological research. Note that the concept of detecting the behavior of individual molecules was not completely new even in the 1980s, as measurements on individual ion channels were already fairly common by then, but the general usefulness of the approach was not widely appreciated until the 1990s. Using single-molecule approaches, many fundamental biological problems that have fascinated scientists for a long time have started to be brought under the microscope. To use a well-known adage, seeing is believing, and many single-molecule techniques allow us to do just that (Selvin and Ha, 2008)! One day, these techniques will become an essential part of laboratory research for routine characterization of many biological phenomena, particularly those involved in the activities of molecular machines.

Among the various single-molecule methods, those based on fluorescence are particularly versatile and useful owing to their superb sensitivity, easy adaptation, and use of common instrumentation (Joo et al., 2008). From the original optical detection of individual biomolecules to the observation of their interactions in real time, and from the conventional diffraction-limited confocal or total internal reflection fluorescence (TIRF) microscopy to the recently developed super-resolution imaging microscopy, the potential of single-molecule fluorescence approaches has risen on many fronts to an unprecedented level and will, with no doubt, advance much further in the years to come (Shi, Lim, and Ha, 2010). In addition to fluorescence, force-based single-molecule tools for manipulation and measurement, such as optical and magnetic tweezers, have provided numerous insights into many important biological problems and can be largely complementary to fluorescence-based methods (Selvin and Ha, 2008). It is conceivable that a combination of different single-molecule approaches will allow researchers to best utilize the merit of each method and, in this way, to explore previously unreachable frontiers in the near future. In the rest of this chapter, we discuss mainly fluorescence-based methods, focusing on single-molecule FRET in particular.

II. SINGLE-MOLECULE FRET

II.1 Principle and Technique

Single-molecule FRET is one of the most, if not the most, general and adaptable single-molecule techniques available today. FRET is a process involving a dipole-dipole interaction between two adjacent fluorescent molecules, through which emission from the lower-energy molecule is observed upon excitation of the higher-energy molecule. As a technique that allows for direct detection of molecular interaction and its dynamics, single-molecule FRET quickly matured after it first went on stage in the mid-1990s (Ha et al., 1996) and has been employed to address fundamental questions about helicase activities (Atkinson et al., 1997; Ha et al., 2002; Myong et al., 2005; Myong et al., 2007; Park et al., 2010), replication (Christian, Romano, and Rueda, 2009; Pandey et al., 2009; Santoso et al., 2010), transcription (Kapanidis et al., 2005; Kapanidis et al., 2006; Margeat et al., 2006), translation (Blanchard et al., 2004a; Blanchard et al., 2004b; Cornish et al., 2008; Cornish et al., 2009; Fei et al., 2008; Fei et al., 2009; Sternberg et al., 2009), reverse transcription (Rothwell et al., 2003; Liu et al., 2007; Abbondanzieri et al., 2008; Liu et al., 2008), and action of DNA-binding proteins (Joo et al., 2006; Roy et al., 2007), to name a few, and the list keeps growing rapidly. It also appears that the practitioners of this technique, originally mostly biophysicists, now encompass members of a much bigger community that includes many biochemists and molecular biologists.

As described by the Förster theory, the rate constant of energy transfer between two fluorophores, termed donor and acceptor, can be expressed as $k_{\mathrm{FRET}} = (R_0/R)^6\, k_{\mathrm{f}}$, where R is the distance between the two fluorophores, R_0 is a parameter often referred to as the Förster distance, and k_{f} is the rate constant of the donor's fluorescence decay in the absence of the acceptor. The value of the Förster distance R_0 is determined by a few key parameters, including the orientation factor that describes the relative orientation of the transition dipoles of the donor and acceptor, the fluorescence quantum yield of the donor in the absence of the acceptor, the overlap integral depicting the degree of spectral overlap between the donor emission and acceptor absorption, and the refractive index of the medium. A more detailed discussion on this topic can be found in a study of energy transfer in yellow fluorescent proteins (Shi et al., 2007). To extract useful information about the distance R between the two fluorophores and its dynamics, one can perform fluorescence lifetime measurements (Edel, Eid, and Meller, 2007; Sorokina et al., 2009) and obtain the FRET rate and, thus, R; however, this would require a pulsed laser source with a picosecond output and sophisticated fast electronics for recording the time-resolved fluorescence signal, both being rare in a single-molecule laboratory, at least for now. As an alternative way to extract the distance information, one can obtain the time-integrated energy transfer efficiency, E, defined as $1/(1 + (R/R_0)^6)$, more conveniently from fluorescence intensity measurements. For a pair of fluorophores such as Cy3-Cy5, FRET efficiency can be approximated by the ratio of the intensity of the acceptor to the sum of intensities of the donor and acceptor, $E_{\mathrm{app}} = I_{\mathrm{A}}/(I_{\mathrm{D}} + I_{\mathrm{A}})$ (Roy, Hohng, and Ha, 2008). It is worth noting that FRET efficiency becomes 50% when the distance R reaches the Förster distance R_0. Because of the sixth-power dependence on distance as shown earlier, FRET efficiency provides an extremely sensitive measure of the relative displacement between the two fluorophore probes in the range close to R_0 (Figure 1.1a) and, more importantly, the dynamics of such distance that is clearly

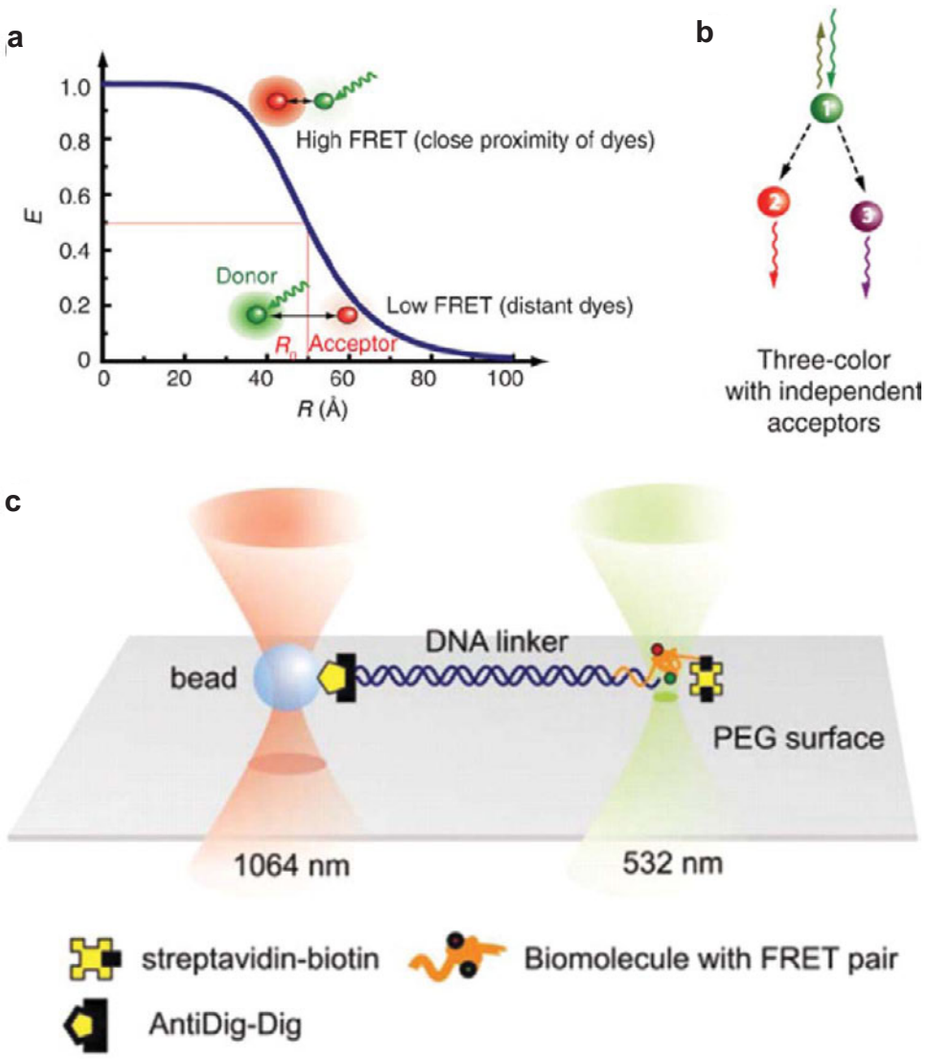

FIGURE 1.1: Single-molecule FRET techniques. (a) Dependence of FRET efficiency on distance R, assuming $R_0 = 50$ Å (from Roy et al., 2008). (b) A two-pair, three-color FRET scheme. 1, 2, and 3 represent the donor, first, and second acceptor fluorophores, respectively (from Roy et al., 2008). (c) A FRET-tweezers hybrid scheme. The two wavelengths, 1,064 and 532 nm, correspond to the lasers for optical trapping and fluorescence detection, respectively. Reproduced with permission from (Hohng et al., 2007).

of interest to researchers studying the motions of biological nanomachines.

As a ratiometric technique, single-molecule FRET eliminates the negative impact of instrument drift, which may not be negligible in some other single-molecule fluorescence measurements, including fluorescence imaging with one-nanometer accuracy (FIONA) (Yildiz et al., 2003) and protein-induced fluorescence enhancement (PIFE) (Myong et al., 2009). The beauty of FRET also lies in its great sensitivity to even a small change in distance in the range close to R_0, typically 3–8 nm depending on the FRET pair used, and this range clearly overlaps with the size of inter- or intra-molecular movement present in many molecular machines. It is noteworthy that, although FRET measurements of individual molecules that are diffusing freely in solution using confocal microscopy are easier to implement, and have been very successful in revealing the presence and distribution of biomolecular subpopulations (Schuler, Lipman, and Eaton, 2002; Rothwell et al., 2003; Kapanidis et al., 2006; Michalet, Weiss, and Jager, 2006), the ability to keep track of a single molecule for an extended period of time from milliseconds to minutes greatly enriches the information content attainable through single-molecule characterization. This improvement was achieved by imaging many surface-immobilized molecules with high throughput (Zhuang et al., 2000; Ha et al., 2002) using TIRF microscopy. Such a single-molecule FRET setup can be assembled by an experienced researcher in

a single day using standard, commercially available optical components that cost as much as a high-end ultracentrifuge (Roy et al., 2008), making the adoption of this technique by most laboratories straightforward. Recent extension of this technique to more than two colors (Hohng, Joo, and Ha, 2004; Clamme and Deniz, 2005) and future incorporation of novel nanotechnologies will definitely add new dimensions of capacity in obtaining useful information about the dynamics of biological nanomachines.

II.2 Comparison to FIONA

Fluorescence imaging with one-nanometer accuracy, abbreviated FIONA, is a super-precision imaging technique for determining the location of a molecule in the laboratory frame from a diffraction-limited image with a precision as high as 1.5 nanometers (Yildiz et al., 2003). This ultrahigh precision is made possible by collecting a large number of photons and curve-fitting the diffraction-limited spot to a Gaussian function, which is similar to finding the absolute peak position of a mountain of defined shape. In comparison, FRET tracks the relative displacement between two interacting objects in the molecular frame with an accuracy on the order of 0.5 nanometer (Cornish and Ha, 2007). Because of their differences in the detection scheme, FIONA would require instrument stability to be at least as good as localization precision while FRET is largely immune to microscope drift. In addition, FRET's sensitivity range depends on the fluorophores selected and is typically 3–8 nm, as previously mentioned. When a motion larger in scale is the subject of interest, for example, the long-distance transport of cargo by a kinesin protein, FIONA becomes more advantageous than FRET. From a user's point of view, the high localization precision of FIONA comes at the price of a much higher photon budget than for FRET (Cornish and Ha, 2007), making it necessary to use fairly bright and photostable fluorescent probes such as quantum dots in the former case. Although FRET is likely a more general method, FIONA can be complementary to it and can be particularly useful when it comes to studies of molecular machines.

II.3 Advances in Single-Molecule FRET

The single-molecule FRET technique discussed thus far has involved only two fluorophores, with which the information about relative displacement can be determined in real time at a high level of accuracy. This scheme has been quite successful in the studies of many molecular machines as discussed later in this chapter, focusing on one particular mode of physical motion. If the biological system of interest involves more components and thus becomes more complex, it is often necessary to acquire additional information simultaneously in order to untangle some of the intriguing riddles. To this end, Ha and coworkers extended the common two-color FRET scheme into one with three colors (Figure 1.1b) by using three spectroscopically distinct cyanine fluorophores, including Cy3 (donor), Cy5 (acceptor 1), and Cy5.5 (acceptor 2), in two parallel energy-transfer pathways (Hohng et al., 2004). Using this multi-color scheme, correlated movements of different arms in a four-way junction structure of DNA (Hohng et al., 2004) and motions of a single-stranded binding protein along DNA (Roy et al., 2009) have been directly observed with little ambiguity. In addition to multi-color detection of fluorescence, a scheme called alternating laser excitation (ALEX) (Kapanidis et al., 2004) has been incorporated into three-color FRET by Weiss and coworkers (Lee et al., 2007), making simultaneous measurements of distances and interactions even more convenient. As a natural extension, a four-color FRET scheme has been realized recently by Hohng and coworkers, by combining multi-color fluorescence detection with pulsed, interleaved excitation. This newest version of multi-color FRET has enabled them to study a complex process of DNA strand exchange mediated by RecA proteins (Lee et al., 2010).

Another, related major technical development is the combination of single-molecule FRET with force-based manipulation and measurement tools (Shroff et al., 2005; Hohng et al., 2007; Tarsa et al., 2007). As demonstrated by Ha and coworkers, the addition of manipulation by force to single-molecule FRET detection (Hohng et al., 2007) made it possible to measure conformational changes of biomolecules through FRET as a function of the force applied by optical tweezers (Figure 1.1c). Using such a FRET-tweezers hybrid, Hohng and colleagues were able to characterize the two-dimensional reaction landscape of the four-way DNA junction structure mentioned earlier. In a related study, Liphardt and coworkers combined FRET with magnetic tweezers to construct a force sensor that provides an optical readout (Shroff et al., 2005). In addition, the importance of mechanical forces in many developmental, physiological, and pathological processes in biology has gained recent recognition and appreciation (Orr et al., 2006). Very recently, Schwartz and coworkers have used the FRET-tweezers duo-spectroscopy for the calibration, with piconewton force sensitivity, of a biosensor based on the vinculin protein that exhibits force-dependence recruitment to cell focal adhesions, which can be used to measure the force across such proteins and examine the regulation of focal adhesion dynamics (Grashoff et al., 2010). We are certain that such a combinatorial approach will be useful in the studies of important biological nanomachines in the near future.

II.4 Advances in Molecule Immobilization and Detection Schemes

As explained earlier, modern single-molecule FRET measurements are more advantageous with surface-immobilized molecules compared to freely diffusing

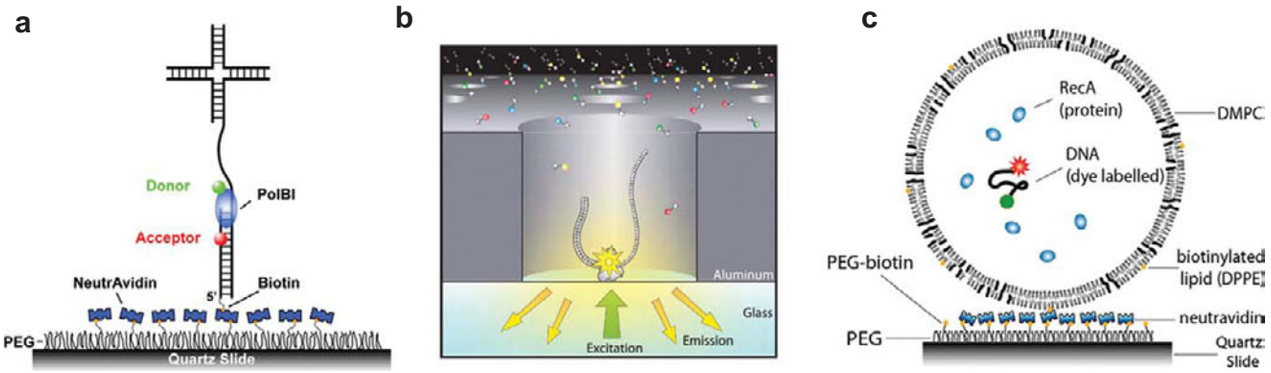

FIGURE 1.2: Advanced molecule immobilization and detection schemes. (a) Surface passivation through PEGylation that allows for specific binding of biomolecules (Shi et al., unpublished results). (b) Illumination based on the zero-mode waveguide (ZMW) nanostructure that gives a reduction of excitation volume down to the zeptoliter (10^{-21} liter) region. Reproduced with permission from (Eid et al., 2009). (c) A porous vesicle containing DNA and proteins that is permeable to small molecules such as ATP at room temperature. Reproduced with permission from (Cisse et al., 2007).

molecules for an extended period of observation on the same molecule. For nucleic acids, immobilization can be achieved fairly easily by using a surface that is passivated by coating with a protein called bovine serum albumin (BSA), a small fraction of which is biotinylated as an anchor for the biotin-streptavidin linkage. For proteins used in single-molecule studies, in contrast, surface coating with the BSA protein was found to be far from adequate in preventing non-specific binding of the protein being studied, and this limitation can be especially detrimental when a fluorescently labeled protein is used, which can easily overwhelm the fluorescence signal contributed by specific interactions. For this reason, a surface-passivation protocol involving the use of polyethylene glycol (PEG) was first adopted in a single-molecule study of DNA helicase unwinding activities (Figure 1.2a) (Ha et al., 2002). Ha and colleagues found that the non-specific binding of Rep helicase was 1,000-fold less on a PEG-coated surface than on a BSA-coated surface, providing a much better reproduction of the enzymatic activity at the single-molecule level. Since then, such a PEG-based passivation procedure has been employed in a large number of single-molecule FRET studies, leading to many interesting discoveries, such as repetitive shuttling of Rep helicase (Myong et al., 2005), spring-loaded DNA unwinding by HCV NS3 helicase (Myong et al., 2007), substrate-directed binding orientations of HIV reverse transcriptase (Abbondanzieri et al., 2008), spontaneous ratcheting of the ribosome (Cornish et al., 2008), and intermediates during nucleosome remodeling by ACF complex (Blosser et al., 2009). As an alternative to PEGylation, a different strategy for preparing an inert surface compatible with single-molecule experiments is passivation by a supported lipid bilayer (Graneli et al., 2006). As has been demonstrated by Greene and coworkers, many molecules can be tethered to fixed locations on such a passivated surface at once, allowing for a parallel recording of the behavior from many individual biological motors simultaneously with a high throughput (Graneli et al., 2006).

With the surface passivated as described earlier and TIRF microscopy, one can carry out many useful single-molecule FRET studies, as long as the concentration of fluorescently labeled biomolecules in the bulk solution is limited to within the low nanomolar range, given the much higher level of fluorescence background that appears above this concentration range. However, many important biological interactions have a much weaker affinity and, thus, are governed by equilibrium dissociation constants in the micromolar range that often necessitate working at micromolar concentrations of fluorescently labeled binding partners in the bulk solution. For these studies, it is necessary to suppress the excitation volume in conventional TIRF or confocal microscopy by several orders of magnitude, as has been achieved successfully using a nanostructure-based device known as the zero-mode waveguide (ZMW) (Levene et al., 2003).

A ZMW is essentially a very tiny hole, tens of nanometers (nm) in diameter, fabricated in a 100-nm thick aluminum film that is deposited on a glass substrate (Figure 1.2b). The small, sub-wavelength size of ZMWs limits the penetration depth of visible light, typically with a wavelength between 400 and 700 nm, to about 30 nm near the bottom of the waveguide. As a result, this nanostructure-based illumination gives a reduction of excitation volume down to the zeptoliter (10^{-21} liter) region; thus, background fluorescence in the presence of an overwhelming pool of fluorescent molecules at micromolar concentrations can be suppressed almost completely. The invention of this technique has facilitated the development of a new-generation DNA sequencing instrument with high throughput and single-molecule resolution (Eid et al., 2009), which is currently being commercialized by Pacific Biosciences, Inc. In a recent study, Puglisi and

coworkers demonstrated the use of ZMWs for observing translation by the ribosome in real time with single-codon resolution using fluorescently labeled aminoacyl-transfer RNA (tRNA) substrates at a physiologically relevant micromolar concentration (Uemura et al., 2010). In the same study, ZMWs were also employed in the measurements of single-molecule FRET between fluorescently labeled tRNA substrates at concentrations up to hundreds of nanomolars, which would be impossible for conventional TIRF microscopy. It is conceivable that such a FRET-ZMW combination will be a very powerful tool in studies of other important biological systems, allowing researchers to characterize weak interactions that are hard to detect otherwise.

In addition to the reduction of observation volume by using a nanostructure such as the ZMW, another promising approach is the encapsulation of single molecules within nanometer-sized compartments such as a lipid vesicle that can be tethered to the surface, which is compatible with conventional, diffraction-limited detection schemes such as TIRF microscopy. Using this strategy, heterogeneous folding pathways of the protein adenylate kinase have been studied in detail (Rhoades, Gussakovsky, and Haran, 2003). One particular feature worth noting in the use of such vesicle encapsulation techniques is that the local concentration of the encapsulated molecule can be sufficiently high for studying transient and weak interactions that are difficult to directly probe by other means.

Although it represents an effective molecule-confinement scheme, vesicle encapsulation is limited by the fact that the lipid bilayer membrane is not permeable to most ions and small molecules, including ATP, in the extra-vesicular solution, making it difficult to probe the effects of these ligands on many interesting dynamic processes. This limitation has been largely ameliorated by taking advantage of the intrinsically porous feature of lipid bilayer membranes near the gel-fluid phase transition temperature. For example, in a proof-of-principle experiment, Ha and coworkers demonstrated that vesicles containing a single-stranded DNA and the recombination protein RecA prepared by using a type of lipid molecule named dimyristoyl phospatidylcholine (DMPC) with a phase transition temperature of 23°C can be made porous at room temperature and permeable to small molecules such as ATP (Cisse et al., 2007). Enabled by the use of porous vesicles, it was possible to modulate the dynamic interactions between RecA and DNA by delivering ATP and its nonhydrolyzable analog across the lipid membrane through the leaky pores, while keeping the larger-sized protein and DNA molecules within the boundary of vesicle nanocontainers (Figure 1.2c). This approach provides a versatile way to supply desired small molecules such as ATP to the inside of the nanocontainer and should be useful for studying the weak interactions in more complex biological systems in a controllable manner.

III. APPLICATIONS

III.1 Applications to Studies of Helicases

Helicases are a class of motor proteins that couple conformational changes driven by ATP binding and hydrolysis to the unwinding of duplex nucleic acids into separate single strands. Like many other molecular machines, helicases move along the molecular track provided by the DNA or RNA substrate and consume the energy stored in ATP molecules. Prior to single-molecule FRET, several other single-molecule techniques including optical tweezers (Bianco et al., 2001) and tethered particle tracking (Dohoney and Gelles, 2001) had been used to study the processive unwinding of DNA by a helicase, *E. coli* RecBCD, with resolutions up to 100 base pairs (bp). However, many helicases exhibit a limited processivity in vitro and often fall off their nucleic acid track long before even reaching 100 bp, making it desirable to use a method with a much higher spatial resolution for studying these enzymes. For this reason, in the early 2000s, Ha and colleagues developed a single-molecule FRET assay for helicase activity that allowed them to measure the unwinding of DNA by a helicase with a resolution better than 10 bp and detect, for the first time, a number of novel helicase events such as unwinding pauses, duplex rewinding, and unwinding restarts, and facilitated a large improvement in our understanding of the underlying mechanisms (Ha et al., 2002).

In this single-molecule FRET assay, two fluorescent probes, Cy3 and Cy5, were attached to the junction between the single-stranded region and the duplex of a partial duplex DNA (Figure 1.3a), of which the unwinding by an *E. coli* helicase called Rep was observed. Because of the short distance between the two reporter fluorophores, the FRET efficiency was almost 100% to start with and decreased as the unwinding helicase proceeded along the duplex, thereby reporting the progress and dynamics of unwinding (Figure 1.3b). Ha and colleagues then observed that, for an 18-bp duplex, unwinding by Rep quickly advanced to completion, whereas for the case of a 40-bp duplex, the situation was much more complex. In the latter case, transient stalls lasting a few seconds were observed in approximately 70% of the unwinding events, followed by either a complete recovery of the initial high FRET, indicative of a duplex rewinding event, or a resumed decrease in FRET, suggesting an unwinding restart event that led to completion of the reaction (Figure 1.3b). Based on the observation that unwinding of DNA by Rep begins only upon the formation of a functional helicase involving more than one Rep monomer (Cheng et al., 2001) and the observation that an unwinding restart event would require free Rep proteins from solution (Figure 1.3c), Ha and colleagues proposed a model in which partial dissociation of the active Rep oligomer complex during unwinding leaves an inactive monomer on the DNA, giving rise to the pauses

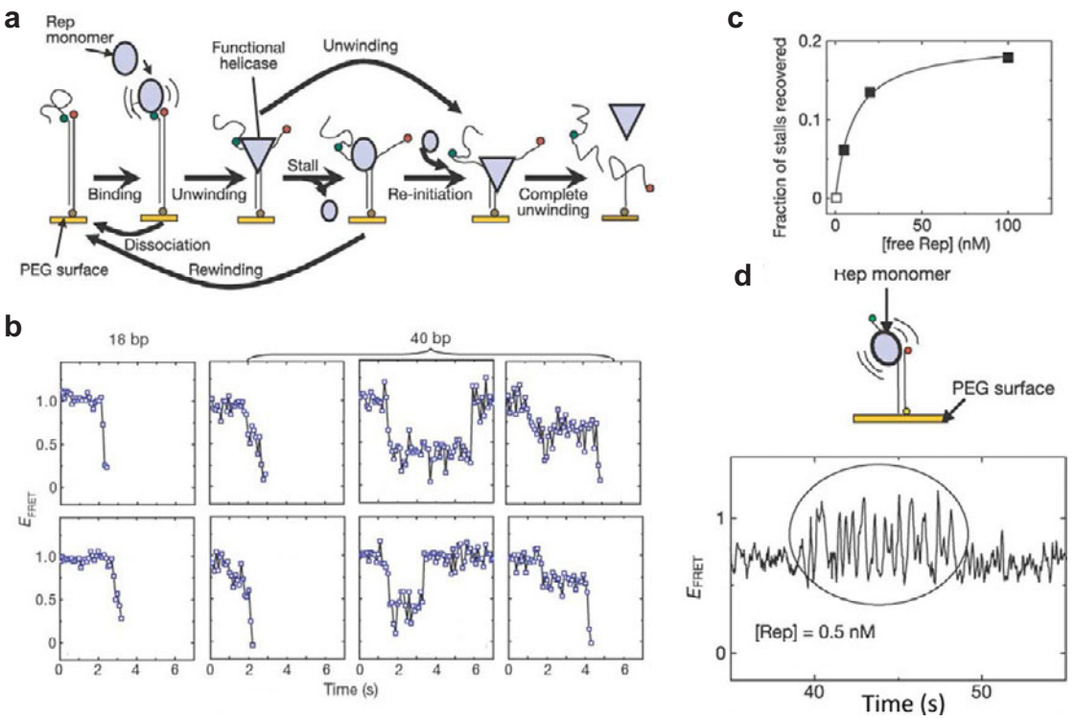

FIGURE 1.3: ***A single-molecule FRET study of*** **E. coli** ***Rep helicase. (a) Experimental scheme and model for Rep unwinding activities. Two fluorescent probes, Cy3 and Cy5, were attached to the partial duplex DNA immobilized on the surface. (b) Typical FRET time trajectories of 18- and 40-duplex unwinding. (c) Dependence of unwinding stalls recovery on Rep concentration, suggesting the requirement of free Rep from solution for unwinding restart. (d) Rapid, ATP-dependent conformational fluctuations of a Rep monomer once in contact with the junction of the partial duplex DNA. Reproduced with permission from (Ha et al., 2002).***

observed (Figure 1.3a). Upon subsequent dissociation of the remaining Rep monomer, the duplex DNA can reform; alternatively, the unwinding can resume upon association of additional Rep monomer(s) to the one stalled on the DNA.

In addition to the mechanism of unwinding of a DNA duplex by the Rep helicase, another interesting discovery made by Ha and colleagues was a rapid, ATP-dependent conformational fluctuation of a Rep monomer, once in contact with the junction of the partial duplex DNA (Figure 1.3d), which did not lead to DNA unwinding. Such a phenomenon became the subject of a follow-up study by Ha and coworkers, as discussed below (Myong et al., 2005). The previously described early study using single-molecule FRET clearly paved the road to answering a series of outstanding mechanistic questions about helicases as molecular machines. For example, in an analogy to the quantification of fuel efficiency of an automobile, how many base pairs are unwound by a helicase per ATP molecule hydrolyzed? What is the fundamental step size involved in the helicase motion? How exactly does a helicase couple the conformational changes induced by ATP binding and hydrolysis to its translocation and DNA unwinding activities? To address these questions more directly, fluorescent labeling of the helicase itself was then employed in single-molecule FRET studies.

Through site-directed mutagenesis, Ha and coworkers achieved site-specific labeling for a series of single-cysteine Rep mutants (Rasnik et al., 2004) and were able to confirm that all of these mutant proteins were fully functional both in vivo, using a plaque assay for bacteriophage replication (Scott et al., 1977; Cheng et al., 2002), and in vitro, as judged by single- and multiple-turnover DNA unwinding activities. As described earlier, PEG-based surface passivation was found to essentially eliminate nonspecific binding of labeled proteins to the imaging surface, allowing the detection of Rep binding to an immobilized partial duplex DNA, as visualized by an abrupt appearance of fluorescence signal in the corresponding channel. With a near-quantitative (above 90%) labeling efficiency, Ha and coworkers showed that at a sub-nanomolar protein concentration, Rep binds to DNA primarily in the monomeric form, based on counting the number of photobleaching steps observed, which could be distinguished clearly from those more complex but rare multiple-protein binding events. For each of the eight single-cysteine Rep mutants, Rasnik and colleagues studied the binding interaction between Rep and DNA in the absence of ATP by