EXPLORATION OF HELICASE MECHANISMS AT SINGLE MOLECULE LEVEL: ANSWERING DECADES OLD QUESTIONS AND BIOENGINEERING SUPER PROTEINS

BY

SINAN ARSLAN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Professor Paul R Selvin, Chair
Professor Taekjip Ha, Director of Research
Professor Karin A Dahmen
Associate Professor Brian L DeMarco

Abstract

In the last 15 years, single molecule techniques have allowed us to observe the transient, heterogeneous, and multi-step behavior of biomolecules often averaged out and lost in ensemble assays. Helicases form a ubiquitous class of enzymes that function in many aspects of nucleic acid metabolism which is central to life. We used single molecule Förster resonance energy transfer (FRET) and optical tweezers force spectroscopy to study translocation and unwinding mechanism of *Escherichia coli* Rep, a model helicase with 3'→5' motor translocation activity on single-stranded DNA (ssDNA) fueled by ATP hydrolysis. When a translocating Rep on a 3'-overhang DNA reaches the DNA duplex junction, instead of dissociating, it snaps back to the 3' end, restarting another shuttling cycle. We investigated the mechanism of repetitive shuttling and discovered the same behavior on various DNA substrates. We concluded that the repetitions are induced by the lack of ssDNA track ahead of the Rep. Using the repetitive shuttling assay, we explored the translocation mechanism of Rep in detail. First, we perturbed the ssDNA binding network via biochemical mutations that led to determination of the key residues that control shuttling speed, ATPase activity, and directionality. Second, we tested the effects of the DNA lesions on Rep translocation, observing that the irregularities encompassing 1-3 nucleotides in the backbone and nucleobases only caused transient stalls. To probe the unwinding mechanism, we developed a conformational control assay which turned wild type Rep monomers with no detectable DNA unwinding activity into super-helicases (Rep-X) via internal covalent crosslinking. Rep-X can unwind thousands of base pairs processively even against large forces. We also showed that partner proteins of a similar enzyme turn on the unwinding activity by stabilizing the active form. Lastly, the orientation dependence of FRET is experimentally shown between nucleic acid conjugated cyanine fluorophores that constitutes the first demonstration of this effect in a biophysical system since its formulation by Theodor Förster in 1948.

To Mother, Father, Sister and Elif

Acknowledgements

I would like to thank my advisor Professor Taekjip Ha for the opportunity to explore the biophysics of life's building blocks. I find him a very supportive mentor and great scientific visionary. He gave me ample freedom for scientific exploration. I am honored to be an alumnus of such a great scientist. I also would like to thank Graduate Executive Heads John Stack and Lance Cooper for their constant support, my PhD committee members, Paul Selvin, Karin Dahmen and Brian DeMarco for their help at the end of this journey, and my preliminary committee board, Nigel Goldenfeld, Benjamin Lev, and the late Bob Clegg for endorsing my candidacy for PhD. I am grateful for their help and support I received from following people. My dear friend Burak Okumus was first to teach me how to perform single molecule fluorescence experiments. I had a chance to be part of his wonderful work and publications. I could never forget the most kind and patient tutorship of Sua Myong who introduced me to the exciting world of helicase proteins and how to work with them. I am grateful for Salman Syed's initial help and continuous scientific and friendly fellowship. Chirlmin Joo is a great person and scientist that I will always remember as a role model. Michelle Nahas, Rahul Roy, Tae Young Yoon and Ben Leslie always took their time to help me. Hamza Balci, Rustem Khafizov and Yuji Ishitsuka were great colleagues with whom I worked on wonderful projects, learned a lot from and published together. Our scientific and friendly discussions with Kaushik Ragunathan, Prakrit Jena, Ibrahim Cisse and others were always a pleasure. Maria Sorokina, Jeehae Park, Jiajie Diao, Cheng Liu, Reza Vafabakhsh, Ruobo Zhou, Mike Brenner, Kyung Suk Lee, Eli Rothenberg, Peter Cornish, Gwanrong Lee, Michael Schlierf, Xinghua Shi, Matthew Comstock, Seongjin Park, Jaya Yodh, Sultan Doganay, Hajin Kim, Thuy Ngo, Ankur Jain, and those that I forget to name here were wonderful colleagues with great scientific prowess. I also want to express my appreciation and thanks to our collaborators, amazing David Lilley, Timothy Lohman, Shu-ou Shan, Christopher D. Thomas, and their students, Kuang Shen and Asif Iqbal. Finally, I thank all my friends and family for their support throughout my PhD journey.

Table of Contents

Chapter 1. Introduction to single molecule biophysics and helicases

In order to understand biological phenomena, and how life works, there is a great need to look at their main components, such as proteins, DNA, RNA, etc., in detail. A living cell is the least understood complex system, despite the great scientific and technological advances of our age. What is hidden from our sight is tremendous compared to what we know, because we are yet to probe the mechanisms, functions and interactions of many biomolecules, mainly the proteins and the complex, adaptive and intelligent systems they form. These macromolecules often exhibit conformational changes, transient interactions, complex behaviors that can be easily missed in ensemble and bulk experiments performed in tubes. In contrast, single molecule biophysics has enabled us to look at these phenomena in detail within the last 15 years, filling the information gap and promising advancements, such as ability to monitor and manipulate biomolecules *in vitro*, *in situ* and *in vivo*.

In this chapter, I will first introduce the single molecules techniques I have used in my research, namely the single molecule fluorescence (or Förster) resonance energy transfer (FRET), total internal reflection fluorescence (TIRF) microscopy and optical tweezers force spectroscopy, their capabilities, and the advantages they provide over bulk methods. Secondly, I will present an introduction of the main biomolecular motor protein I have focused during my PhD research, a model superfamily 1 helicase, *Escherichia coli* Rep, and the aspects of it that we studied; the helicase enzyme that unwinds double-stranded DNA (dsDNA) and the translocase that tracks single-stranded DNA (ssDNA). Lastly, I will give an overview of how we bioengineered super-active helicases by conformational control.

1.1 Single molecule vs ensemble methods

Ensemble techniques are essential and valuable tools. However, single molecule techniques allow us to distinguish subpopulations of different dynamics and behaviors that are often averaged out in ensemble measurements, resulting in over/underestimated parameters of biomolecules. A good example of this is how single molecule measurements revealed that PcrA helicase, a homolog of Rep, takes 1 nucleotide (nt) steps along the ssDNA during translocation as indicated by structural biology studies. In contrast, the ensemble kinetic studies, such as

stopped-flow techniques predicted that the kinetic step size of PcrA like helicases 3-6 nt. This long standing debate of kinetic step size of helicase proteins for the ssDNA translocation and dsDNA unwinding was resolved by measuring the period of shuttling for an individual helicase molecule on a certain length of ssDNA track (40-100 nt) several hundred times (*1*). Gamma distribution fit of the periods of this multistep Poissonian event from individual PcrA molecules revealed that the numbers of rate limiting steps are equal to the number of nucleotides in the ssDNA track. This result effectively proved single nucleotide step size for this class of helicases. It also showed, as it was shown over and over in many other single molecule studies of proteins, that these macromolecules are inherently heterogeneous in their dynamics, kinetics, reaction rates, etc. The PcrA study concluded that ensemble averaging of heterogeneous quantities can cause overestimated step sizes, which explains the results of ensemble methods.

1.2 Single molecule fluorescence resonance energy transfer (smFRET) microscopy

FRET is a non-radiative energy transfer phenomenon between two fluorophore molecules, a donor-acceptor pair. After the donor is excited by a photon, for example by laser source, electronic energy state is converted from a singlet ground state (S_o) to an excited electronic singlet state (S_1). S_o and S_1 states are comprised of multiple of vibrational energy levels which determines the excitation and emission spectra of the fluorophore. The absorption occurs very quickly and on the order of 10^{-15} s, if no other interacting fluorophores, such as an acceptor molecule present, the donor dissipates the energy and return to S_o (after a fluorescence lifetime of $\tau = 10^{-9}$-10^{-7} s) usually via thermal relaxations and a photon emission. Energy lost in the internal heat conversion leads to a red-shift in the emission spectrum compared to the absorption spectrum of the molecule (Stokes shift ref2). There are also non-radiative relaxation pathways that limit the number of photons generated per excitation photon, a ratio of which is defined as quantum yield (Φ) of the fluorophore. A good fluorophore for single molecule FRET experiments should have a relatively high Φ.

FRET occurs as another non-radiative energy transfer pathway for the donor in the presence of a nearby acceptor fluorophore. Dipole-dipole coupling between S_1 states of the donor to the acceptor allows energy to be transferred to the acceptor S_1 state, resulting in acceptor emission, which is further red-shifted. This interaction is first characterized by T.

Förster in 1948, as the dipole-dipole interaction of the S_1 electronic states between the fluorophore pair. He formulated that the FRET efficiency, E_{FRET}, i.e. the ratio of the acceptor emission due to FRET to total emission intensities of donor and accepter molecules,

$$E_{\mathrm{FRET}} = \frac{I_{Acceptor}}{I_{Donor} + I_{Acceptor}}$$

has a distance dependence in the following form (*2*):

$$E_{\mathrm{FRET}} = \frac{1}{1 + \left(R / R_0 \right)^6}$$

where the R_o is the Forster radius at which point the E_{FRET} is 0.5, and it is specific to the FRET pair (Figure 1.1). R_o is usually in the 2-9 nm range, for example R_o for Cy3-Cy5 pair, two well-known cyanine fluorophores, is about 5-6 nm. Hence the E_{FRET} is sensitive to the changes in R within 3-8 nm range which provides a suitable nano-ruler that can measure the distance and size in many biomolecular reactions and structures, such as DNA-protein interactions.

One of the features of the FRET is its dependence on the relative orientation of the donor and the acceptor, and it is one of the parameters in R_o (see Chapter 4, for the background theory). In most experimental cases, fast rotational thermal fluctuations average out this dependence. However the local environment of the conjugated fluorophores can restrict the rotational fluctuations of both fluorophores. In such cases, E_{FRET} can deviate from the widely used fluctuation averaged values, solely due to fixed orientation. In Chapter 4, we demonstrate this effect on DNA conjugated Cy3-Cy5 pair.

Single molecule fluorescence studies require reduced the background to obtain good signal to noise ratio from single molecule emissions. This is accomplished by only exciting a small volume of the sample, using techniques such as total internal reflection (TIR). In TIR, the excitation laser beam is reflected from the interface between glass/quartz and aqueous medium, creating an approximately 100-200 nm thick evanescent excitation field in the aqueous medium. Thus only the molecules tethered to the solid surface or trapped in this excitation layer are excited, and the background is dramatically minimized. There are basically two experimental designs for TIR fluorescence microscopy, an objective based or prism based TIR microscope (Figure 1.2, Figure 1.3). The idea of using total internal reflection to illuminate cells contacting

the surface of glass was first described by Ambrose in 1956 (*3*). It was then extended by Daniel Axelrod in early 1980s (*4*).

1.3 Single molecule force spectroscopy, optical tweezers

smFRET is a powerful tool, but it has limitations, for example the scales outside 2 -8 nm are inaccessible. The optical tweezers on the other hand can access to ranges from 1 Å to many microns, at high time resolutions in sub-millisecond (*5*). In addition to high spatiotemporal resolution, biomolecules can be studied under applied force approximately in the range of 1-250 pN which is advantageous compared to AFMs because of the low force regimes of biomolecules that AFMs cannot access. The optical tweezers utilize momentum of tightly focused laser light to trap micron-sized dielectric beads. A polystyrene bead placed at the focal point will be in a 3-dimesional potential well (*6*) which act as a Hookean spring that can used to manipulate macromolecules, such as DNA, protein, viruses, etc. Some of the examples are studies of nucleic acids structures such as hairpins, ribozymes and G-quadruplex (*7, 8*), and various DNA-binding proteins such as RecA (*9-11*), SSBs (*12*) and others (*13*). We used it to monitor the DNA unwinding of Rep, PcrA and their bioengineered versions in Chapter 3.

1.4 Introduction to Helicases

Helicases derive their name from the catalytic ability to separate energetically stable double helix structure of DNA into two complementary single stranded components. They form a large class of motor proteins that utilize the energy from hydrolysis of nucleotide triphosphates (NTP), to unwind double stranded (ds) nucleic acids (DNA, RNA) (*14-16*) and to translocate on ssDNA processively (*17*). These enzymes function in all aspects of nucleic acid metabolism including replication, repair, recombination, chromatin remodeling, removal of proteins from DNA and RNA, propagation of Holliday junctions, bacterial conjugation, transcription termination, and others (*14, 15, 18-28*). Helicases are abundant in all living organisms, and comprise the 1-2% of the whole genomic composition.

Understanding the helicase mechanism is essential to characterize nucleic acid metabolism in cell. They are also an important class of enzymes for medical research. Many diseases are attributed to mutations in helicases, such as Werner's syndrome, Xeroderma

pigmentosum (XP), Bloom's syndrome, Cockayne's syndrome, tricothiodystrophym, thalassemia, mental retardation, premature aging and increased propensity to cancer (*29-34*). These biomedical consequences of genetic helicase deficiencies and defects in cancer and hereditary diseases, as well as being potential drug targets have brought great attention to helicase studies.

1.5 Unwinding and translocase function of helicases

Some helicases can unwind DNA with very high speeds that is in range of progression speed of replication fork (500-1000 bp/s) (*35*). However, many non-hexameric helicases have slower unwinding rates (30-300 bp/s). These helicases also have very short unwinding processivity. Hence it has been proposed that the translocase activity of these motor proteins can be the main function in vivo, since they cannot effectively unwind DNA on their own. Translocation of a helicase, as in the unwinding activity, is coupled to binding and hydrolysis of NTP, which is mainly ATP, and subsequent release of by products, NDP and inorganic phosphate (Pi) (*14, 36-39*).

Most DNA helicases have specific binding orientation to ssDNA, which results in polarized translocation upon ATP hydrolysis. Even though crystal structures have provided great insight, single molecule experiments can shed considerable light on detailed aspects of translocation mechanism and how it relates to unwinding of nucleic acid duplexes.

Some helicases form a hexameric ring around the DNA, such as replicative helicases, DnaB or MCM, whereas others work as monomers or homo/hetero oligomers, for example, bacterial helicases, Rep, PcrA and UvrD (*14, 37, 38, 40, 41*).

1.6 Helicase superfamilies

Helicases are grouped in different ways, for example, according to the conserved sequential motifs, directionality, substrate type, etc. Main classes of helicases are called superfamilies, namely, superfamily 1, 2, 3 , 4 and 5 , or in short, SF1, SF2, SF3, SF4 and SF5 according to the aminoacid sequence and conserved folds based on Gorbelenya's work (*42*). Gorbelenya's method defines seven conserved sequence motifs that are the basis of the two largest superfamilies, SF1 and SF2. Figure 1.4 shows the location of these motifs on a SF1 helicase, Rep from *E. coli*. Main structural feature of these two superfamilies is the two RecA like

folds that act as the motor domains (RecA is a ubiquitous homologous recombination protein in *E. coli*). Outer surface of these motor domains contain the ssDNA binding motifs. The cleft formed between them has the motifs that bind and hydrolyze nucleotide triphosphates, such as ATP. These ATP binding motifs are famously referred as Walker motifs. Among the examples of SF1 helicases are *E. coli* Rep, UvrD, RecB, RecD, *B. stearothermophilus* PcrA, and some viral examples such as T4 Dda, HSV-1 UL5. Famous examples of SF2 helicases *are E. coli* PriA, human RecQ helicases such as BLM and WRN, eIF-4A, and NS3 helicase from Hepatitis C virus. SF3 helicases are only identified by three conserved motifs. SF4 and SF5 are rather small families of hexameric helicases. They form a hexameric ring around the ssDNA/RNA during their activity. SF4 hexameric helicases contain five conserved motifs and carry a DNA primase domain. SF5 hexameric helicases have structures homologous to F-1 ATPase (*42*).

1.7 *E. coli* Rep, a superfamily 1 DNA helicase

The *E. coli* Rep has 673 amino acid long single-chain protein with 76.4 kD size (*43-45*). There are about 50 copies per *E. coli* cell (*46*). Rep is not an essential protein for E. coli, genotypes of *rep* knockout exhibit slow replication fork progression coupled with increased number of replication stalls (*19, 45, 47, 48*). Rep has been shown to interact with DnaB, the main replicative helicase of *E. coli*, and seemed to facilitate uninterrupted progression of replication fork (*49-51*). In a similar work it was also shown that where DnaB helicase is unable to dislodge a downstream Lac repressor, Rep can easily bypass the Lac repressor(*52*). These observations may explain the frequent pauses at the replication fork in *E. coli rep* mutants when the replication fork encounters protein-DNA complexes. It has been proposed that Rep facilitates the progression of the replication fork by dislodging DNA-bound proteins (*53, 54*).

Rep is a model superfamily 1 (SF1) helicase with a 3’ to 5’ ssDNA translocation activity. This behavior is shared by UvrD and PcrA which are also SF1 helicases (*55-60*). Rep has an extensive structure (90%) and sequence homology (40%) with *E. coli* UvrD and *B. stearothermophilus* PcrA (*61-63*). It functions in restarting the stalled replication processes and unwinding of several phages such as *ϕX*174, M13, P2, etc. (*19, 64-66*).

1.8 Unwinding activity of Rep helicase

Rep, PcrA and UvrD are poor DNA unwinders; a detectable activity occurring only at high concentrations (> 10 nM) *in vitro*. Furthermore, none of these homologous helicases possess

unwinding processivity above 200-300 bp in even higher saturated concentrations, and they do not exhibit any detectable helicase activity in monomeric form *in vitro* (*40, 55, 59, 67*). They require association with other partner proteins to turn into processive helicase units. The mechanism of this activation has not been answered. Chapter 3 explains how we unraveled this elusive question.

A crystal structure of *E. coli* Rep (*61*) in two conformations with DNA substrates are shown in **Figure 1.5**. These structures reveal a flexible swiveling domain, named 2B, and RecA like motor domains, 1A and 2A, and another relatively small stationary domain, 1B. Rep homologs *E. coli* UvrD (*68, 69*), and B. stearothermophilus PcrA (*62, 63*) have been crystallized in similar conformations as well. These two extreme conformations of 2B subdomain are named open and closed forms. The position and orientation of other three subdomains (1A, 1B and 2A) remain essentially unchanged in the open and closed conformations. A mapping between the open and closed conformations of Rep, based on the positions of the α-carbons of 1A, 1B, and 2A domains, results in almost a perfect match with an RMSD of 3.6 Å. Degree of swiveling around the hinge region that connects the 2B subdomain to the 2A subdomain ranges between 130° and 160° in the x-ray structures for Rep, UvrD and PcrA.

In order to study the biological role of 2B subdomain, a 2B deletion mutant, RepΔ2B was made and tested (*70*). RepΔ2B was still a functional helicase, supporting *ϕX*174 replication *in vivo*, and retaining helicase activity *in vitro*. Pre-steady state ensemble unwinding reactions revealed that RepΔ2B is a faster unwinder than the wild type Rep, but with shorter processivity (*70, 71*). We also confirmed that RepΔ2B is a faster translocase as well. This result suggested that the 2B subdomain could be a protein-protein interaction domain with a regulatory role, but how that is accomplished has not been answered since the structural discovery in 1997. We used the open and closed form structures of 2B subdomain as a basis to bioengineer two stabilized conformers using biochemical crosslinking agents, and characterized their helicase activities for Rep and PcrA (Chapter 3) which answered this long-standing question for SF1 helicases.

Crystal structures of PcrA and UvrD bound to a dsDNA with a short 3' ssDNA tail both show the 2B subdomain in the closed conformation; making contacts with the duplex DNA which led to various proposals about the significance of the 2B subdomain-duplex DNA interactions in DNA unwinding (*63, 68*), even though it does not contain any of the conserved

helicase motifs and shows significant variations in both size and sequence among SF1 and SF2 superfamilies (*16*). Our previous study of FRET triangulation showed that Rep monomer binds to the DNA with a short 3'-overhang DNA in the closed conformation (*72*), similar to the PcrA and UvrD observed in the crystal structures in complex with 3'-overhang DNA (*63, 68*). 2B subdomain of PcrA and UvrD monomers seemed to anchor the protein to the dsDNA with a 5'-overhang. The anchored protein then translocates in a 3' to 5' direction along the ssDNA tail so that the enzyme reels in the tail; forming a ssDNA loop (*1, 73*). Overall, these observations were interpreted as the 2B subdomain having a regulatory role, rather than having an active functional role in DNA unwinding. For example, it has been suggested that the 2B subdomain interaction with the dsDNA may function to prevent wild type Rep monomers from unwinding or otherwise invading dsDNA, i.e. an auto-inhibitory role (*17, 55*). Then it was speculated that the 2B-induced unwinding inhibition and regulation may be lifted by the interaction with a cellular partner protein *in vivo*, while *in vitro* studies suggested that the protein-protein interaction between two or more Rep monomers is enough to suppress the proposed auto-inhibitory effects of the 2B subdomain-dsDNA interaction. A detailed understanding of the 2B subdomain-duplex DNA interactions during the ATP hydrolysis cycle would facilitate an understanding of the role of these interactions. Systematic crystal structure studies, such as those reported for PcrA and UvrD complexes with DNA (*63, 68*) are not available for the Rep helicase, making fluorescence based assays that can probe the structure of Rep-pdDNA complexes particularly important. The study we report in Chapter 2.6 was designed to determine the relative conformation of a Rep monomer in complex with a partially duplexed DNA (pdDNA) at different intermediates of ATP hydrolysis using a FRET based triangulation assay. The crystal structure studies of Rep-ssDNA complexes (*61*) and our previous FRET studies on a Rep-pdDNA complex without ATP (*72*) served as a basis for this study.

1.9 Translocase activity of Rep helicase

ssDNA/ssRNA dependent translocase activity is integral feature of the many helicases. A completely new type of translocation behavior for Rep helicase, namely the repetitive shuttling was discovered by our group. Rep translocates on the ssDNA track which is characterized by the ability of the Rep to rebind to the beginning of the ssDNA track upon reaching the end of the track, on DNA constructs such as 3' tail of a pdDNA, surface tethered ssDNA and stalled

replication fork analogue (*58*). Other studies revealed that PcrA (*1*) and UvrD (unpublished) exhibits similar repetitious behavior.

As it will be described in Chapter 2.1, Rep can translocate repetitively on large variety of DNA substrates without dissociation. Based on our observation and its low copy number in cells, it is plausible that an important role of Rep *in vivo* is ssDNA translocation, which was proposed to keep ssDNA free of unwanted proteins. Repetitive shuttling can facilitate this role with enhanced efficiency. These substrates provided a more comprehensive understanding of how robust and widespread the repetitive shuttling phenomenon can be.

Rep, like PcrA, UvrD and many other helicases bind the ssDNA in a specific polarity. N-terminal motor domain (1A) is located on the 3'-end side and the C-terminal motor domain (2A) faces the 5'-end. The crystal structures revealed the important residues that make contact with ssDNA via hydrogen bonding and π-stacking (Figure 1.6). These contacting residues mainly reside in motifs Ia, III, IVa, and V, and certain residues outside these motifs but conserved among smaller subset of SF1 helicases. Thr56 and Asn57 of motif Ia, Tyr248 of motif III, Arg350, Gly351 and Asn352 of motif IVa and, His558 and Thr556 of motif V seem to interact with phosphate backbone of ssDNA whereas Trp250 of motif III and highly conserved non-motif residue Phe183 (the "gatekeeper" residue) are involved in π-stacking interaction with the nucleobases. Certain residues also make hydrogen bonds with the nucleobases as well, such as Arg251 of motif III, His85 of TxGx motif, and non-motif residues, His580 and Ser582. We have investigated how these residues are involved translocation mechanism by site directed mutagenesis and testing the mutants using the repetitive shuttling assay. The mutations revealed a whole spectrum of effects, from fast translocations to slow ones, or completely abolished translocation and/or DNA affinity and/or ATPase activity. These results are presented in Chapter 2.4.

1.10 Bioengineered versions of Rep and PcrA helicases

As mentioned above Rep, PcrA and UvrD helicases present two distinct conformations named as "closed" and "open" forms in their x-ray crystal structures. These two conformations are result of 130°-160° swiveling rotation of the 2B subdomain relative to the fairly rigid body of the protein. Single molecule fluorescence and ensemble biochemical studies show that opening

and closing dynamics of 2B subdomain is coupled to the translocation, ATP hydrolysis cycle, and presence of duplex DNA, suggesting that the movement of the 2B subdomain might be functionally relevant in regulatory capacity. For example 2B subdomain might facilitate the dimerization of the Rep protein (*61*). Interestingly the 2B subdomain does not contain any helicase motifs. Dispensability of this domain in RepΔ2B studies make its role even harder to pinpoint. Despite all these findings, what 2B subdomain does still remains an open question today.

In order to determine the functional and/or regularity role importance of the 2B subdomain, we have successfully crosslinked 2B and 1B subdomains to arrest the protein in a perpetual "closed" form. We performed this first by replacing all the native cysteines and introducing two new cysteines at specific positions on 2B and 1B subdomains respectively. In one double cysteine mutant, these residues very closely face each other, when the protein is in the closed form. Upon covalent crosslinking of these two cysteines with a short homobifunctional bismaleimide crosslinker, the protein is arrested in "closed" form. We bioengineered the closed form crosslinked version of Rep and PcrA, which we named Rep-X and PcrA-X. Among the most amazing results, these modifications converted the monomers with no detectable unwinding activity into super-active monomeric helicase units that can unwind 6-kbp DNA in a single reaction. This result is important because it proves that the closed form is a functional form of these helicases, and it also showcases first time a super active enzyme can be engineered by conformational control of protein structure. In contrast, Rep arrested in the open form (Rep-Y) was similar to the wild type Rep in its helicase characteristics. This work is presented in Chapter 3.

1.11 Figures

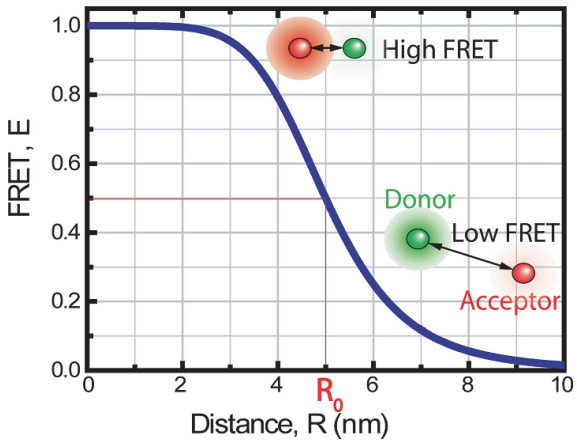


Figure 1.1 FRET efficiency-distance dependence. E_{FRET} plot is shown for a donor and acceptor pair (green and red balls respectively). Shading represents the relative emission intensity. At the Forster radius (*Ro* = 5 nm), E_{FRET} is at 50%. The plot also shows that E_{FRET} most sensitive to the changes in the range of 3-8 nm.

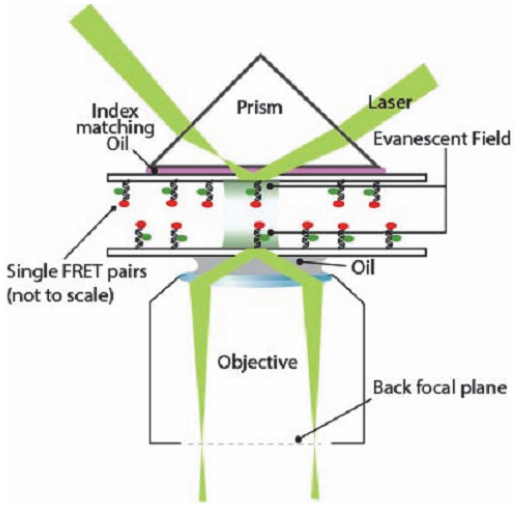


Figure 1.2 Types of total internal reflection fluorescence (TIRF) microscopy. Diagram of the objective and prism type TIRF designs show how the excitation laser beam (green) with an incident angle greater than the critical angle can be reflected from inner surfaces of the slide/coverslip to form an evanescent field layer in the aqueous sample chamber (green gradient), hence eliminating background by only exciting a ~200 nm layer in the chamber. Sample molecules with a FRET pairs are tethered to the slide/coverslip surface. Objective type TIR is usually obtained by focusing the laser light at the back focal plane whereas in the prism type TIR, the laser is focused approximately at the reflection surface.

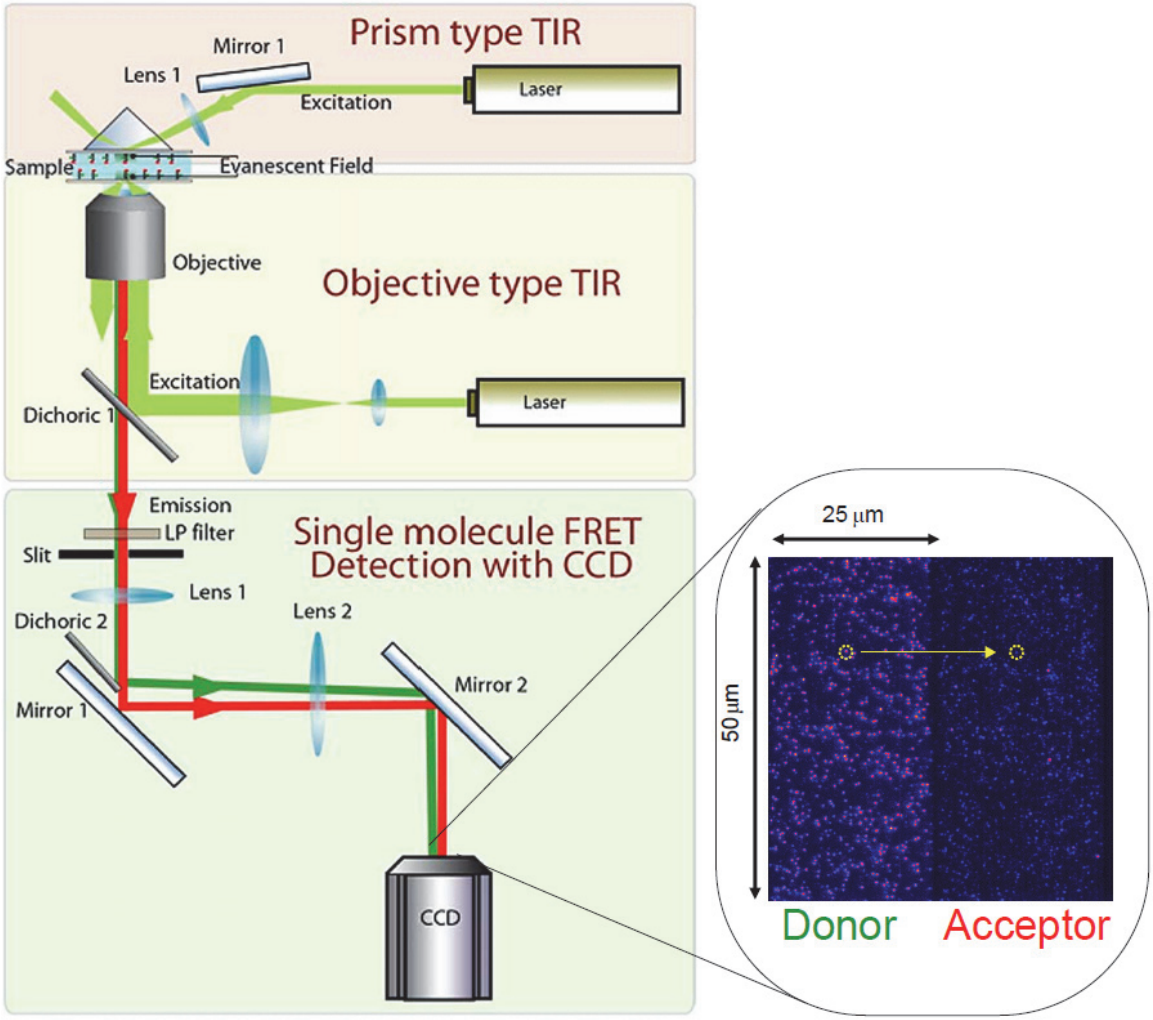


Figure 1.3 Single molecule total internal reflection (TIR) fluorescence setup. Top and middle panels show the emission pathways for the prism and objective type TIR designs, respectively In objective type, laser light is expanded, and reflected by dichroic 1 mirror with a cut off wavelength slightly above that of the laser. Middle and bottom panel shows the emission pathway for donor and acceptor signal (dark green and red) collected by the objective lens, first cleaned by a long pass filter to reject scattered excitation laser and a dichroic 2 mirror to separate the donor and acceptor emissions for side by side detection at the CCD camera. A sample recorded image shows both the channels with a molecule marked in both channels

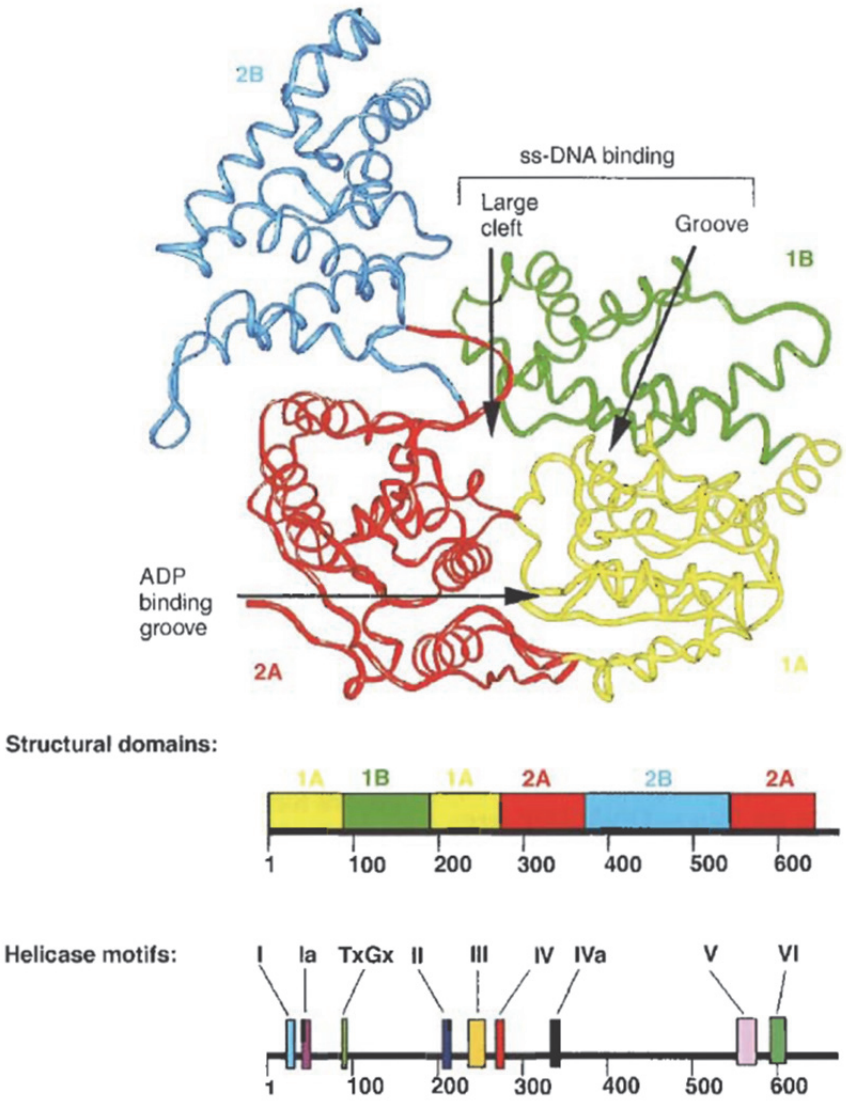


Figure 1.4 Conserved helicase motifs. Four subdomains of Rep are color coded in the open form crystal structure and color coordinated with the top sequence map (PDB ID: 1UAA). Relative positions of conserved helicase motifs are shown in the bottom sequence map. While the ATP binds to the cleft between the 1A (yellow) and 2A (red) motor subdomains, ssDNA binds to the top surface of the same subdomains. (Reprinted from Lohman 2004 with permission)

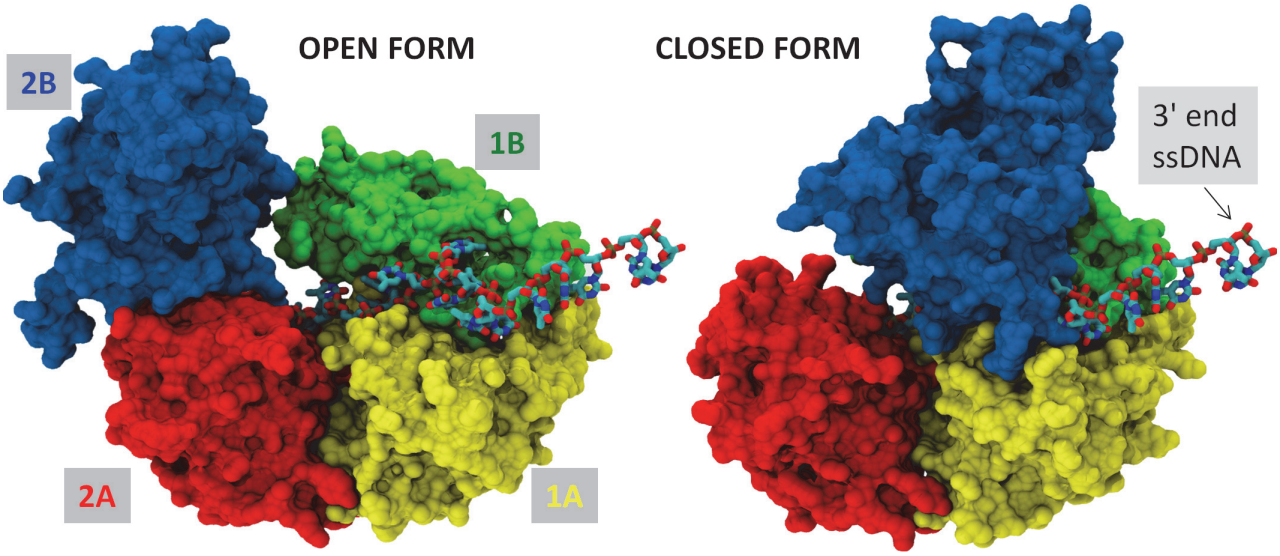


Figure 1.5 Open and Closed form of Rep helicase. Crystal structures of the open and the closed conformations (surface representation) are aligned on the same section of the ssDNA (multicolored licorice). Subdomains colored as in Figure 1.4 (PDB ID: 1UAA). The terms open and closed refer to the relative conformation of the dynamic 2B subdomain (blue).

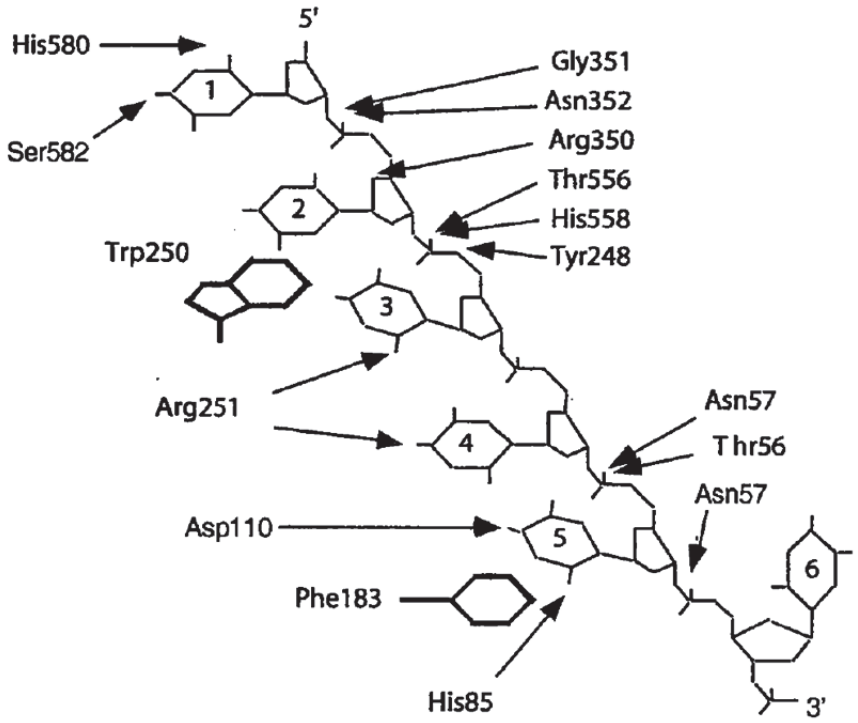


Figure 1.6 ssDNA interaction network of Rep helicase. Some of the residues of Rep that interacts with ssDNA backbone and nucleobases are shown. (Reprinted from Korolev, Hsieh et al. 1997 with permission)

Chapter 2. Translocation mechanism of Rep helicase

2.1 Repetitive shuttling mechanism

A novel phenomenon for Rep helicase, namely the repetitive shuttling on ssDNA had been discovered in our laboratory (*58*). In that study, Rep translocations repeated successively on a partial duplex DNA (pdDNA) with a free 3' overhang (Figure from Sua's paper). Discovery hinged on a curious behavior Rep exhibited once it reached the end of the ssDNA track, i.e. the junction of duplex DNA where it could not move further. At that stage, Rep helicase did not dissociate from the DNA, but somehow jumped to the 3' end of the track, and repeated the translocation cycle multiple times. Myong *et. al.* showed that during the jumps Rep co-binds the 3' end of the DNA, transiently forming a loop of ssDNA, and then releasing the junction to completely transfer to the 3' end followed by subsequent translocation cycle. These jumps were called snapbacks. Myong *et. al.* showed that Rep does the same on two other substrates, ssDNA tethered to surface at 5' end, and stalled replication fork. Only on the 3' end tethered ssNA, they did not observe repetitions. Based on such persistent shuttling activity, the ability to prevent binding of recA, and prior reports of no apparent unwinding ability of monomers led Myong *et. al.* to propose an *in vivo* translocase function for Rep monomers. So they concluded that the repetitive shuttling keeps the exposed ssDNA free from unwanted toxic effects of certain proteins.

Myong *et. al.* also hypothesized that the repetitions were blockade induced, since the entire set of DNA structure they tested contained a blockade at the downstream of translocation direction. This was supported by an internal conformational change that was coupled to the translocation. The translocation toward the blockade was accompanied with a gradual closing of swiveling 2B subdomain and sudden opening upon snap-back to the start position. Hence it was concluded that in the blockade induced closed state, Rep has increased affinity to the starting end of the ssDNA track which enables the snap-backs. Following each snap-back Rep repeated the same cycle until full dissociation from the DNA which happened after anywhere from 1 to 50 cycles of translocation on 80-nt long track.

In order to study this phenomenon, we tested different DNA substrates, and found that Rep can also repetitively translocate on variety of other DNA structures, such as pdDNA with a 5' overhang, free ssDNA, and ssDNA bound by dsDNA at both ends. It should be noted that two of these structures have no end blockades to induce repetitions, but simply a 5' end of a ssDNA track. As for the intra-molecular conformational change, we observed the same gradual closing of the 2B subdomain as Rep approaches the 5' end of the ssDNA track. Based on these results, we propose a more comprehensive repetition model that shortening of ssDNA ahead of the Rep induces a conformational change that results in snapback to the beginning of the track.

Although all the DNA substrates that showed repetitive translocation have a free 3' end at the snapback site as a common feature, we did not find that 3'-OH was essential for the snapbacks. All of these structures, except for the stalled replication fork, contained a ssDNA track with a free 3' and/or 5' end. To see if this was a requirement for the repetitive shuttling, we looked at the Rep behavior on a stretch of ssDNA flanked by duplex DNA at both ends. Rep exhibited the repetitive shuttling on this substrate as well, hence indicating that a ssDNA with a free end is not required the repetitive shuttling.

Based on the diverse DNA structures Rep can repetitively shuttle on and its low copy number in cells, it is plausible that one of the main roles of Rep *in vivo* could indeed be repetitive ssDNA translocation.

We have monitored Rep activity on these various DNA substrates via molecule Forster Resonance Energy Transfer (smFRET), protein induced fluorescence enhancement (PIFE) and time correlated single photon counting (TCSPC) on TIRF and confocal microscope stages (Chapter 1: Introduction to single molecule biophysics and helicases). Preparation and assembly of reaction chambers, fluorophore labeled DNA substrates Rep samples, detailed buffer conditions, and the details of the experimental assay to reproduce the results are described in 2.1.8.

2.1.1 Rep translocation on 5' overhang DNA

First, we designed an assay in which a single Cys mutant of Rep protein was labeled with a donor dye at the N-terminal domain (residue 43 or 97 with Cy3) and the acceptor dye (Cy5) was located on the junction of the pdDNA with 80-nt 5' poly-dT overhang as shown in Figure 2.1a. The biotinylated pdDNA was surface tethered to a polymer-coated (to prevent non-specific

protein adsorption) quartz slide via neutravidin. Upon injection of reaction buffer containing Rep (100 - 500 pM) and ATP (1 mM) into the reaction chamber, binding of a labeled Rep molecule to the surface-tethered DNA was observed as a sudden appearance of fluorescence signal. FRET efficiency exhibited a series of gradual decrease phases followed by sudden jumps in repetitive manner (Figure 2.1a), indicating Rep's directional movement away from the junction and a sudden return to the junction. A histogram of translocation periods, Δt, measured between adjacent snapback events, showed an average translocation speed of ~80 nt/s, similar to values obtained by Myong *et. al.* To see if sudden return to the junction is a snapback in the manner Myong *et. al.* discovered, we monitored the FRET efficiency between acceptor labeled Rep (Cy5 at C-terminal domain residue 333) and the donor (Cy3) labeled 5' end (Figure **2.1**b, cartoon). As a result, FRET efficiency pattern was reversed, i.e. cycles of gradual increase and sudden drop, suggesting 3' to 5' translocation on ssDNA without staying bound to the 5' end or the duplex DNA followed by snapbacks in the same manner as it did on pdDNA with 3' overhang.

Figure 2.1c represents an otherwise identical experiment but with an unlabeled Rep on the 5' tailed pdDNA with Cy3 at the 5' end. During the translocation, we observed a series of sawtooth shaped Cy3 intensity changes resulting from Cy3 emission enhancement induced by proximity of Rep (PIFE) (*74*). The sawtooth pattern showed a gradual increase and an abrupt decrease of intensity, consistent with the Rep translocation cycles observed above. As another indicator of protein movement, we obtained the single molecule trajectories of fluorescence lifetime of Cy3 during the same experiment (Figure 2.1d) using a time-correlated single photon counting (TCSPC) assay on a confocal microscope stage (2.1.8). The fluorescence lifetime of Cy3 showed repetitive cycles of increased gradual increase from about 1.2 ns to 1.8 ns and abrupt drops to 1.2 ns, paralleling the sawtooth pattern observed in the emission intensity in Figure 2.1c.

A histogram of translocation periods are plotted for all four assays, showing similar Δt_{ave} values (average 0.87 s, standard deviation 0.11 s). Slight differences can be attributed to the room temperature variation (T_{room}=22±2^{o}C), and to the inherent biochemical of single Cys mutants. Here we showed that, first time, upon encountering the free 5' end of the ssDNA track, without any physical blockade, Rep does not dissociate from the DNA, but snaps back to the

junction and repeats the translocation cycle multiple times. Our finding indicates a different mechanism of repetitive shuttling than previously conceived blockade induced mechanism.

2.1.2 Repetitive shuttling on free ssDNA

We have shown that replacing streptavidin anchored 3'-end of a ssDNA that has a free 5' end with a duplex DNA junction recovered the snapbacks that was found to be lacking (*58*). Now that we discovered free 5' end of a ssDNA can induce snap backs depending on the target site of the snapback, we wanted to see how Rep behaved on a free ssDNA oligonucleotide that has no duplex DNA junction or a protein block at either ends. For TIRF microscopy imaging of Rep behavior on free ssDNA, we trapped pairs of ssDNA and Rep in small unilamellar vesicles (SUV) of 200 nm diameter that were tethered to surface as shown in Figure 2.2 (*75*). Phospholipid SUVs exhibit spontaneous porosity at the phase transition temperature that allows passage of small molecules such as Mg^{2+} ions and ATP necessary for Rep translocation, so we used DMPC (1,2-Dimyristoyl-*sn*-Glycero-3-Phosphocholine) lipid as its phase transition temperature (23°C) is suitable working at room temperature. Donor labeled Rep (Cy3 at residue 333) and the acceptor labeled 98-nt ssDNA (5'-Cy5) were co-encapsulated in vesicles as described in the 2.2.3.4: Rep Helicase Encapsulation. The vesicles were tethered to surface by biotinylated lipid moieties that were added to DMPC. Upon introduction of ATP into the reaction chamber, we observed the repetitive shuttling signature, i.e. the cycles of gradual FRET efficiency increase and abrupt decrease, similar to those observed on pdDNA tethered directly to the surface, with similar translocation speed. This data suggest that Rep is able to translocate repetitively on free ssDNA, so it is capable of snapbacks from a free 5' end to a free 3' end. However, vesicle encapsulation studies alone could not rule out the possibility that the repetition may simply be due to rebinding of Rep to the same DNA after its complete dissociation (since effective concentration of a single Rep molecule in a 200 nm wide vesicle is around 400 nM and binding rate of Rep to a ~20-nt ssDNA is around 4 10^{-2} $nM^{-1}s^{-1}$ (*40*), rebinding would likely to occur within 10 ms whereas our integration time was 16-50 ms). Therefore, we performed a complementary experiment by surface-immobilizing the biotinylated Rep protein (2.1.8, Figure 2.3). 98-nt long ssDNA labeled with Cy3 at the 5' end was then added along with ATP. When a free ssDNA molecule bound to Rep, the protein exhibited repetitive translocation on the ssDNA as observed through PIFE of Cy3 signal (Figure 2.3). Average translocation periods are in agreement with the ones obtained on 3'and 5' overhang pdDNA

substrates. These experiments confirm the ability of Rep to snapback from the 5’ end to the 3’ end of the same ssDNA without dissociating. Behavior of Rep on the 5’overhang pdDNA and the free ssDNA supports a mechanism of repetitive shuttling that is facilitated by lack of ssDNA ahead of the helicase. It is still curious that when 3’ end of a ssDNA is blocked by streptavidin it inhibits the snapbacks yet a duplexed DNA at that end does not.

2.1.3 Repetitive shuttling on ssDNA flanked by duplex DNA at both ends

One of the most common feature of all the DNA substrates on which Rep was able to do repetitive shuttling had been a free ssNA end (3’ or 5’) except for the stalled replication fork analog where one end contained a ssDNA/duplex DNA junction (Okazaki fragment) and the other end with a triple junction of ssDNA and two duplex DNA stems (template and the leading strand) (*58*). In order to expand on this, we designed another substrate that is biologically relevant but did not have any free ssNA ends. It was a stretch of ssDNA is flanked by dsDNA at both ends (ds-ss-ds DNA) (Figure 2.4), basically a similar substrate to the stalled replication fork but without the upstream leading strand duplex. Acceptor label (Cy5) was located at the end of the translocation track that was a 40-nt long poly-dT sequence. smFRET traces of donor labeled Rep (Cy3 at residue 333) were recorded demonstrated that Rep was able to do the same repetitive shuttling on the ds-ss-ds DNA as the other substrates, further supporting a snapback mechanism that does not require a free ssDNA.

2.1.4 Free in-trans ssDNA stops Rep translocation on surface tethered DNA

Mechanism of repetitive shuttling hinges on the snap back of Rep to the 3’ end via transient looping of ssDNA. A good question would be whether a ssDNA provided in trans could eliminate shuttling from the surface-immobilized DNA. If our model of looping via 3' end recognition is correct, then addition of free ssDNA at very high concentration should stop the Rep translocation on the tethered DNA. Free ssDNA in the solution provides additional 3’ end for snap back that competes against the 3’ end of the surface-tethered DNA, subsequently causing Rep to leave the surface. To test this hypothesis, we flew in 100 μM of ssDNA of dT_{45} while the Rep-Cy3 molecules were translocating on the Cy5 labeled partial duplex DNA with dT80 3’

overhang (Figure 2.5). Introduction of high concentration of ssDNA resulted in fast dissociation of Rep molecules from the surface tethered DNA. A histogram of the dissociation FRET efficiencies showed that helicase molecules predominantly left the surface during the snap back. This result suggests that Rep recognizes the 3' end when it reaches the end of the ssDNA track, then it snaps back to the 3' end region, and repeats the shuttling process.

2.1.5 Conformational dynamics of Rep during the repetitive shuttling on 5' overhang DNA

Dual labeled Rep on a 3' overhang pdDNA showed a repeating pattern of conformational change coupled to the repetitive shuttling in Myong *et. al.*'s work. During the translocation flexible 2B subdomain was gradually closing and abruptly opening with a periodicity similar to that of translocation cycles. Hence it was proposed as the structural mechanism of snapback that the protein assumed the closed form, a secondary binding site for the snapback target site on the protein was exposed to facilitate transition of Rep from one site to another site on the DNA. We wanted to see if the same conformational change occurred during translocation toward a free 5' end with no blockade.

Based on the two conformations in the crystal structures of Rep in complex with ssDNA, the 2B subdomain of Rep was proposed to be flexible and to swivel 130° between the close and open forms (*61*). We used a Rep mutant with two cysteines (residues 473 and 97) that are stochastically labeled with a mixture of Cy3 and Cy5 (*58*). Residues 473 and 97 are located at the tips of 2B and 1B subdomains respectively such that the closed conformation should display significantly higher FRET than the open conformation. Single molecule time traces of this doubly labeled Rep moving on an unlabeled partial duplex DNA with a 5' ssDNA tail displayed repetitive pattern of a gradual FRET increase followed by an abrupt FRET decrease (Figure 2.6), suggesting that the 2B subdomain is primarily in the open conformation when it is near the junction with the duplex DNA, and becomes closed as it translocates towards the 5' end of the ssDNA. In the previous work (*58*), similar cycles of gradual closing and sudden opening of 2B subdomain were observed as the enzyme was moving on a 3' ssDNA tail towards the duplex junction. Combining the two experiments, it is evident that the conformational state of Rep is mainly governed by the ssDNA track ahead rather than its contact with junction or a form of blockade. In a way, Rep "senses" the length of the ssDNA ahead of it.

This conformational change was originally observed on substrates with blockade at the end of translocation tracks. Here we saw the same behavior as Rep moves on ssDNA track with free 5' end. In all of these cases, what naturally followed upon closing of 2B subdomain was snapback. Hence the closed conformation of Rep could be facilitating the snap backs, perhaps by exposing a secondary binding site on the protein for the translocation start site, i.e the snapback target site, as proposed by Myong *et. al.*

2.1.6 3'-OH is not an essential factor in repetitive shuttling

Repetitive shuttling was observed in all types of DNA substrates tested in this study and in the previous study with one exception. On a ssDNA with a free 5' end and its 3' end linked to the surface via biotin-streptavidin, only one cycle of translation was observed with no repetition (*58*). We previously interpreted the result as evidence for a requirement of a physical blockade at the 5' end of the track for snapback to occur. However, in light of the new findings here that showed repetitive shuttling from DNA substrates with free 5' end, it appears that it is the lack of ssDNA track ahead of the translocase that signals for snapback, not a bulky blockade. What blocks the snapback then appears to be the lack of accessible 3' end which is blocked by its linkage to streptavidin via biotin. In all other substrates, the 3' end is accessible, either as a single-stranded end or as a recessed 3' end. If Rep indeed contains a secondary binding site for a 3' end, it may recognize the 3'-OH group. A binding pocket for 3'-OH was described for another *E. coli* helicase that functions in stalled replication restart, priA (*76*). Some exonucleases also recognize the target oligonucleotides through the 3'-OH group (*77*). If Rep indeed contains a similar secondary binding site for a 3' end, it may recognize the 3'-OH group.

In order to test if the 3'-OH is required for snapback, we modified the end of the 3' overhang of a pdDNA with a phosphate group (Figure 2.7). We observed that Rep was still able to translocate repetitively on this modified substrate. The number of successive repetitions before the dissociation (or photobleaching) was similar to that of unmodified substrate with (Figure 2.7). We concluded that a 3'-OH is not essential to the snapback phenomenon.

2.1.7 Conclusion and discussion

We tested various DNA structures for Rep helicase translocation behavior, especially for the gaining more insight into the repetitive shuttling characteristic. We have found that Rep can repetitively shuttle on the 5' overhang of a pdDNA, free untethered ssDNA and ds-ss-ds DNA.

DNA with 5’ overhang and free ssDNA showed that snapback, the key event for the repetitions, can occur at 5’ end of a ssDNA track, as well as at the blockades in form of a duplex DNA or a protein block (such as Neutravidin). Thus, we propose that repetitions are induced by lack of ssDNA ahead of the protein rather than the specific structure at the end of the ssDNA translocation track. Repetitive shuttling on ds-ss-ds DNA substrate supported this idea, along with all the other substrates, by showing that Rep can transition from any end of ssDNA track to a snapback target site, whether it is a free 3’ end of a ssDNA, duplex junction with a receded 3’ end, or stalled replication fork junction with 3’ end of the leading strand. The only substrate on which the repetitions were inhibited was a ssDNA track tethered to the surface at the biotinylated 3’ end (snapback target site) (*58*). In comparison, 5’ tailed pdDNA replaces the neutravidin-biotin end of the inhibited substrate with a duplex DNA (with recessed 3’ end) and it recovers the repetitions. Hence it seemed that as long as 3’ end is present at the snapback site, Rep was able to transition when it encounters end of ssDNA track. Thus it became evident that the 3’ end containing DNA structure at the target snapback site governs the transition of Rep to the beginning of translocation track. However interaction of the Rep with the 3’ DNA end at the snapback site was not specifically facilitated by the 3’-OH group, since replacing it with a phosphate group did not have any effect.

Dual labeled Rep experiments showed that the synchronous conformational change of Rep in the form of gradual closing of 2B subdomain occurred on both the 3’ and the 5’ overhang DNA. This suggests that the conformational change is caused by the lack of ssDNA track ahead rather than what lies at the end of the track. At the end of each translocation cycle, prior to snapback, 2B subdomain attained the closed form. Myong *et. al.* suggested that affinity of Rep to free 3’-end increased in the closed conformation by exposing a secondary binding site for it. Since we observed the similar conformational change on the 5’ overhang DNA, such a site should have versatility to accommodate binding to a recessed 3’-end structure. Hence the other half of the repetitive shuttling mechanism lies in finding the functional residues that directs the helicase to the beginning of the translocation track. Our biochemical mutagenesis study of ssDNA interacting residues of Rep can shed more light on the location of the snapback facilitating binding site for the 3’ end. Briefly, we found that the certain mutations of ssDNA interacting residues seemed to cause long stalls at the end of the translocation cycles, indicating an impaired ability of snapback. Most notable of these residues is the Asn352, and to the some degree Arg350, Thr556, His558 and His580 all of which are located on 2A domain. Thus the

binding location of the target snapback site could very well be the some of the ssDNA binding residues on the leading edge motor domain (2A). Crystal structures show that these residues are more exposed in the closed form of 2B subdomain. See 2.4: Role of functional helicase residues in translocation and directionality for more details.

In a final note, our discovery of new modes of repetitive shuttling on the variety of substrates, along with previous results clearly shows the robustness of this phenomenon. Such variety of DNA structures that enables repetitive shuttling can be an indication of a greater role of Rep-like proteins as translocases *in vivo*.

2.1.8 Materials and Methods

Translocation of individual Rep molecules on DNA was monitored by smFRET between a donor (Cy3) and an acceptor (Cy5) fluorophores conjugated to Rep and DNA structures as shown in Figures. We have also employed protein induced fluorescence enhancement (PIFE) method to do the same on certain experiments as indicated in the text. For PIFE, we used a single fluorophore (Cy3) labeled DNA and an unlabeled protein (*56, 57, 74*). We also measured the fluorescence lifetime of DNA conjugated Cy3 during translocation of unlabeled Rep by time-correlated single photon counting (TCSPC) on a confocal microscope stage.

In some PIFE experiments, we used surface tethered Rep-biotin conjugates and supplied non-biotinylated labeled DNA substrate *in trans*. Surface tethering of DNA or protein employed well established biotin-neutravidin scheme.

Fluorophore labeled Rep molecules were prepared as described in (*72*). Cy3 or Cy5 maleimide dyes (GE Healthcare) were conjugated to Rep mutants at specific positions as indicated in the text. Their *in vitro* activity (ATPase, ssDNA translocation and/or dsDNA unwinding) and *in vivo* functionality (phage replication plaque assay) were confirmed previously (*72*). Residues 43 and 97 are located on the trailing end during the translocation and residue 333 is located on the leading edge. The labeling positions were selected to maximize the FRET efficiency change during translocation cycles. For surface tethering, Rep-biotin conjugates (at residue 43 or 97) were prepared as described in (*78*). For monitoring the conformation of 2B subdomain, a dual labeled version, Rep97/473-Cy3/Cy5 was prepared as described in (*58*). Residue 97 is located on the 1B domain, and 473 is located at the tip of the 2B subdomain, so

the closing of the 2B subdomain would increase the FRET efficiency as the distance between two residues reduced from 8.6 nm to 3.2nm.

For 2.1.1: Rep translocation on 5' overhang DNA, two biotinylated 5'-overhang DNA substrates were prepared. First was prepared by annealing oligonucleotides 5'-Cy3-T_{80}-GCC TCG CTG CCG TCG CCA-3' and 5'-biotin-TGG CGA CGG CAG CGA GGC- 3'. Second was prepared by annealing oligonucleotides 5'-T_{80}-GCC TCG CTG CCG TCG CCA-3' and 5'-biotin-TGG CGA CGG CAG CGA GGC-Cy5-3'. For 2.1.2: Repetitive shuttling on free ssDNA , two 98mer oligonucleotides of same sequence were used, 5'-Cy5-T_{80}-GCC TCG CTG CCG TCG CCA-3' and 5'-Cy3-T_{80}-GCC TCG CTG CCG TCG CCA-3'. For 2.1.3: Repetitive shuttling on ssDNA flanked by duplex DNA at both ends, we designed a ds-ss-ds DNA substrate had 18-bp dsDNA at both ends with 40nt poly-dT ssDNA in between. The recessed 5'-end carried the Cy5 label. We annealed three oligonucleotides to form ds-ss-ds DNA substrate: 5'-ACT CGC GCC GTC GCT CCG-T_{40}-GCC TCG CTG CCG TCG CCA-3', 5'-biotin-TGG CGA CGG CAG CGA GGC-3' and 5'-Cy5-T CGG ACG GAC GGC GCG AGT-3'. For 2.1.5: Conformational dynamics of Rep during the repetitive shuttling, we used a unlabeled 5' overhang DNA by annealing oligonucleotides 5' -T_{80}-GCC TCG CTG CCG TCG CCA-3' and 5'-biotin-TGG CGA CGG CAG CGA GGC- 3'. For 2.1.6: 3'-OH is not an essential factor in repetitive shuttling, we used a 3'-overhang DNA that was prepared by annealing a 3'-end phosphorylated oligonucleotide and 5'-TGG CGA CGG CAG CGA GGC- T_{80}-3'-Phos with 5'-Cy5-GCC TCG CTG CCG TCG CCA-biotin-3'.

Reaction chambers were formed by quartz slides and glass coverslips passivated with poly-ethyleneglycol (PEG) and 1% biotinylated PEG (mPEG-SC and bio-PEG-SC respectively, Laysan Bio, Arab, AL) to reduce non-specific adsorption of protein (*79*), followed by 5 min incubation with Neutravidin (Thermo Scientific, Newington, NH) for immobilization of biotinylated molecules or vesicles on the chamber surface as described below.

In the single molecule experiments where the DNA was surface-tethered, the reaction chamber was first incubated with 50 pM-100 pM biotinylated DNA substrate in T50 buffer (10 mM Tris-HCl pH8.0, 50 mM NaCl) for 5 min, followed by flushing excess DNA. Translocation was initiated by flowing 0.1- 1 nM Rep protein and 1 mM ATP in the reaction buffer (10 mM Tris-HCl [pH 8.0], 10 mM $MgCl_2$, 15 mM NaCl, 10% (v/v) glycerol, 1% (v/v) gloxy and 0.2% (w/v) glucose, an oxygen scavenging system (*80*) and 3-4 mM Trolox, (*81, 82*)). In the surface-tethered Rep experiments, the reaction chamber was incubated with 0.1-1.0 nM Rep-biotin in the reaction

buffer, followed by flushing excess protein. The translocation was initiated by flowing 0.1-1.0 nM non-biotinylated DNA and 1 mM ATP in the reaction buffer. For translocation experiments performed in vesicles, 5'-Cy5 labeled 98-nt ssDNA and Rep-Cy3 were encapsulated in small unilamellar vesicles (SUV) and tethered to surface as described in 2.2: Use of phospholipid vesicles for prolonged observation of single Rep molecules. The translocation was started by injecting 1 mM ATP in the reaction buffer.

All smFRET experiments were conducted on a custom-built prism type inverted TIRF microscope (Olympus IX71). For excitation, we used a 532 nm Nd:YAG laser. Signal from individual molecules in a 70 µm x 35 µm imaging area was collected by 60X NA water objective, split into two side-by-side images of donor and acceptor channels using a home-built emission optics pathway and acquired by an EMCCD camera (iXon DV 887-BI, Andor Technologies, CT) at 16-100 ms resolution using a custom C++ software as described in (*83, 84*). Movie acquisition, analysis and generating donor/acceptor signal time series were done using custom C++ and IDL codes. As needed, E_{FRET} was calculated as described in (*83*).

Maria Sorokina and I performed single molecule Cy3 fluorescence lifetime measurements to monitor unlabeled Rep translocation on surface tethered Cy3 labeled 5' overhang DNA on a home-built time correlated single photon counting (TCSPC) setup. For details of the setup, software and analysis, see (*85*). Briefly, a confocal microscope piezo stage was excited with a mode locked picosecond pulse laser that was split and detected by a photodiode for synching. Individual photon emissions detected by an ultra-fast avalanche photodiode (APD). Photons were time correlated by a TCSPC card using signals from both photodiodes. Lifetimes were calculated by fitting the time distribution histogram of 500 photons. An initial raster scan of a preset area was done to record the positions of DNA-Cy3 molecules. Once the translocation reaction started, piezo stage moved to the position of each molecule, data was acquired until the Cy3 photobleached, and then the stage moved to the next molecule and so on.

2.2 Use of phospholipid vesicles for prolonged observation of single Rep molecules*

In this topic I will elaborate on the vesicle encapsulation technique, that we used for observing the helicase translocase activity presented in this chapter and for demonstrating K2 effect on FRET in Chapter 4. Encapsulation of a biological molecule or a molecular complex in a vesicle provides a means of bio-friendly immobilization for single molecule studies that allows collecting data from individual molecules until fluorophore labels photobleach. We used a strategy for making porous vesicles using a bacterial pore-forming toxin, α-hemolysin published in (*75*). This approach was tested by co-encapsulating *E. coli* Rep helicase protein and its single-stranded DNA track. The DNA translocation activity of Rep helicase inside vesicles was fueled by ATP provided outside the vesicle, and a dramatically higher number of translocation cycles could be observed due to the minuscule vesicle volume that facilitates rapid rebinding after dissociation. These pores are known to be stable over a wide range of experimental conditions, especially at various temperatures, which is not possible with the previous method using DMPC vesicles.

2.2.1 Introduction

Single molecule fluorescence techniques are revolutionizing biological inquiries both in vitro and in vivo (*86-89*). In order to observe single molecule reactions for an extended period, molecular movements via diffusion have to be slower than the time scale of the method used to track their movements. Although it is possible to track particles in one dimension, for example cytoskeleton motor proteins moving on actin filaments (*90*) or microtubules (*91*), or in two dimensions, as in the case of membrane protein diffusion (*92*), there are many other examples where three-dimensional diffusion would preclude an extended observation time window. Therefore, several methodologies have been developed to keep the molecules of interest under observation for prolonged periods. Immobilization via biotin−streptavidin anchors on bare (*93*) or polymer-grafted surfaces (*94*) and confinement within agarose or polyacrylamide gels (*95*) are very useful strategies, yet care should be taken to ensure minimal perturbation to the

* The content in this subsection was reproduced with permission from "Single molecule nanocontainers made porous using a bacterial toxin", Okumus B., Arslan, S., Fengler, S. M., Myong, S. & Ha, T., Journal of the American Chemical Society, 131 (41), 14844-9 (2009). Copyright 2009 American Chemical Society.

activity of biological molecules. A recent report (*96*) described an elegant technology for the confinement of individual molecules in free solution by means of anti-Brownian electrokinetic trapping (ABEL). ABEL, however, does not allow changing the solution condition while keeping the molecule trapped. An alternative approach entraps a molecule within a small unilamellar vesicle (SUV) which is then tethered on surfaces (*97-99*). Vesicle encapsulation has been used to study protein folding (*100*), RNA folding (*97*), DNA folding (*93*), DNA−protein interaction (*101*), protein−protein interaction (*99*), and protein conformational changes (*102*). In principle, vesicle encapsulation should provide a more biofriendly environment (*103*), which may help avoid artifacts sometimes caused by direct surface tethering of proteins (*103, 104*), because the immediate environment is composed of lipid molecules rather than an artificial solid substrate such as a glass slide. Vesicle encapsulation also obviates the need to modify biological molecules for surface tethering (e.g., biotinylation), offering ease and flexibility to the researcher. For example, we utilized it to determine the effect on FRET of dipole orientations of the fluorophores stacked at DNA duplex ends (*105*). However, the inability to exchange buffer across the vesicle wall was noted as the major drawback (*94*) because buffer exchange is commonly practiced either for examining molecules under various conditions, for example by varying concentration of certain chemical agents between measurements (*93*), or for triggering certain reactions on demand (*78*). Rendering the vesicles porous is therefore essential for fully realizing the potential of the encapsulation approach. This is already achieved for bulk assays (*106-108*), but their adaptation for single molecule experiments has been limited.

Porous vesicles should also be useful if the measurement condition is not suitable for vesicle formation. In such circumstances, vesicles can be formed under conditions favorable for encapsulation, and the solution condition can then be changed to the desired imaging conditions. Furthermore, because porous vesicles allow buffer exchange without washing out the encapsulated molecules, it would be possible to monitor the interaction between the same pair of biomolecules under various conditions. Finally, the presence of pores prevents the vesicle-to-vesicle variation of intravesicular conditions, for example ion concentrations, emerging from stochastic encapsulation and vesicle size distribution, which may be important for studying systems that display strong heterogeneity (*98*). In this regard, pores ensure that the intravesicular condition is in equilibrium with the bulk solution, hence remains identical between vesicles. Clearly, introducing pores on vesicles should greatly expand the venues for single molecule vesicle encapsulation assays.

In order to fully harness the capabilities of vesicle encapsulation and overcome the impermeability barrier, our colleagues recently exploited the thermotropic properties of lipid membranes (*101*) where the membrane permeability of solutes reaches a maximum around membrane's melting temperature (T_m). They showed that the vesicles made of the phospholipids DMPC (dimyristoylphosphatidylcholine) (T_m = 23 °C) can readily allow the passage of ATP and its analogues (*101*) at room temperature. Nevertheless, the pores rely on transient formation of lipid packing defects, and their presence and properties may be affected by many factors. An obvious limitation is that the working temperature range is narrow with the membrane being porous only within 3 °C of T_m (*109*), making it laborious to carry out measurements at various temperatures. For steady-state measurements, one would need to incubate the sample with the desired buffer at the T_m of the lipid membrane before taking data at a new temperature. Moreover, data acquisition at temperatures different from T_m would be problematic if the encapsulated reaction consumes substrates (e.g., ATP), as their consumption and product accumulation will alter the dynamics. Finally, nonequilibrium kinetic experiments, for example via flow, at various temperatures would require several different lipid compositions with T_m values that match those of the temperatures to be used. We therefore utilized a bacterial toxin, α-hemolysin (aHL), to induce stable pores on SUVs, which had been previously used for bulk measurements in an attempt to build an artificial cell (*106*). Our colleagues proposed using aHL in an earlier report (*101*), and the actual implementation is the subject of this section.

aHL is a bacterial toxin secreted by *Staphylococcus aureus* as water-soluble monomers which incorporate into lipid membranes to form a heptameric pore (*110*). There are several reasons for our choice of aHL: (i) Owing to spontaneous incorporation of aHL into the membrane, such pores are easier to make than many other membrane proteins which require difficult reconstitution procedures. (ii) Once formed, aHL pores are stable in a wide variety of buffer conditions, pH, and temperature. (iii) The aHL pore with a narrowest opening of 1.8 nm is not selective, allowing passive passage of molecules up to 3kD both ways (*111*). (iv) Finally, available engineered mutants of aHL may be further utilized to achieve unique experimental assays.

My colleagues, Okumus and Fengler, and I carried out the proof-of-concept experiments using smFRET (single molecule fluorescence resonance energy transfer) (*83, 112*) on a home-built total internal reflection fluorescence setup (Figure 2.10) using three different model

systems. Okumus and Fengler carried out experiments pertaining the first two systems, and Okumus and I tested the last system. First, the hairpin ribozyme whose single molecule behavior under various conditions is well characterized (*113*) served as a molecular reporter for the intravesicular buffer condition, in this case magnesium ion concentration. Second, the RNA four-way junction was also used as a well-characterized test system (*114*). As the third model system we encapsulated *E. coli* Rep helicase together with its DNA substrate in order to demonstrate the passage of ATP through aHL pores and to showcase the ability to study transient protein−DNA interactions for a prolonged period. Here I will present our finding on the third system.

2.2.2 Results

We encapsulated an acceptor (Cy5)-labeled partial duplex DNA with a donor (Cy3)-labeled Rep helicase monomer in SUVs (Materials and Methods and Figure 2.11 a) in order to confirm the ATP passage through aHL pores. A Rep monomer, which cannot unwind duplex DNA by itself (*40, 78*), can translocate on a single-stranded DNA in the 3′ to 5′ direction. When Rep encounters a blockade (such as a double-stranded DNA) on its track, it snaps back to the 3′ end on its substrate. These repetitive translocation events (termed shuttling) continue until the Rep monomer dissociates from DNA (*58*). Such measurements were previously carried out on DNA molecules tethered to a quartz surface coated with polymers. For the surface experiments, shuttling takes place a limited number of times (typically less than 10) due to eventual dissociation, and the number of translocation events is thus inadequate for inferring statistically significant information on the shuttling properties from a single Rep helicase (Figure 2.11 b). In contrast, we expected that encapsulation would enable long-term monitoring of shuttling for a single pair of Rep and DNA, as the molecules would quickly rebind within the minuscule SUV volume.

Upon surface immobilization of SUVs, and subsequent ATP injection, shuttling events were observed within SUVs (Figure 2.11 c), indicating the ATP passage through aHL. The translocation activity was observed only when aHL was incorporated and ATP was present. Number of repetitive translocation events that can be observed from an individual Rep was substantially higher in the SUVs compared to the surface assays as anticipated. When encapsulated, a single Rep−DNA pair exhibited shuttling essentially until photobleaching of the fluorophores. Such enhancement for the number of shuttling events makes it possible to

observe more than 100 cycles of translocation from an individual Rep−DNA pair and to obtain the kinetic parameters reliably free of complications arising from single molecule heterogeneity (Figure 2.11 d).

In conclusion, in order to fully harness the capabilities of the vesicle encapsulation, we described here a generalizable methodology for making robustly porous vesicles. aHL pores utilized in this study allowed rapid passage of Mg^{2+} as well as ATP through membranes of phospholipids SUVs. Compared to the thermothropically rendered defects, aHL pores are well-defined and stable at various temperatures. In the future, engineered aHL mutants of which opening/closing can be controlled via different chemical or optical means (*111, 115, 116*) may be exploited to design novel experimental platforms.

2.2.3 Materials and Methods

2.2.3.1 FRET Measurements

All experiments were carried out using a prism-based two-color total internal reflection fluorescence microscope capable of single molecule FRET analysis (*83*) at room temperature and with 100 ms exposure time unless otherwise stated. A 532 nm laser was used for donor excitation in FRET measurements. A 633 nm laser was used to excite the acceptor directly to confirm colocalization of a donor-labeled Rep helicase and an acceptor-labeled DNA. FRET efficiency was calculated as the acceptor intensity divided by the total intensity of donor and acceptor fluorescence averaged over 1 s. Data from hundreds of molecules were pooled to obtain FRET histograms. Noise in the measurements could yield FRET values below 0 and above 1 due to negative values after background subtraction for the acceptor and donor intensities, respectively. The leftmost peak in FRET histograms represents donor only population, hence was ignored in data interpretation.

2.2.3.2 Purification and Characterization of aHL

Initial attempts by our colleague Yuji Ishitsuka with the commercial aHL (Calbiochem) were not successful in forming pores despite our efforts that explored a wide variety of experimental conditions. As this failure is likely due to the low concentration of commercially available aHL (the practical maximum concentration we could use was ~1 μM in monomer), he produced aHL in-house. We overexpressed hexahistidine-tagged aHL from a plasmid in *Escherichia coli* and purified the protein using a Ni-NTA column as described in (*75*).

Recombinant aHL concentration was estimated to be ~36 μM by quantitative Western blotting. For long-term storage, aHL was divided into aliquots, flash frozen, and kept at −80 °C.

2.2.3.3 aHL Incorporation on SUV

A 30−50 μL volume of aHL solution with the 36 μM monomer concentration was injected directly into the flow cell containing surface-immobilized vesicles. Unless stated otherwise, the vesicles were incubated with the aHL containing T50 solution (10 mM Tris (pH 8.0) with 50 mM NaCl) at room temperature for at least 15 min. Excess aHLs were washed away with the buffer (*75*).

2.2.3.4 Rep Helicase Encapsulation

Partial duplex DNA (400 nM, double-stranded DNA of 18 base pairs with a 3′ $(dT)_{80}$ and a Cy5 attached to the junction) and a 400 nM concentration of Rep molecules (*58*) labeled with a donor fluorophore (Cy3) mixed in a total volume of 200 μL of buffer R (20 mM Tris pH 8.0, 500 mM NaCl) were used to hydrate 2.5 mg of lipid films. In order to achieve, on the average, one molecule of each kind (i.e., one protein vs one DNA) in a 100 nm diameter vesicle, a concentration of 400 nM was chosen. Buffer R contains 500 mM to ensure that the labeled Rep does not aggregate, as when high concentrations of dye-labeled Rep monomers are dissolved under low salt conditions. Upon forming the multilamellar vesicles, 100 or 200 nm (concentrations were scaled accordingly) diameter vesicles were formed via extrusion. Finally, unencapsulated Rep molecules were separated from the vesicles by a Ni-NTA column via the histidine tag on the protein as described below. While buffer R had to be used during encapsulation to avoid clumping of Rep, it had to be exchanged with the imaging buffer (*88*) (10 mM Tris-HCl, pH 7.6, 1 mM ATP, 12 mM $MgCl_2$, 15 mM NaCl, 10% (v/v) glycerol, 0.4% (w/w) glucose, anti-blinking agents such as 2-mercaptoethanol or Trolox, and oxygen scavenging system utilizing catalase and glucose oxidase) during data acquisition because high salt concentration in buffer R does not allow Rep to bind to the DNA.

A Bio-Rad column was packed with 2 mL of Ni-NTA bead suspension. After adding the Ni-NTA beads, the column was first washed with 3 mL of distilled water, then ~10 mL of buffer R The column material was thus equilibrated with the buffer in which the vesicles were prepared. Having the beads all settled down, ~150 μL of the sample was gently loaded on the Ni-NTA bead matrix. Initially, 3-4 drops (volume of one drop was determined to be typically ~50 μL) came out of the column due to the volume displaced by the sample loading. After waiting 5 min. to make

sure that the binding to column is equilibrated, 3 mL of buffer R was used to elute the sample. The separation was first characterized using solutions of Rep or vesicles only. We found that up to 1 μM of Rep could totally be captured by the Ni-NTA column, and the vesicles eluted between ~10-15th drops. For the Rep encapsulation sample, drops between 10-15 (counting includes the first 3-4 drops during the sample loading) was collected in a test tube, and used immediately. The molecules are successfully encapsulated inside vesicles. Vesicles were diluted by 20 to 200 times in T50 buffer (10 mM Tris pH 8.0, 50 mM NaCl) and then incubated with the coated quartz slide. PEG coated slides were used for the Rep experiments. When the slide surface was not covered with neutravidin, insignificant binding was observed. In sharp contrast, same dilution of vesicles yielded a good coverage of fluorescent spots on the surface after neutravidin incubation. We therefore concluded that the binding was specific to only the biotinylated objects (i.e. the vesicles), and that the non-specific binding was undetectable.

The headgroup of EggPC lipids which were used for making the encapsulation samples is zwitterionic. Although there is no particular reason to think that the Rep helicase would interact with the lipids, we tested the binding of Rep helicase directly on supported lipid bilayer (SLB) or blank vesicles immobilized on SLB. 1 nM of Rep showed very few binding the surfaces provided that a high quality SLB is formed. Although such measurements cannot rule out transient interactions between Rep and lipid membranes, evidently the Rep activity is not altered inside vesicles.

The freeze-thaw cycle is not essential but is thought to make encapsulation more efficient. We had previously characterized the effect of freeze-thawing on the unwinding activity of Rep helicase at the bulk level. 10 freeze-thaw cycles decreased the DNA unwinding amplitude by only 5% (data not shown) when the thawing was carried out in room temperature water bath. However, single molecule measurements done on vesicles prepared with freeze-thawing showed no Rep activity (translocation). We therefore skipped the freeze-thawing step during the encapsulation only after when Rep translocation was observed for encapsulation samples. The distribution of the number of encapsulated molecules was not uniform, i.e. some vesicles had only one Rep molecules whereas others had multiple (e.g. four or five of them) encapsulated within (The number of Rep molecules could be estimated by counting the digital photobleaching events.). Such non-uniform encapsulation efficiency could be due to aggregation of Rep monomers or reflect the fact that the freeze-thawing makes the encapsulation not only more effective, but also more homogenous. For the single molecule analysis, only the traces

that showed proper signatures - e.g. single step photobleaching, and constant total (donor + acceptor) intensity - were considered.

2.3 Activation energy of Rep translocation*

The enzymatic translocase activity of Rep helicase was measured at various temperatures using vesicle encapsulation technique (*117*). Temperature dependence of the translocation rate of Rep helicase on ssDNA was used to calculate the activation energy of stepping.

2.3.1 Results

Labeled Rep helicase and partial duplex DNA without biotin were coencapsulated within a vesicle to observe the rate of translocation at four different temperatures (16, 22, 30, and 33 °C). Surface-immobilized vesicles were incubated with aHL within the channel as described above. Cy3-labeled Rep helicase showed a characteristic sawtooth repetitive translocation pattern along the single-strand portion of the Cy5-labeled partial duplex DNA upon addition of 1 mM ATP. No activity was observed when the vesicles were not treated with aHL or when ATP was not added. Figure 2.12 shows four representative FRET time traces at different temperatures. The period for each sawtooth, Δt, clearly showed a correlated increase as the temperature was increased. Histograms of the Δt values were fitted with Gaussian functions to obtain average Δt values of 2.4, 1.2, 0.5, and 0.4 s for 16, 22, 30, and 33 °C, respectively. The Arrhenius fit of these points gave an apparent activation energy of 82 ± 3 kJ mol^{-1} or ~1 kJ mol^{-1} $base^{-1}$ (Figure 2.13).

Unlike thermotropically porous vesicles, aHL porous vesicles are permeable independent of the temperature (*117*). This characteristic is useful to measure enzymatic activity at multiple temperatures. Our proof-of-concept measurements of Rep helicase translocations along ssDNA inside aHL porous vesicles at four different temperatures clearly showed this capability. In the bulk study, denatured enzymes in the system negatively contribute to the overall rate as the temperature increases. However, the single-molecule approach allows rate measurements purely from the active enzymes. Due to the highly effective

* The content in this subsection was adapted with permission from "Temperature-independent porous nanocontainers for single-molecule fluorescence studies", Ishitsuka, Y., Okumus, B., Arslan, S., Chen, K. H. & Ha, T. *Analytical chemistry* **82**(23), 9694-701 (2010). Copyright (2010) American Chemical Society.

local confinement of ssDNA and Rep helicase, we were able to observe these translocations more consecutively than we would have outside of the nanocontainer. Such a high concentration is not currently attainable due to added high background signals as well as nonspecific binding of Rep helicase for the conventional surface measurement. As mentioned before, encapsulation also allows measurement of interactions between the same set of molecules due to the prolonged observation times.

2.3.2 Materials and Methods

2.3.2.1 Vesicle Preparation

Egg phosphatidylcholine (eggPC), 1,2-dipalmitoyl-*sn*-glycero-3-phosphoethanolamine-*N*-(cap biotinyl) (sodium salt) (biotinPE), and 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) were purchased from Avanti Polar Lipids, Inc. (Alabaster, AL). Cholesterol (Chol) was purchased from Sigma-Aldrich (St. Louis, MO). Solvents were obtained from Fisher Scientific (Pittsburgh, PA). All chemicals were used without further purification. All vesicles were prepared by the extrusion method. Lipids dissolved in chloroform were measured out using a gastight syringe (Hamilton, Reno, NV) and mixed to make the final desired molar ratio (99:1 eggPC/biotinPE or 33:66:1 eggPC/Chol/biotinPE). The chloroform was dried off by a gentle stream of nitrogen, leaving a thin lipid film. This film was further dried under vacuum for at least 3 h.

2.3.2.2 Rep Helicase−DNA Coencapsulation

Rep helicase−DNA coencapsulation was done as described previously with some minor modifications (*118*). In brief, the lipid film was hydrated with 400 nM partial duplex DNA (18-base-pair double-stranded DNA with a Cy5-conjugated single-stranded 3'-$(dT)_{80}$) in buffer R (20 mM Tris (pH 8.0), 500 mM NaCl) to obtain a final lipid concentration of 12.5 mg/mL. DNA-containing multilamellar vesicles were subjected to five freeze−thaw cycles. Donor dye (Cy3) labeled Rep helicase (*58*) was added to give a 400 nM concentration just before the extrusion step (39 passes, 200 nm pore size) to minimize the deactivation of proteins from freeze−thaw cycles. These concentrations were chosen to achieve, on average, one DNA molecule or one DNA−helicase pair per vesicle for a given size of the vesicle. The number of molecules per vesicle was quantified by counting the photobleaching steps from single-molecule time traces as mentioned above. To minimize the nonspecific surface binding, non-encapsulated Rep molecules were removed using a Ni−NTA agarose column via the histidine tag on the protein as before.

2.3.2.3 Single-Molecule Fluorescence Measurements

After the assembly of the flow cell, the vesicle sample was immobilized on the surface by first incubating the surface with 0.2 mg/mL NeutrAvidin (Pierce, Rockford, IL) in 10 mM Tris (pH 8.0) with 50 mM NaCl for 5 min, followed by a buffer wash and incubation of the vesicle sample for 15 min. Excess vesicles and nonencapsulated molecules—if any—were flushed away by the buffer solution. smFRET measurements of the Rep helicase−DNA system were done using 10 mM Tris (pH 7.6), 1 mM ATP, 12 mM $MgCl_2$, 15 mM NaCl, 10% (v/v) glycerol, 0.8% (w/w) dextrose, antiblinking agents such as 2-mercaptoethanol or Trolox, and an oxygen-scavenging system, utilizing catalase and glucose oxidase as mentioned above.

2.3.2.4 aHL Expression and Purification

aHLs were expressed and purified in-house by Yuji Ishitsuka as described previously (*75*). Briefly, histidine-tagged aHL was overexpressed in *Escherichia coli* (BL21-DE3) and purified it using a Ni−NTA agarose column. The recombinant aHL concentration was estimated to be ≈36 μM by quantitative Western blotting. Aliquoted aHL stock solutions were flash frozen and stored at −80 °C. Unless stated otherwise, all experiments were done using in-house aHL. A commercial aHL was purchased from Calbiochem (San Diego, CA).

2.3.2.5 aHL Incorporation

A 30−50 μL volume of aHL solution with the indicated concentration was injected directly into the flow cell containing surface-immobilized vesicles. Unless stated otherwise, the vesicles were incubated with the aHL solution (10 mM Tris (pH 8.0) with 50 mM NaCl) at room temperature for at least 30 min. Excess aHLs were washed away with the buffer (*75*).

2.3.2.6 smFRET Data Acquisition and Analysis

A prism-based total internal reflection fluorescence inverted microscope (IX70, Olympus) equipped with a water immersion objective lens (60×, NA = 1.2, Olympus) was used for imaging. The details on the optical setup may be found elsewhere (*87*). The sample temperature was controlled by heating or cooling the slide stage, prism holder, and objective lens with recirculating water (Neslab Instruments, Portsmouth, NH). Mapping calibration, matching donor spots to the corresponding acceptor spots, was achieved using surface-adsorbed fluorescent beads (0.2 μm, 540/560 nm, Invitrogen, Carlsbad, CA). Fluorescence data, in the form of successive images, were recorded by custom software. The donor and acceptor intensity time trace information was extracted out using custom IDL (Research Systems Inc.)

programs. Background and donor leakage subtracted donor (I_D) and acceptor (I_A) intensities were used to calculate the apparent FRET efficiency ($E = I_A/(I_D + I_A)$) using custom Matlab programs (MathWorks, Natick, MA). To quantitate the permeability of the vesicle, the relative population of unfolded GQ-DNA molecules was used. The number of molecules with *E* values between 0.15 and 0.5 and larger than 0.5 were counted as unfolded and folded GQ-DNA molecules, respectively. The molecules with *E* less than 0.15 were DNA molecules without acceptor dye or with inactive acceptor dye and were excluded from the analysis.

2.4 Role of functional helicase residues in translocation and directionality

Directional translocation of helicases and motor proteins are essential to their function, and it is a result of concerted modulation of the complex network of interactions of the protein with the ssDNA track, couple to the ATP hydrolysis. So far, we have only the static pictures of this interaction network from the structure studies, and very little of atomic level understanding of the dynamical picture as to how the directional translocation is produced. One way to understand the dynamic picture is to remove its elements one at a time to see the resulting effect on the activity, so we performed site directed biochemical mutations of ssDNA interacting residues of Rep, tested their translocation activity using the repetitive shuttling assay. We chose the target residues based on crystal structures of Rep, UvrD and PcrA which share genetic and structural homology to a great extent (~40% sequence homology and >80% structural homology). The ssDNA interacting residues of the Rep are mostly part of the conserved helicase motifs, and some residues are highly conserved even across several superfamilies. They are mainly located on the two RecA like folds which are the motor domains of these helicases, named 1A and 2A (red and yellow domains in Figure 1.4 and Figure 1.5). How their binding affinities are modulated by ATP hydrolysis forms the atomic level basis of the translocation. Crystal structures showed that the relative positions of the ssDNA interaction network on 1A and 2A are 1-nt closer to each other in the ATP bound state, compared to the no-nucleotide state. These observations suggested an inchworm model that each of the two motor domains moves 1-nt alternatingly to translocate unidirectionally on the ssDNA track. Upon ATP binding weakly bound (to ssDNA) motor domain (let's say domain 1) is pulled toward the strongly bound other motor domain (domain 2). This is most likely a compressed state held together by the ATP

and the Mg^{2+} ion. Then, ATP hydrolysis occur, byproducts (ADP and Pi) are released. After ATP hydrolysis, binding strength of the motor domains are reversed, relaxation back to the no-nucleotide state occur by 1-nt movement of domain 2, because the domain 1 is now more tightly bound to the DNA. In case of Rep, domain 1 and 2 are 1A and 2A respectively, resulting in 3'→ 5' translocation, but in case of RecD2 or Dda (*119, 120*), Rep homologs with reverse polarity, domain 1 is 2A, and domain 2 is 1A, resulting in 5'→3' polarity. We performed site directed mutagenesis of ssDNA interacting residues in each motor subdomain, we wanted to show how the translocation and directionality was affected, so that we can pinpoint important residues and their roles.

2.4.1 Mutagenesis of ssDNA binding residues of Rep

An interaction network of Rep residues and ssDNA is shown in Figure 1.6 that was indicated by the crystal structure (*61*) From these, among our target sites, Thr556 and Thr83 are two of the backbone binding residues on the forward (2A) and the rear (1A) motor domains respectively (nomenclature is based on polarity). These threonines are conserved across many SF1 and SF2 helicases. Trp250 and Phe183 (the gatekeeper residue) forms aromatic stacking with the bases. Although these two residues are on the rear motor domain, Trp250 is at the tip of a loop that extends toward the front motor domain, positioning it between 1A and 2A. Nearby Tyr248, like Trp250, is situated against the ssDNA nucleobases in between the two motor domains. His580 and His85 are interacting with the base via hydrogen bonding. His580 is not a part of any motif, but it is conserved in PcrA which was crystallized in an aromatic stacking interaction with oligonucleotide base, and it is replaced with serine in UvrD. His85 on the other hand is conserved across the SF1 helicases, as part of the TxGx motif. The same is true for Thr556, that is part of motif V. His353 is not conserved residue, and it was not indicated in any structure study, however it was in the close proximity of Arg-Gly-Asn (350-352) block (motif Iva). Since this residue is already an Ala in PcrA and UvrD, His353Ala mutation is unlikely to have any effect Rep, but it may change the translocation speed. Another set of conserved residues in motif Ia are Thr56, Asn57 and Lys58. Lastly, we made the Leu595Ile/Gly599Ala double mutation that was proposed to change the translocation directionality based on molecular dynamics simulation studies of PcrA (*121, 122*).

2.4.2 Results

Results of alanine replacement mutations of these residues were varied from complete inhibition of translocation to that is faster than the wild type or loss of directional translocation. Cy3 labeled mutants were synthesized as described in reference (*72*). For the translocation activity we used the partial duplex DNA with 3' tail that is a 80-nt long poly-dT (pdT80). The partial duplex DNA was labeled with Cy5 at the junction. Cy3 labeled Mutant Rep behavior was recorded using the single molecule assay described in as described in (*58*) and, sections 2.1.8 and 2.4.4. Materials and Method. Representative time traces are shown for each mutant in Figure 2.8.

Thr556Ala, His85Ala, His353Ala, His580Ala, Lys58Ala, Leu595Ile/Gly599Ala mutations preserved the repetitive directional translocation activity, albeit with slower or faster velocities with respect to the wild type. Among those Thr556Ala, His580Ala mutants were slower 4-5 times than the wild type. His85Ala and His353Ala mutants were faster than the wild type 3 and 1.7 times respectively. The gatekeeper residue mutant Phe183Ala, a motif Ia mutant Ly58Ala, and the mutant with a proposed reverse directionality did not have any distinguishable translocation property than the wild type. A comparison plot of translocation speeds for all the mutants is shown in Figure 2.9.

Trp250Ala mutant, on the other, hand lost the strict polarity of the translocation which in turn yielded rather irregular FRET traces, instead of regular sawtooth pattern, an indicator of unidirectional movement. Mutation of nearby Tyr residue, Tyr248Ala also seemed to be no longer strictly unidirectional. These results present a strong indication of determination of translocation directionality by the residues that are transferring the ssDNA nucleobases from forward to rear motor domain.

The other motif Ia mutants on the rear motor domain, Thr83Ala, Thr56Ala, and to a large degree Asn57Ala had no translocation activity. However Thr83Ser mutant conserved the wild type activity, pointing the importance of specific hydrogen bonding between alcohol sidechains of Thr and Ser with the backbone phosphate. Despite the inhibited translocation, Thr83Ala was still able to show ssDNA dependent ATP hydrolysis, but such activity was completely abolished for Thr56. This finding indicates that Thr56 is part of the network that links the ATP hydrolysis to the ssDNA binding. Thr83Ala mutant was very likely attempting the 1-nt steps proposed by the inchworm model during ATP binding and hydrolysis, but the moving

motor domain was probably sliding back because the Ala at this position is not providing necessary DNA affinity immediately after ATP hydrolysis. It is also noteworthy that the His85Ala mutant of nearby His in the same motif as Thr83 was the fastest Rep mutant we discovered.

2.4.3 Discussion

The inchworm model is based on alternating binding affinity of forward and rear motor domains, in synchronization with ATP/Mg^{2+} binding, hydrolysis and release of APD.Pi/Mg^{2+}. However the exact release sequence of the byproducts is unknown. In the no-nucleotide state, two domains are separated by 4 nt measured from Thr556 to Thr83. As the ATP binds, rear domain moves single nucleotide on the DNA, since the forward domain is bound to ssDNA more tightly in the rigor state. In the ATP bound state, called substrate state, forward domain binding to ssDNA weakens, and rear domain becomes tightly bound. In this state the distance between the domains is 3 nt. After or during the ATP hydrolysis and release of products, the two domains relax back to the no-nucleotide state as the weakly bound forward domain is pushed one nucleotide forward. In this model, alternation of affinity of the two motor domains to ssDNA is the essence of the directional translocation mechanism. Exact sequences of events are not known. Model based on crystal structures need verification with dynamic data. Using the single molecule assay and repetitive shuttling as readout, we investigated how the perturbations to the proposed interactions based on crystal structure would affect the translocation mechanism. The mutagenesis results could help us determine the residues that build up the binding affinity of each domain. We have determined which residues are important for determining the translocation speed and directionality.

Now we are closer to the dynamic picture of this model, based on the mutant behavior we observed. But further investigation of the correlation between ATPase and translocation speed of the mutants and structural molecular dynamics analysis are required to propose more comprehensive model for the directional translocation. Our understanding of the translocation mechanism based on this model would be ultimately tested if we can reverse the directionality of a helicase. Our preliminary results suggest that we already achieved bi-directionality on a directional helicase in case of Trp250Ala, and Tyr248Ala mutants.

2.4.4 Materials and Methods

Site directed mutagenesis, expression and labeling of Rep helicase were performed as described in (*72*). Partially duplex DNA with 80 nt long 3' overhang (pdT80) was prepared as described in (*58*). Rep translocation assay of mutants using smFRET TIRF microscopy was performed as described previously (*58, 78*). ssDNA ATPase rates of mutants were measured using Enzchek kit (Invitrogen), in the presence of 1uM dT_{45} ssDNA substrate, and 2mM ATP, 15 mM NaCl, 10mM $MgCl_2$, 10% glycerol (v/v) in a 10mM Tris.HCl pH8.0 buffer.

2.5 Effect of DNA modifications on Rep translocation

Effects of duplex nucleic acid irregularities on helicase unwinding have been investigated (*123-126*), but how the helicase translocation is influenced by these irregularities is so far missing. Therefore we probed the helicase-ssDNA interactions, by perturbing the DNA structure, to see how the translocation activity was affected. Crystal structure of Rep helicase with ssDNA suggests particular pi stacking and hydrogen bonding interactions between nucleobases and certain protein residues as well as characteristic hydrogen bonding to the DNA backbone which has been seen in other backbone tracking helicases (*61, 63, 121*). For the purpose disrupting these interactions, abasic sites (nucleotide without base) and PEG linkages were inserted to ssDNA track (Figure 2.14 a). We complemented our results with biochemical mutations of the Rep residues involved in these interactions to see how the translocation behavior would change.

2.5.1 Results

We compared the translocation of Rep on 5 different modified DNAs. Modified DNAs contain one to three abasic sites interspersed (substrates 1A, 2A and 3A), 3 consecutive abasic sites (substrate 3Ac) and PEG_6 chain inserted to a region approximately 3-nt long (substrate PEG25) (Figure 2.14 a). All the DNA substrates had the similar contour lengths to that of a 80 nt long oligonucleotide. Exact location of modification sites are listed in Figure 2.14 a. All the measurements involved partial duplexes with Cy5 labeled junction and Rep labeled with Cy3 at position 333 that is on the leading end during the translocation (*72*). FRET traces indicate that Rep is able to bind and translocate repetitively on all the modified partial duplex DNAs, as it does on the unmodified pdDNA (*58, 127*).

During the translocations, FRET efficiency increases from a minimum to a maximum value which correspond to the start and stop points of each shuttling event, namely the 3' end and the junction where the Cy5 is. Superimposition of FRET efficiency traces are shown in Figure 2.15. FRET efficiency maxima were same for translocations on all the modified and unmodified DNAs. This means that the modification sites did not inhibit the translocations nor prematurely truncated them at the modification sites, however, Rep took longer on the modified DNAs to traverse the whole length of the DNA. This is reflected in the shift seen in ensemble distribution of the translocation periods (Figure 2.14 b). Accordingly, the average speed of Rep was smaller on the modified DNAs (1A, 2A, 3A, 3Ac, PEG25) than that on the unmodified DNA (pdT80).

Despite the slowdown of translocation by 100 – 200 ms due to a single modification site, we were not able to identify distinct stall plateaus on the FRET traces. In order to investigate whether the slowing effect was global or local, we employed another type of analysis, namely, we looked at the FRET efficiency gradient (dE_{FRET}/dt vs. E_{FRET} plots). We calculated the ensemble averaged gradient of FRET efficiency curves from unmodified DNA (Figure 2.16). Since the FRET efficiency profile of each translocation event resembles a skewed S shape, its derivative is a bell shaped curve. Introduction of a stall site is expected to yield a dip in the bell shaped gradient curve (dE_{FRET}/dt vs. E_{FRET}) representing a local stalling rather than a global slowing down. Such localized deviations of modified DNAs from the unmodified pdT80 gradient curve were observed confirming the local nature of stalling due to the specific modifications. Our results suggest that the addition of each non-neighboring abasic site on the ssDNA track (1A, 2A and 3A) further delayed Rep movement with respect to the previous one. On the other hand, in terms of the translocation speed, the effect of the 3 neighboring abasic sites (3Ac) was smaller than 3 non-neighboring abasic sites (3A). PEG25 modification that has an equivalent contour length of 3 nts, had less effect than 3Ac. The pronounced effect induced by abasic sites shows that the local disruption of base-protein interaction is enough to stall the translocation, but not to stop its movement completely.

2.5.2 Conclusion and discussion

Single molecule observations yielded that abasic sites and PEG linkages prolonged the translocation; however these sites did not block the Rep movement. We have also shown that the prolongations were caused by stalling at the irregularities rather than an overall slowing of the helicase molecule. Single abasic nucleotide caused an average stall that lasted shorter than

one caused by three consecutive ones. However overall effect of the three interspersed abasic nucleotides was more than three adjacent ones. Absence of a single nucleobase (abasic site) seemed to have greater stalling effect then PEG linkages. These results also point to the robustness of the Rep translocation despite the mutations to the DNA structure. Along with the variety of DNA substrates Rep can shuttle repetitively, its ability to overcome damages to ssDNA track emphasizes the importance of translocation and fills a gap of information in this regard for helicases.

2.5.3 Materials and Methods

2.5.3.1 Experimental assay and smFRET TIRF

Glass cover slips and quartz slides were coated with PEG polymer brush as described (*78*). Approximately 1% of the PEG chains are biotin functionalized. Reaction chambers are constructed by using double sided tape to form channels on the quartz slide and closing the cover slip on top of the tape (*79*) . For surface tethering of partially duplex DNA molecules, one end of the substrate is biotin conjugated. DNA substrates were attached to surface using biotin/neutravidin scheme. After flushing the excess DNA, 100-500 pM Rep helicase is injected in reaction buffer that contains oxygen scavenging system and blinking suppression agents BME or trolox [ref] or trolox [ref] (10mM Tris pH7.6-8.0, 10mM $MgCl_2$, 15mM NaCl, 10% (v/v) glycerol, 0.4% glucose, 1mM ATP, 1% (v/v) gloxy, 140mM BME or 2mM trolox).

After these chamber preparation stage, the slide is placed on upside down prism type TIR microscope (Olympus X71 inverted microscope), in coverslip down orientation. The excitation of donor Cy3 molecules is achieved by evanescent field of 532nm laser (Crysta Laser). Emission from donor (Cy3) and acceptor (Cy5) dyes are collected by oil immersion objective, filtered for excitation and separated by dichroic mirror to create an offset in imaging plane between two channels. Resulting side by side image of donor and acceptor emission wavelengths is focused on EMCCD (Andor Xion) camera for recording. Donor and acceptor signal from each spot are converted into time series by homemade IDL code. Corresponding FRET efficiency curves, and temporal measurements on the time series data is performed by homemade MATLAB code.

2.5.3.2 Rep protein and DNA substrates

We used single cysteine mutant of Rep is labeled with Cy3 at position 333 for modified DNA translocation experiments. Residue 333 of Rep is located at the leading end during the translocation for maximum FRET efficiency change during translocation. All the mutants were confirmed for their in vivo and in vitro activity. Partial duplex DNA substrates comprise a 18-bp double strand DNA region and a 3' polythymidine tail dT_{80}. Cy5 is located at the junction of duplexed DNA and ssDNA. Modified strands obtained from IDT DNA, Iowa City, Iowa and Tim Lohman group at Washington University. Fluorescently labeled and biotinylated DNA oligos were purchased from IDT DNA. Oligos were annealed by using slightly more molar ratio of non-biotinylated strand (1.2x) and slowly cooling from 90C to room temperature in a heating block.

2.6 Conformational dynamics of 2B subdomain during ATP Hydrolysis*

The conformational states of *E. coli* Rep helicase undergoing ATP hydrolysis while bound to a partial duplex DNA (pdDNA) were studied using single-molecule FRET(*128*). Crystallographic studies showed that Rep bound to single-stranded DNA can exist in open and closed conformations that differ in the orientation of the 2B subdomain. FRET measurements between eight Rep mutants, donor-labeled at different residues, and pdDNA, acceptor-labeled at the junction, were conducted at each of the four nucleotide states. The positions of donor-labeled residues, based on crystal structure, and FRET measurements between these donor molecules and the acceptor fluorophore at DNA junction were used to predict the most likely position for the DNA junction using a triangulation algorithm. These predicted junction positions are compared with the crystal structure to determine whether the open or closed conformation is more consistent with the FRET data. Unexpectedly, our data revealed that there are two distinct Rep-pdDNA conformations in the ATPγS and ADP states. The primary conformation is similar to that observed in nucleotide-free and ADP.Pi states, and the secondary conformation is a novel conformation where the duplex DNA and 2B subdomain moved as a unit by 13 Å relative to the rest of the protein.

* The content in this subsection was reproduced with permission from "Single-Molecule Nanopositioning: Structural Transitions of a Helicase-DNA Complex during ATP Hydrolysis", Balci, H., Arslan, S., Myong, S., Lohman, T. M. & Ha, T.. *Biophys J*, **101**(4), 976-984 (2011). Copyright (2011) Elsevier.

2.6.1 Triangulation using FRET Nano-positioning System

We used prism-type total internal reflection microscopy (see Figure 2.18 a and Materials and Methods for details of the experimental assay) and smFRET (*112*) based triangulation. Nucleotide-free, ATPγS, ADP-vanadate (ADP.Vi), and ADP states were studied as ATP hydrolysis intermediates. ATPγS is a slowly hydrolysable ATP analog that mimics the ATP state, and ADP.Vi mimics the ADP.Pi state (*129, 130*). Eight Rep mutants, each labeled at a different site with a single Cy3 fluorophore (see Figure 2.17 b), and a pdDNA labeled with a single Cy5 at the duplex-single strand junction were used (see Materials and Methods for DNA sequence and fluorophore location). Labeling sites were distributed on Rep as shown in Figure 2.17 b: two residues within the 1A subdomain (43 and 233), one residue within the 1B subdomain (97), three residues within the 2A subdomain (310,316, and 333) and two residues within the 2B subdomain (473 and 486). All eight single cysteine mutants are functional in vivo and in vitro (*72*). Initially, we studied binding of Rep with pdDNA of various tail lengths to find an optimal tail length (Supporting Materials). A pdDNA having an 18-bp duplex and a 16 nucleotide 3' ssDNA tail (referred to as pdT16) was selected as the optimal substrate. Even though this tail length, 16 nt, is long enough to in principle accommodate two Rep monomers, the low protein concentration (1 nM) makes this unlikely. In addition, all traces were individually screened and those that showed more than one donor photobleaching step were excluded from analysis, ensuring binding of a single protein to DNA. Overall, eight Rep mutants and one DNA construct were studied in four different nucleotide states, resulting in 32 different FRET measurements.

Our analysis has two sets of input; donor fluorophore locations determined by Rep crystal structures and the measured efficiencies of FRET between these donor fluorophores and the acceptor fluorophore. These two inputs are used to determine the location of the acceptor fluorophore, which represents the location of the DNA junction, relative to Rep. We can also determine if the results are more consistent with the open or closed conformations using quantifiable metrics. The consistency check of the FRET data with the open or closed conformation is based on the agreement between the predicted junction position from triangulation analysis and the expected location based on the crystal structure. This procedure was repeated for all four ATP hydrolysis intermediates.

Figure 2.18 b shows an example time trace of fluorescence intensities of the donor and acceptor signal for Rep labeled at residue 333 in the nucleotide-free state. Most traces have

constant FRET efficiencies and do not show transitions to other FRET efficiencies. Figure 2.19 shows the FRET efficiency histograms, each representing several hundred Rep-pdDNA complexes, for Rep labeled with a Cy3 at a different residue and DNA labeled with a Cy5 at the junction. Each row represents a different Rep construct (in Rep43, Rep97, Rep233, Rep310, Rep316, Rep333, Rep473, Rep486 order from row A to row H, respectively), and each column represents a nucleotide state (in the nucleotide-free, ATPγS, ADP.Vi, and ADP order from the first to fourth column). The peak positions are summarized in Table 2.1. The histograms showed distributions that could be described with one or a sum of two Gaussian functions. In the nucleotide-free state, 7 out of 8 Rep constructs and in the ADP.Vi state all 8 constructs showed a single peak. In contrast, 5 out of 8 constructs showed two peaks for both the ATPγS and ADP states, indicating at least two different conformations. Single molecule time traces showed occasional transitions between the two states, which were more frequent during the initial binding of the protein to the DNA, i.e. the protein binds in one conformation and transitions into the other conformation after a short dwell (<1 sec). Examples of these transitions are shown in the Figure 2.24. Interestingly, Rep labeled on either residue within the 2B subdomain (473 and 486) shows a single peak in both ATPγS and ADP states (Table 2.1). Therefore, in the two conformations, the 2B subdomain may have the same relative position against the partial duplex junction. Because the crystal structures of UvrD and PcrA bound to a 3'-tailed pdDNA showed a 2B-dsDNA interaction, we suggest that the two conformations we observed here involve a coordinated motion of 2B and dsDNA relative to the rest of the protein composed of the 1A, 2A and 1B subdomains.

The single molecule measurements made here revealed multiple conformations which had not been observed previously. However, the presence of multiple conformations presents a technical challenge, especially for the triangulation analysis because in principle there are 32 ($=2^5$) possible ways of linking two different FRET states for each of the five Rep constructs that showed two populations. Presently, there is no generalizable solution for this problem. In an earlier study, Choi et al (*131*) chose the FRET state with the dominant population for each construct to perform triangulation for the major population. However, in our case, the two FRET states show similar populations for most constructs. Instead, we assumed that one of the two conformations in the ATPγS or ADP states is similar to that of the ADP.Vi state (primary conformation). With this assumption, we picked the FRET states with FRET values closer to those of the ADP.Vi values and grouped them into Set 1. The remaining FRET states were grouped into

Set 2. The residues that have a single peak are grouped in both sets. For example, for Residue 43 ATPγS state, the two peaks are 0.48 and 0.78. The ADP.Vi peak for Residue 43 is 0.52. Since 0.48 is closer to 0.52, it is grouped into Set 1 and 0.78 is grouped into Set 2. Hence, Set 1 for ATPγS state is (in increasing residue number order): (0.48, 0.65, 0.53, 0.66, 0.81, 0.70, 0.80, 0.79) and Set 2 is: (0.78, 0.89, 0.77, 0.66, 0.55, 0.85, 0.80, 0.79). Similarly, Set 1 of ADP is: (0.50, 0.69, 0.51, 0.59, 0.79, 0.71, 0.74, 0.81) and Set 2 is: (0.72, 0.94, 0.71, 0.59, 0.57, 0.89, 0.74, 0.81).

Triangulation analysis was then applied to each set to determine the junction position relative to the protein using the method developed by Muschielok *et al.* (*132*). The method is called FRET Nano-Positioning System (FRET NPS) and it employs a Bayesian algorithm to calculate the most likely position, e.g. the positions at which the probability distribution has the maximum value, and the probability distribution of possible positions. In our case we used eight FRET values and FRET NPS to determine the position of the acceptor fluorophore, hence DNA junction, and the uncertainty associated with this position. The uncertainty arises from several factors including the uncertainty in the measured FRET values, the available volume for the fluorophore on the protein surface, the length of the linker connecting the fluorophore to the protein or to the DNA, and the distribution of the Förster radius. To estimate the uncertainty, we modeled each donor position based on the Rep crystal structure to determine the available volume for each fluorophore. In addition, we used an uncertainty of ±0.05 for all experimental FRET efficiency values, and modeled the distribution of the Förster radius for each donor-acceptor pair based on anisotropy measurements of the corresponding labeling site on the protein. Finally, we used a linker length of 5 Å for both the Cy3-Rep linker and the Cy5-DNA linker.

The results for Set 1 of the ADP state are shown in Figure 2.20a when the Rep structure in the closed conformation is used for the analysis (white mesh), in Figure 2.20b for the open conformation (cyan mesh) and in Figure 2.20c for the case in which the 2B residues were excluded from the analysis (orange mesh). The mesh structure represents an envelope of possible locations determined using an isovalue of 1.0 which corresponds to the smallest possible envelope that encloses the sought position with 68% certainty. Imposing a smaller isovalue results in a larger envelope as higher certainty is required. Due to the difficulty in conveying the 3D distribution as a 2D projection in a figure, we refer the reader to 360° videos in Supporting Material.

The common characteristic of all the nucleotide states is that when the 2B residues are not included in the analysis, the probability density spans a broad arch as in Figure 2.20 c. Including the 2B residues in the closed or open conformations localizes this arch into a single small pocket (as in Figure 2.20 a) or into two small pockets (as in Figure 2.20 b). If we then increase the isovalue, both envelopes shrink but one disappears earlier than the other. What remains as we increase the isovalue is considered the most likely pocket. In this way, we determined the most likely position for all nucleotide states and conformations (Figure 2.21 a-b and Table 2.2). As detailed at the end of this section, the mesh size can be taken as an indicator of resolution and the most likely positions are used to infer whether our data is more consistent with the open or closed conformation as we discuss next.

2.6.2 Comparison to Open and Closed Conformations

We performed two related analyses in order to decide whether the open or closed conformation is more consistent with our data. The first analysis is based on the distance between the most likely junction position (Table 2.2) and the 5'-end of the ssDNA in the crystal structure of Rep. These distances are given in the third column of Table 2.2. Because our DNA is a 3'-tailed partial duplex, we expect that the partial duplex junction will be close to the 5' end of the ssDNA seen in the crystal structure of Rep. The junction positions predicted by the closed conformation are concentrated in a region 19 Å away from the 5'-end of the ssDNA. On the other hand, the junction positions predicted by the open conformation are populated in two regions both of which are about 32 Å away from the 5'-end of ssDNA. The 19 Å deviation in the closed conformation is within the uncertainly dictated by the dsDNA diameter and the linker.

In the second analysis, the junction position determined for the closed conformation of Rep is compared with that obtained in the crystal structures of UvrD-pdDNA complexes. To facilitate this comparison we mapped the UvrD-pdDNA complex onto the Rep-ssDNA complex based on the alpha-carbons in the 1A and 2A subdomains of the two proteins. Figure 2.22 shows this mapping, along with the predicted fluorophore position from our triangulation analysis. UvrD is shown in gray, the pdDNA in magenta, the predicted position for the ADP.Vi state in white mesh, and the Rep domains are shown in the same colors we used in the other figures. As can be inferred from Table 2.2 the most likely positions for the other nucleotide states of the closed conformation would also occur in similar positions. As shown in Figure 2.22, there are remarkable structural similarities between the two proteins. In fact, even such a simple mapping

results in an almost identical path for ssDNA within the two helicases (data not shown). The ssDNA-dsDNA junction position, assumed to be the center point of the 20 Å wide dsDNA is 16 Å from the predicted position for the closed conformation. Given that the fluorophore is attached to the junction via a 5 Å long linker and that we are comparing two similar but different helicases, the 16 Å distance between the two positions is an acceptable result. A similar analysis using the open conformation data gave significantly larger deviations ranging between 25-44 Å for different nucleotide states, as given in the last column of Table 2.2. This analysis further supports the better consistency of the closed conformation with our data.

2.6.3 Multiple Conformations for ATPγS or ADP

An interesting and unexpected finding of this study is that we observe multiple conformations of the Rep-DNA complex in the ATPγS and ADP-bound states. Here, we consider these multiple conformations in more detail, focusing on the closed conformation as it is more consistent with our data. The predicted junction positions for Set 1 of the ATPγS data and Set 1 of the ADP data are identical to each other and are in close proximity, within 6 Å, to the positions predicted by the nucleotide-free and ADP.Vi states (Table 2.2). This is expected because of the choices made in selecting Set 1. The positions predicted by Set 2 of the ATPγS data and Set 2 of the ADP data are also almost identical to each other (2 Å deviation), but are significantly different, by about 15 Å, from the positions that are prevalent in the other nucleotide states. Furthermore, the junction positions predicted by Set 1 and Set 2 differ by about 15 Å, which is much larger than what one would expect from rearrangement of the two RecA-like domains of Rep during ATP binding and hydrolysis. Therefore, these two cases should be considered different conformations of the Rep-pdDNA complex and they will be referred to as the primary (result of Set 1 peaks) and secondary (result of Set 2 peaks) conformations. Interestingly, the residues in the 1A, 1B, and 2A subdomains all show double FRET peaks when bound to ATPγS and ADP whereas the residues within the 2B subdomain show only single FRET peaks in these nucleotide states (Table 2.1). We interpret this to indicate that the 2B subdomain moves in synchrony with the duplex DNA, and hence has the same relative distance to the junction in both conformations, resulting in a single FRET peak. On the other hand, the other subdomains move relative to the duplex DNA in the primary and secondary conformations.

The different junction positions predicted for the primary and secondary conformations are also reflected in the distances of these positions to the 5'-end of ssDNA or to the junction

position predicted from Rep UvrD mapping (last two columns of Table 2.2). In addition to these quantitative differences, the positions predicted by the secondary conformation (Set 2) in the ATPγS and ADP states are buried within the 1B subdomain of Rep, which is not an acceptable position for the junction. Therefore, the secondary conformation is likely to be significantly different from what is observed in the crystallographic structures. One possibility is that the 2B subdomain itself may have an orientation distinct from both the open and closed conformations. A clash between the predicted junction location and the 1B subdomain also exists for the primary conformation but at a much smaller extent. In primary conformation case, the most likely position is outside the 1B subdomain and does not clash with it, even though the mesh volume partially overlaps with 1B subdomain (Figure 2.20 a). Therefore, the primary conformation we deduced here is consistent with the closed conformation of Rep observed in the crystal structure.

2.6.4 Discussion

The surprising finding of two FRET populations for the 1A, 1B, and 2A subdomain residues in the ATPγS and ADP states has been interpreted as multiple conformations in these nucleotide states. The single FRET populations in the 2B residues (473 and 486) have motivated the interpretation of a coordinated movement of the 2B subdomain with the duplex DNA. As the acceptor fluorophore is located at the ss/dsDNA junction, alternative scenarios including junction dynamics or dsDNA melting at the junction should be discussed as alternative explanations of our data. The first scenario to consider is whether junction dynamics, e.g. junction stabilizing in two different locations or transitioning between two different locations while Rep remains essentially static, could be the reason for multiple FRET peaks. In this scenario, as Rep is static while the junction is moving, we would expect multiple FRET populations for all residues regardless of where on Rep they are located. However, our data shows multiple populations for 1A, 1B, and 2A residues, while the 2B residues demonstrate a single population, making this scenario unlikely.

Another scenario to consider is whether different levels of Rep-induced duplex junction melting in different nucleotide states could result in multiple FRET peaks or influence the FRET results in significant ways. Numerous single molecule measurements have been performed on Rep-pdDNA complexes and junction melting due to Rep monomer binding has not been observed to take place at detectable levels (*58, 72*). In addition, crystal structure studies on

UvrD and PcrA in complex with pdDNA and different ATP hydrolysis intermediates do not show any junction melting upon helicase binding. Even if we consider melting of a base pair at the junction, the junction would be displaced by about 3 Å from the Rep helicase. This would not be adequate to explain the 13Å separation we observe between the primary and secondary conformations. In fact, 13Å separation would require 4-5 base pairs to unwind upon Rep binding, which has not been observed before. Finally, junction melting would affect all the Rep mutants similarly, regardless of where the fluorophore is, hence the different behavior for the 2B and non-2B residues cannot be explained by this scenario either.

Another possibility to consider is whether partial occupancy of ATPγS, i.e. binding to some Rep monomers but not others, could be the reason for two conformations in the ATPγS states. As we demonstrate in Supporting Materials Figure 2.25 10 μM ATPγS is enough to saturate Rep binding, and at the experimental concentration we use, 500 μM ATPγS, all Rep molecules should be bound by ATPγS. A related scenario that could be considered is whether slow ATPγS hydrolysis gives rise to two different conformations. Single molecule traces do show occasional transitions between the two FRET states for some Rep constructs (Rep 43, 97, 233, 316, and 333), as demonstrated in Figure 2.24, but not for all of them (not observed for Rep310, 473, and 486). If ATPγS hydrolysis were the reason for the observed dynamics, we would have expected to observe it for all constructs. Hence, we do not think this scenario is consistent with our data.

An important question that should be considered is what the potential biological significance is for having two conformational states in ATPγS and ADP states while a single conformation is observed in the ADP.Vi state. Because the ATPase reaction would proceed from the ATPγS state to the ADP.Vi state and then to the ADP state, we can speculate that the ATPase reaction may proceed in two pathways upon ATP binding. The two pathways would then merge upon ATP hydrolysis and then diverge again when Pi is released. These multiple pathways may provide a mechanism of regulation of helicase activity, depending on the functional context. If multiple pathways and conformations already exist in the enzyme-DNA complex, one can imagine that changes in the environment or functional context can enable the selection of one of the existing conformations without having to induce a new conformation.

The resolution of our method can be estimated from the size of the mesh structures, i.e. size of one of the pockets of probability distribution in Figure 2.20 a. Quantitatively, the

resolution is characterized by the standard deviation of the most likely position of the probability distribution. For the closed conformation a single pocket of probability distribution is obtained in 5 out of 6 cases, whereas this number is 2 out of 6 for the open conformation. This is another reason why the open conformation is less likely because it resulted in a broader probability distribution compared to the closed conformation. In all single pocket cases of the closed conformation, the mesh structures have a standard deviation of about 10 Å which corresponds to a standard deviation of about 6 Å in x, y, and z directions.

Finally, we can examine the consistency of the triangulation analysis by including different numbers of residues in the analysis using the ADP state as an example. The standard deviation in the most likely position of the junction gradually increases from 8 Å to 12 Å when the number of residues was decreased from eight to three, about 1 Å per eliminated residue. However, using five residues or less resulted in a second pocket appearing in the probability distribution for all studied cases. Not including both 2B residues in analysis was an exception as it resulted in an abrupt decrease in resolution, which we believe is due to non-uniform distribution of the labeled residues. On the other hand, the most likely positions did not change significantly regardless of the number of constraints used and were all within 8 Å each other, demonstrating the robustness of the analysis and consistency of the data.

2.6.5 Conclusion

This study has several important conclusions for the conformational state of Rep helicase when in complex with pdDNA and an ATP hydrolysis intermediate (nucleotide-free, ATPγS, ADP.Pi and ADP studied). First of all, our data is more consistent with Rep being in a closed conformation with respect to its 2B subdomain, than the open conformation, for all ATP hydrolysis intermediate states. This conclusion is also consistent with the crystal structure studies on related helicases such as UvrD and PcrA. However, we show that in the ATPγS and ADP bound states the Rep-pdDNA complex show evidence of at least two conformations whereas the crystal structures of UvrD showed only one conformation. Hence, studies of static crystal structures do not capture the fully dynamic aspects of this complex under these conditions. During this reconfiguration between the two conformations, the 2B subdomain rotates with the duplex DNA. This may result in a tension change between the single stranded portion and the enzyme's motor domains (2A and 1A) although at present we do not know the functional roles of these movements. A speculative biological significance of having multiple

conformations could be the possible ability to select one pathway over another, which could in turn regulate helicase activity, depending on the environmental conditions or functional context, without having to induce new conformational states. Finally, we demonstrate that FRET based triangulation analysis is capable of providing 10 Å resolution in absolute position, which should be enough to address many other biological questions. In particular, we believe a triangulation analysis would be particularly powerful in structural problems involving multiple biological entities (*131*), e.g., dimerization conformation of two proteins or assembly conformation of a multi-protein complex on a DNA substrate, since such problems could be very difficult to address using more conventional methods such as X-ray crystallography.

2.6.6 Materials and Methods

2.6.6.1 DNA substrate and Labeled Rep Mutants

The DNA substrates were formed by annealing ssDNA of sequence 5'-Cy5-GCCTCGCTGCCGTCGCCA-Biotin-3' with a ssDNA of sequence 5'-TGGCGACGGCAGCGAGGCTTTTTTTTTTTTTTT-3'. The ssDNA molecules were ordered from IDT DNA. A detailed protocol of protein purification and labeling has been published elsewhere (*72*) hence will not be detailed here. Briefly, the five native cysteines were replaced with C18L, C43S, C167V, C178A, and C612A, and a *"Cys-free"* mutant was made. Then eight different *"Cys-light"* mutants each containing a single cysteine were produced (pRepS43C, pRepA333C, pRepS316C, pRepA310C, pRepS233C, pRepA97C, pRepA473C, and pRepS486C). None of these eight sites are within the known helicase motifs, they are not shared among Rep, UvrD and PcrA, and they are well exposed on the surface of the protein making them ideal for labeling. These *Cys-light* mutants were then labeled by mixing them with excess amounts of Cy3 maleimide with about 90% efficiency. The ATP hydrolysis and DNA unwinding activities of the Cys-light proteins and labeled proteins was measured and found to be at least 75% of the wild-type proteins (*72*).

2.6.6.2 smFRET Assay

Single molecule measurements were performed using a prism type total internal reflection microscope (Olympus IX50) and an electron multiplying charge-coupled device camera (iXon DV 887-BI EMCCD, Andor Technology, CT). Both the quartz slide and the glass cover slip that formed the sample chamber were coated with biotin functionalized polyethylene glycol (PEG) in order to reduce non-specific binding of DNA and proteins. All single-molecule measurements were performed in a saturated Trolox (≈2 mM, Calbiochem) buffer containing an

oxygen scavenging system [0.1 mg/ml glucose oxidase, 0.02 mg/ml catalase (Roche, Indianapolis, IN), 0.8% (w/w) dextrose, and 1% v/v 2-mercaptoethanol], 10 mM $MgCl_2$, 15 mM NaCl and 0.1 mg/ml BSA. This buffer increases the photo-stability and suppresses the blinking of the fluorophores. 500 μM of ATPγS, ADP.Vi or ADP were used as ATP hydrolysis intermediates. Fluorescence images were acquired at an integration time of 100 msec. Movies of 700-1200 frames were recorded at 8-bit format and analyzed using custom programs written in Matlab. Finally, crystal structure data and results of our triangulation analysis were superimposed using the software Visual Molecular Dynamics (VMD, http://www.ks.uiuc.edu/Research/vmd).

2.6.6.3 Triangulation

Triangulation is a method in which the location of a point is determined by using the distance between this point and three other points. As in Global Positioning Systems, three distances are the minimum required constraints in order to reduce the possible locations to two points, one of which is typically not relevant to the system of interest. At the molecular level, FRET based triangulation has been used to obtain structural information in physiologically relevant solution conditions (*72, 132-134*). In our case, we measure the FRET efficiency between a Cy3 labeled residue on Rep and a Cy5 labeled nucleotide on DNA. Under these conditions, we typically get a distribution of FRET efficiency values, suggesting that a range of donor-acceptor separations is sampled for each protein-DNA pair. Multiple binding sites on DNA, the flexibility of linkers that attach the fluorophores to DNA or protein, and possible dynamics in the binding conformations of the Rep-DNA complex may all contribute to the distribution. Under such non-ideal conditions, using more than 3 constraints would make the analysis more rigorous and less prone to potential errors that might occur. In our case, we determine the position of a Cy5 acceptor by measuring the FRET efficiency between this acceptor and 8 independent donors, distributed on various residues on Rep. We used the FRET NPS program by Muschielok *et al.* for this analysis (*132*).

2.7 Figures

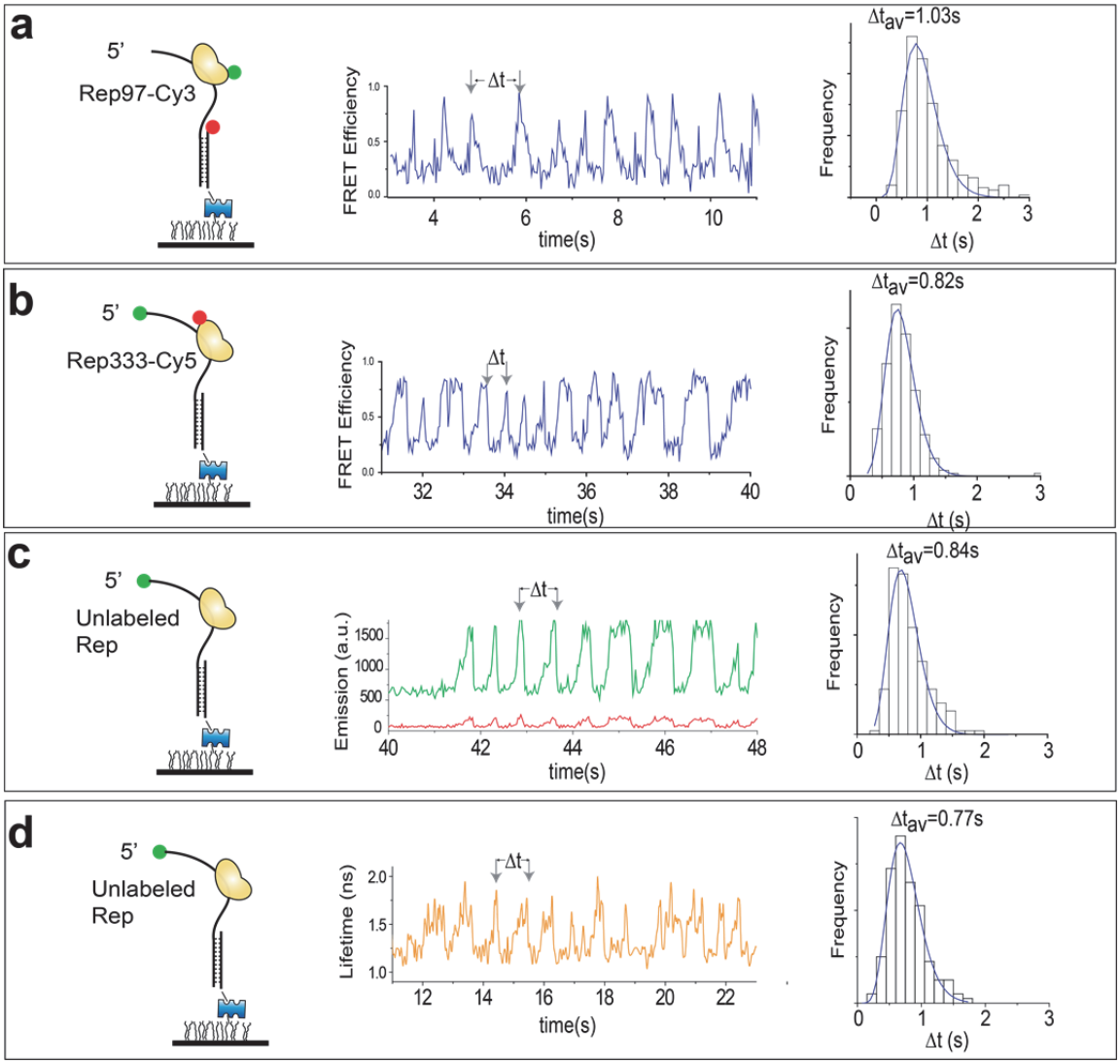


Figure 2.1 Rep helicase repetitively shuttles on 5' overhang DNA. Cartoons represent the experimental designs. Red and green balls represent Cy5 and Cy3. Rep molecules are shown in yellow. Histograms for the durations of the full translocation cycles, Δt, are plotted for each experiment. a. Cy3 labeled Rep and Cy5 label at the junction produced reverse sawtooth pattern (blue) in E_{FRET}. Reverse sawtooth pattern shows that the dye to dye distance gradually increases and abruptly comes back to original position b. Cy5 labeled Rep and Cy3 label at the 5' end yield regular sawtooth pattern. This is indicative of gradual decrease of distance between the dyes followed by sudden increase. c. Same as b except unlabeled Rep was used. Protein induced fluorescent emission profile of Cy3 shows Rep slowly approaches to 5' end and suddenly jumps back to the start position to repeat the process. d. Cy3 fluorescence lifetime is measured on a TCSPC setup. Repeating pattern has similar periods to those shown above.

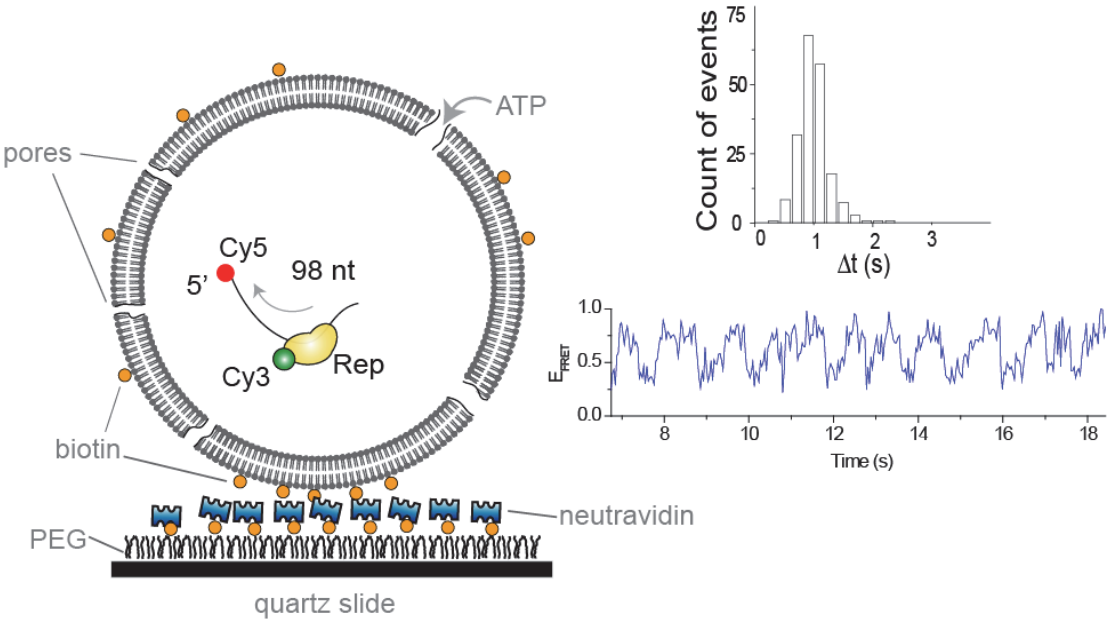


Figure 2.2 Encapsulation of Rep and ssDNA. Cy5 labeled free ssDNA and Cy3 labeled Rep is trapped in a vesicle. Rep can repetitively translocate on free ssDNA as shown in the fluorescence traces. Histogram shows the distribution of shuttling periods.

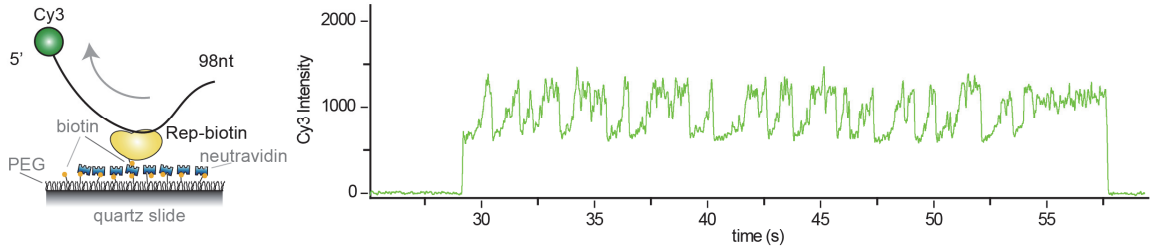


Figure 2.3 Repetitive shuttling on free ssDNA. Surface bound Rep binds and translocates repetitively on free ssDNA. Cy3 fluorescence trace shows the binding of DNA and repetitive shuttling indicated by gradual increase and sudden drops. Fluorescence enhancement of Cy3 occurs due to approaching protein.

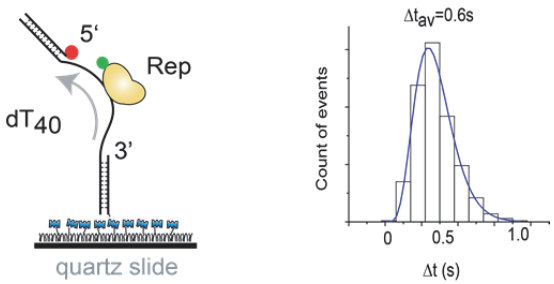


Figure 2.4 Repetitive translocation on ssDNA flanked by duplex (ds-ss-ds DNA). Cy3 labeled Rep exhibits repetitive shuttling on 40 nt ssDNA stretch in between the duplex DNA ends. Histogram shows the distribution of translocation periods.

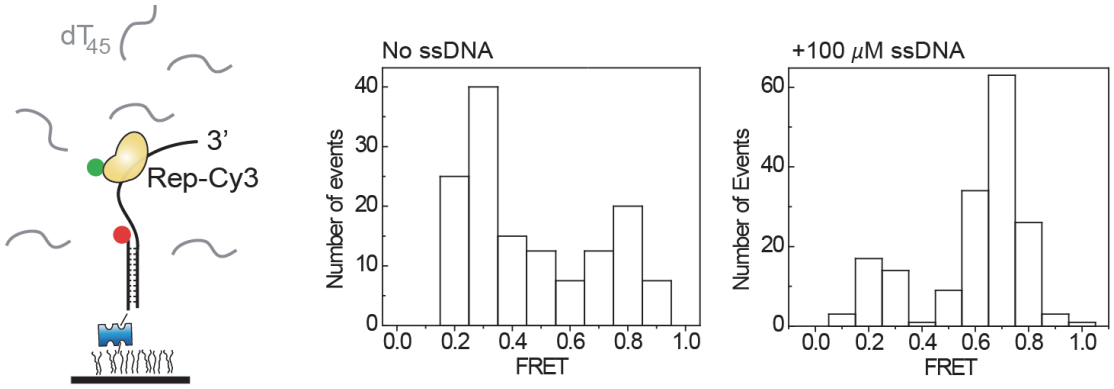


Figure 2.5 3' end recognition facilitates repetitive shuttling. Cartoon represents the experimental assay. 100 uM poly-dT ssDNA was injected during the repetitive translocation of Rep. The histogram on the left shows the distribution of dissociation FRET efficiency values when buffer with no DNA was flowed in. In comparison, the histogram on the right shows the same distribution after the buffer with 100uM ssDNA was injected. ssDNA clearly promotes Rep dissociation when Rep reaches the junction that coincides with the snap back events.

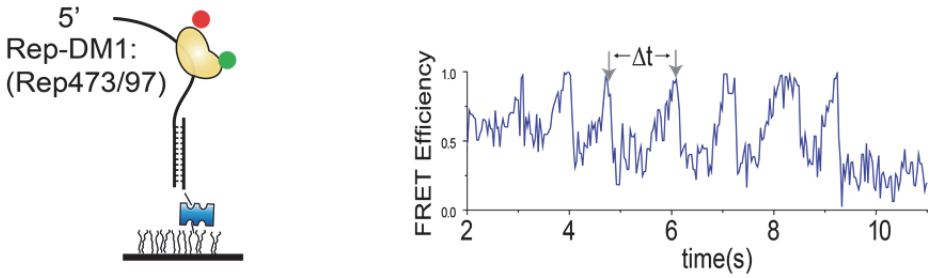


Figure 2.6 Conformational change during translocation 5'. Dual labeled Rep (DM1) translocates on unlabeled 5' overhang substrate. Gradual closing and sudden opening of 2B subdomain is evident in the FRET efficiency traces. Period of the cycles is similar to the translocation period.

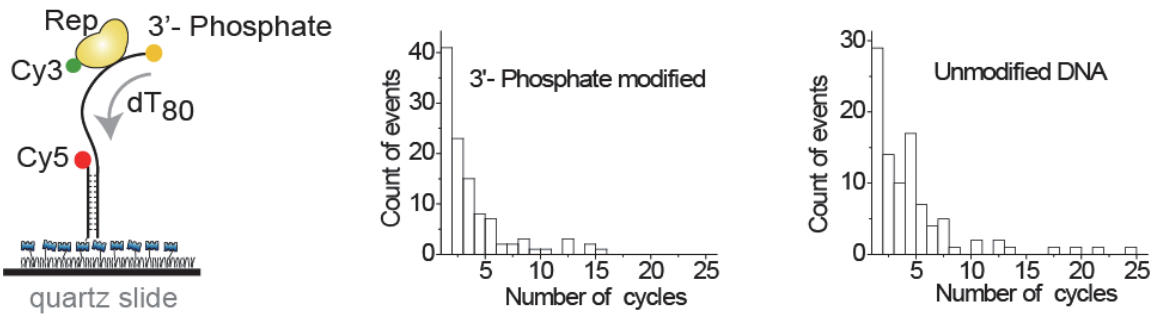


Figure 2.7 3'-Phosphate modification. 3'-OH is replaced with a phosphate group on this substrate. However distribution of number of consecutive shuttling cycles remains similar to that of unmodified substrate.

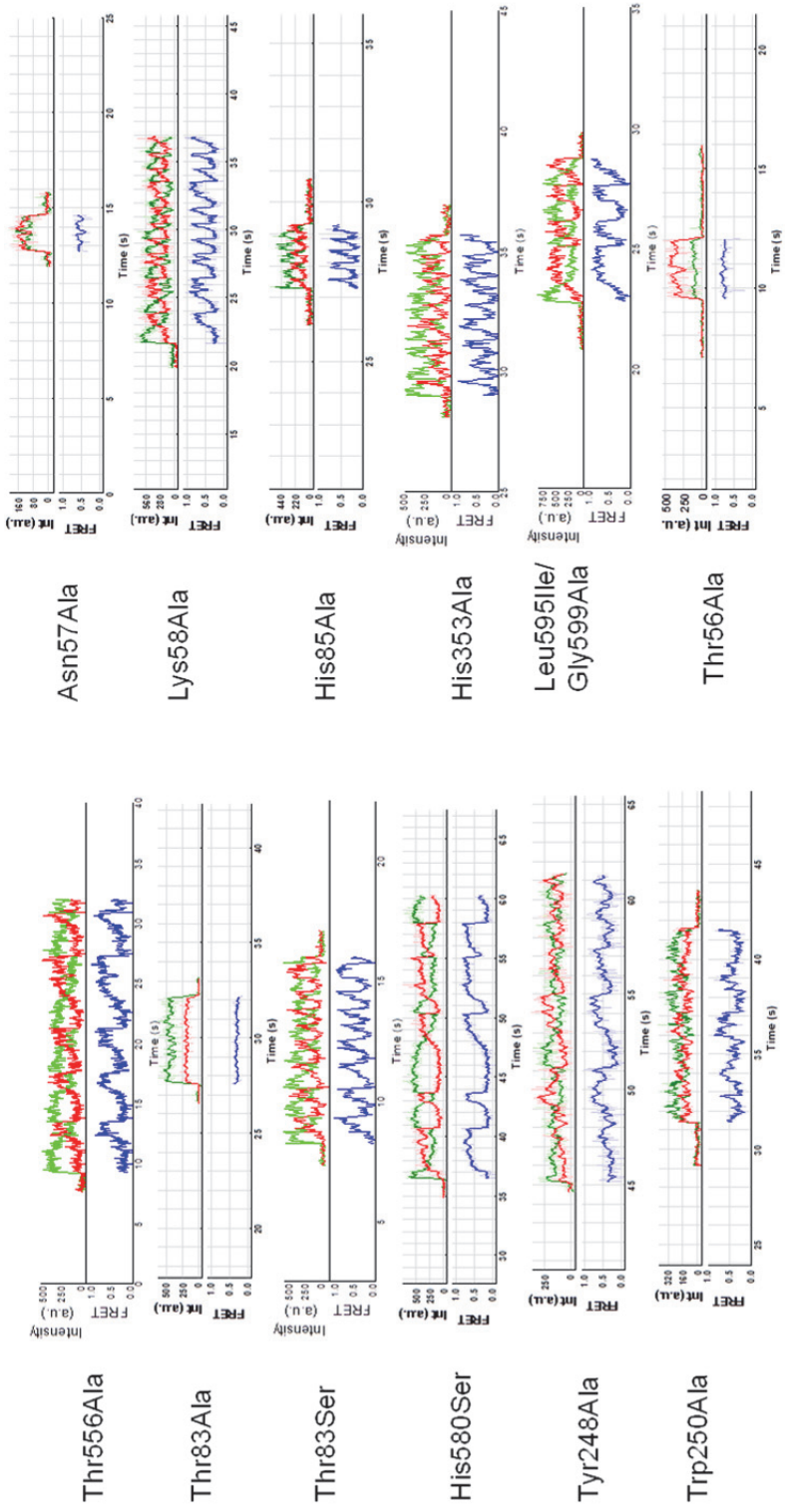


Figure 2.8 Representative repetitive translocation traces of biochemical mutants of Rep. Repetitive shuttling traces of Cy3 labeled mutants that were obtained on 3'-overhang pdT80 substrate show faster, slower and inhibited translocation characteristics.

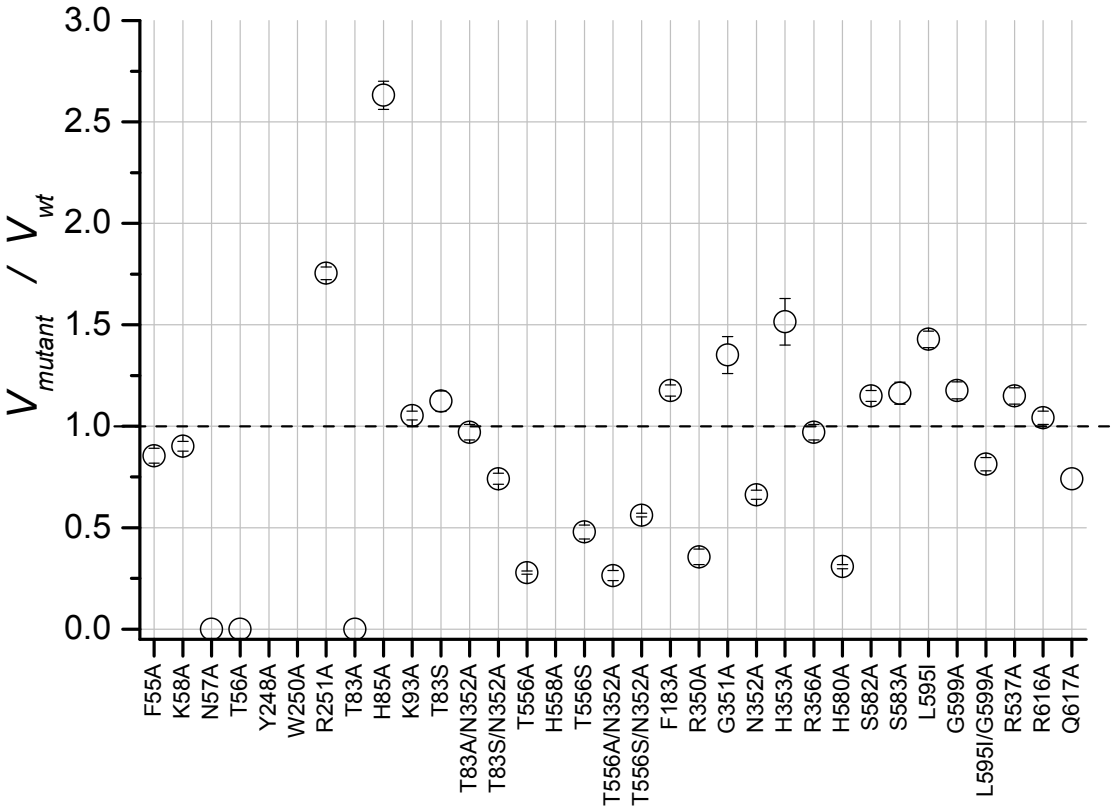


Figure 2.9 Translocation speeds of biochemical Rep mutants. Normalized translocation speeds of mutants are plotted. Speeds were measured on smFRET repetitive shuttling assay and normalized by the speed of the baseline Rep (RepA333C, a single cysteine wild type like mutant (*58*)).

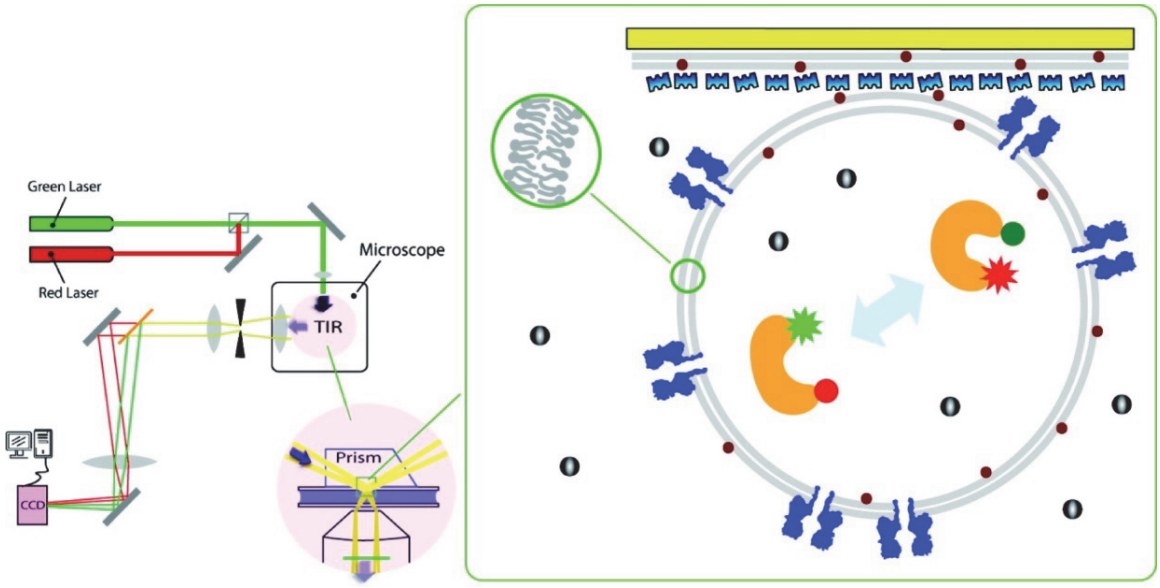


Figure 2.10 A diagram summarizing the experimental scheme (not to scale). (Left) A home-built prism-type total internal reflection microscope with two-color fluorescence detection. (Right) The encapsulation assay. A supported lipid bilayer (a single membrane layer is depicted as one gray stripe) is first formed on a quartz coverslide (yellow) as a cushion. SUVs are then immobilized on the lipid cushion through biotin (brown dots)–streptavidin or neutravidin (blue) linker. The aHL pores on the SUV are shown also in blue. Encapsulated inside SUVs is the biomolecule (orange), depicted in folded and unfolded conformations at two and eight o'clock positions, respectively. The molecule is labeled with a donor (green) and an acceptor (red) which serve as FRET probes. The star shape represents the brighter fluorophore. Black circles represent the chemical agent (e.g., Mg^{2+} or ATP) relevant to the reaction under study.

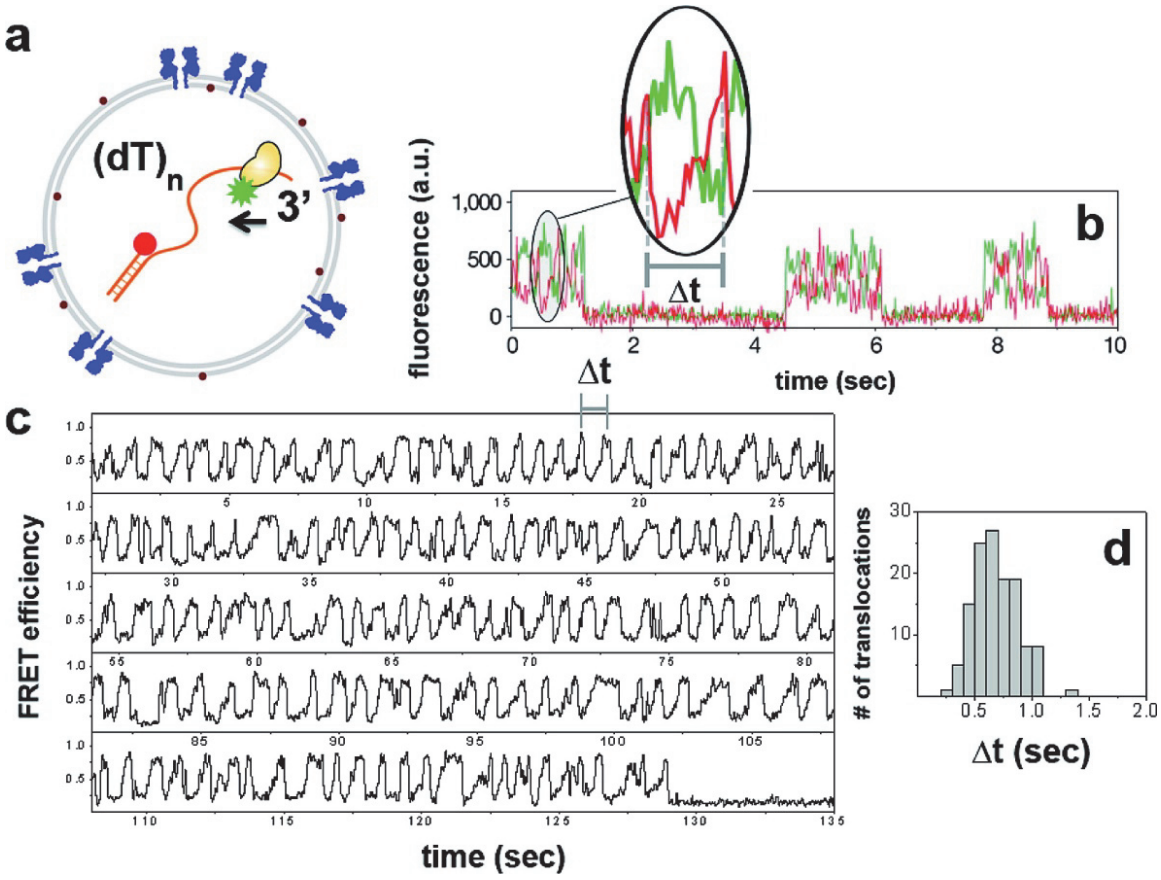


Figure 2.11 (a) Rep helicase labeled with donor, and DNA labeled with acceptor at the junction depicted as coencapsulated within a porous SUV. Partial duplex DNA is an 18-bp double strand carrying a $(dT)_{80}$ tail. (b) Rep shuttling data obtained from surface-tethered DNA (modified from ref. (*58*)) with a same length tail. Inset shows a single translocation event where the gradual increase in the acceptor signal (with accompanying decrease in the donor signal) reflects translocation of Rep toward the junction, whereas the abrupt drop marks the snapping back to the 3′ end. The translocation time (Δt) is defined as the period between consecutive snapping events. Individual cycles of shuttling separated in time are due to different Rep molecules from solution binding on the same surface-attached DNA and exhibiting limited repetitive translocations until dissociation. (c) FRET trace of Rep shuttling on a $(dT)_{80}$ DNA within 100 nm diameter porous SUVs. In marked contrast to the surface experiments, a single Rep−DNA pair shows over 140 translocation events until acceptor photobleaching near 130 s). Time resolution was 30 ms. (d) Δt histograms can be built from a single Rep−DNA pair. Similar statistics for surface measurements (b) can only be obtained by merging data from translocations events from many shuttling cycles exhibited by different Rep molecules.

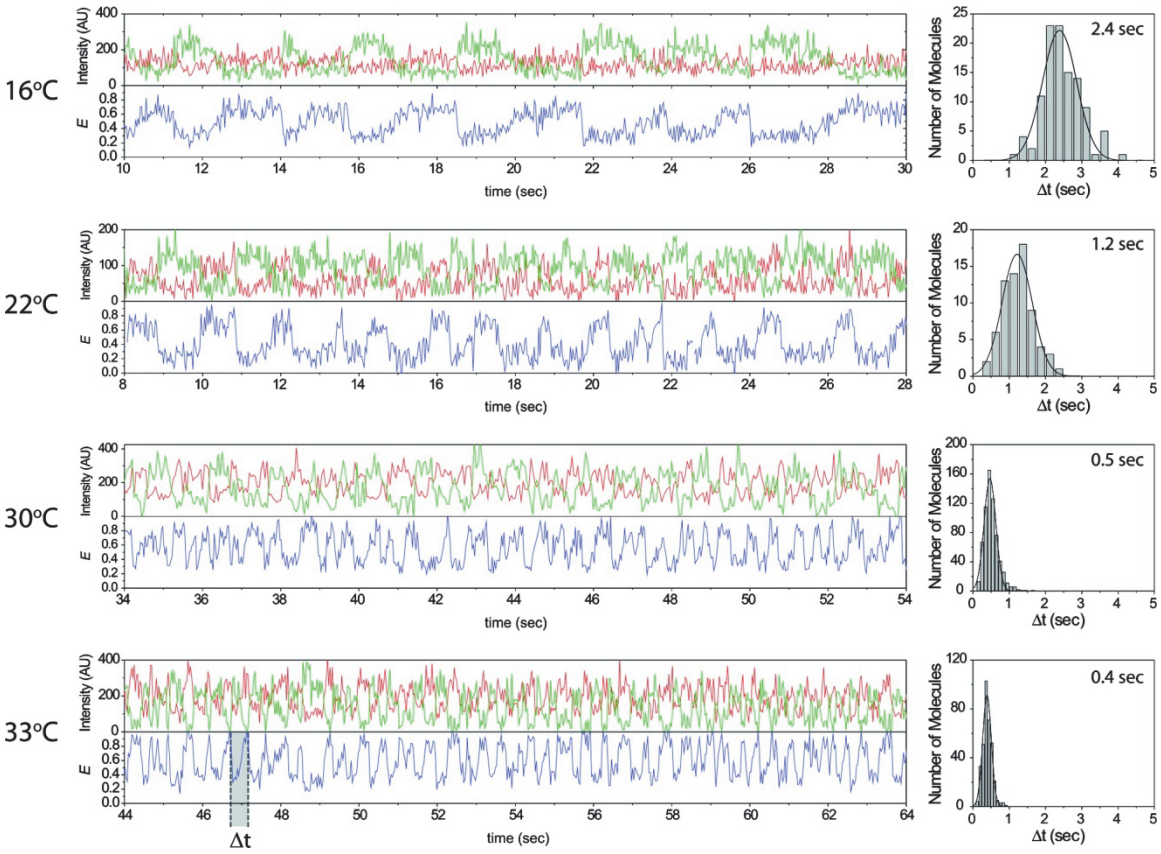


Figure 2.12 Rep helicase translocation inside a porous nanocontainer. Cy3 (donor) labeled Rep helicase and Cy5 (acceptor) labeled partial duplex DNA have been coencapsulated within eggPC/Chol vesicles. The data taken at 16, 22, 30, and 33 °C are shown at the top. From these time traces, the translocation period, Δ*t*, was measured and plotted as a histogram. Average Δ*t* values from Gaussian fits of these histograms are shown on the plot.

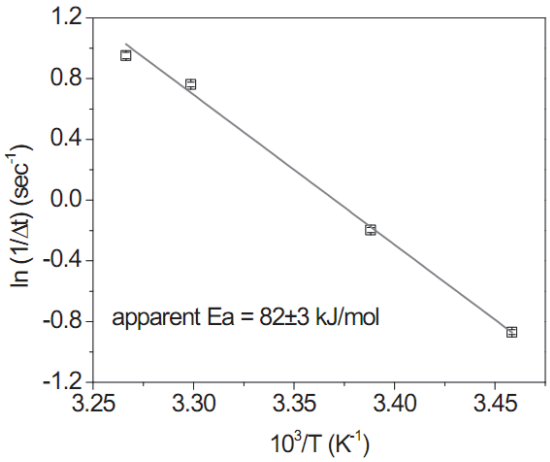
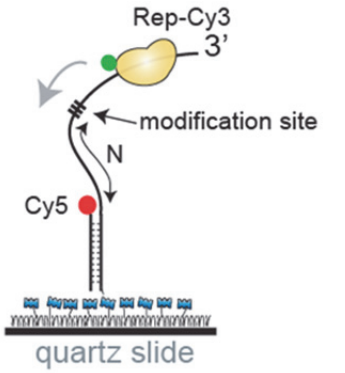


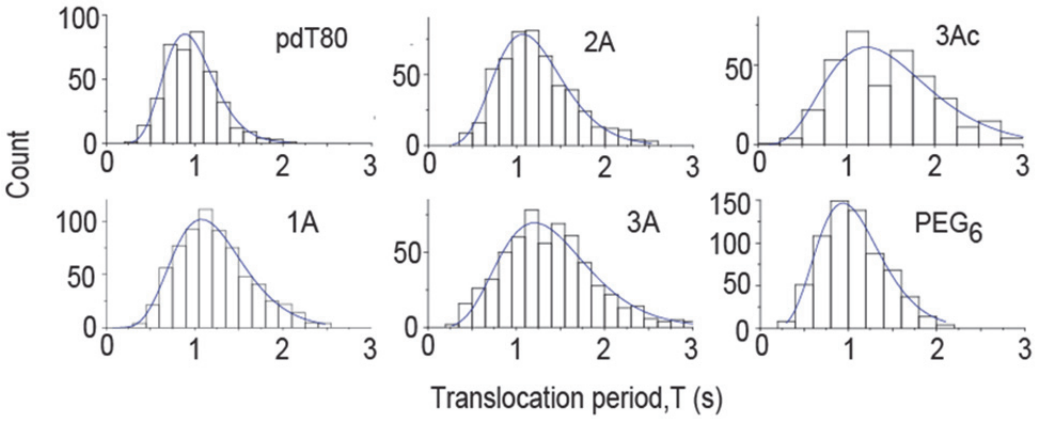
Figure 2.13 Rep translocation apparent activation energy. Arrhenius plot of Rep helicase translocation time (Δt) along partial duplex $(dT)_{80}$ and its fit are shown above. The data points have been fitted to a line and obtained the apparent activation energy of 82±3 kJ/mol.

A



Name	modification	N, position (nt)
pdT80	--	--
1A	Abasic	40
2A	Abasic	20,50
3A	Abasic	20,40,60
3Ac	Abasic	26,27,28
PEG25	PEG_6	26-28

B



C

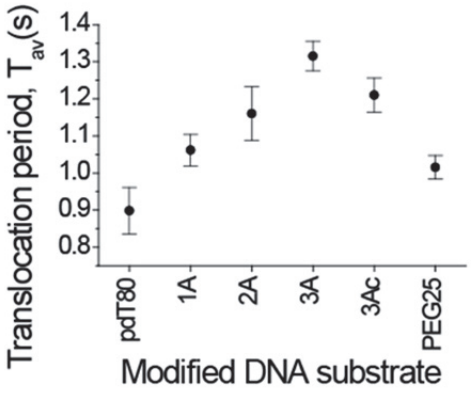


Figure 2.14 Effect of DNA modifications on Rep translocation. a) Cartoons represent the assay that a Cy3 (green ball) labeled Rep molecule is translocating on the 3' overhang of a partially duplex DNA labeled with Cy5 (red ball) and modification sites are listed in the table. B-C) Histograms of the shuttling period distributions and the plot of the average translocation time on modified DNA with equivalent 80 nt poly-dT length.

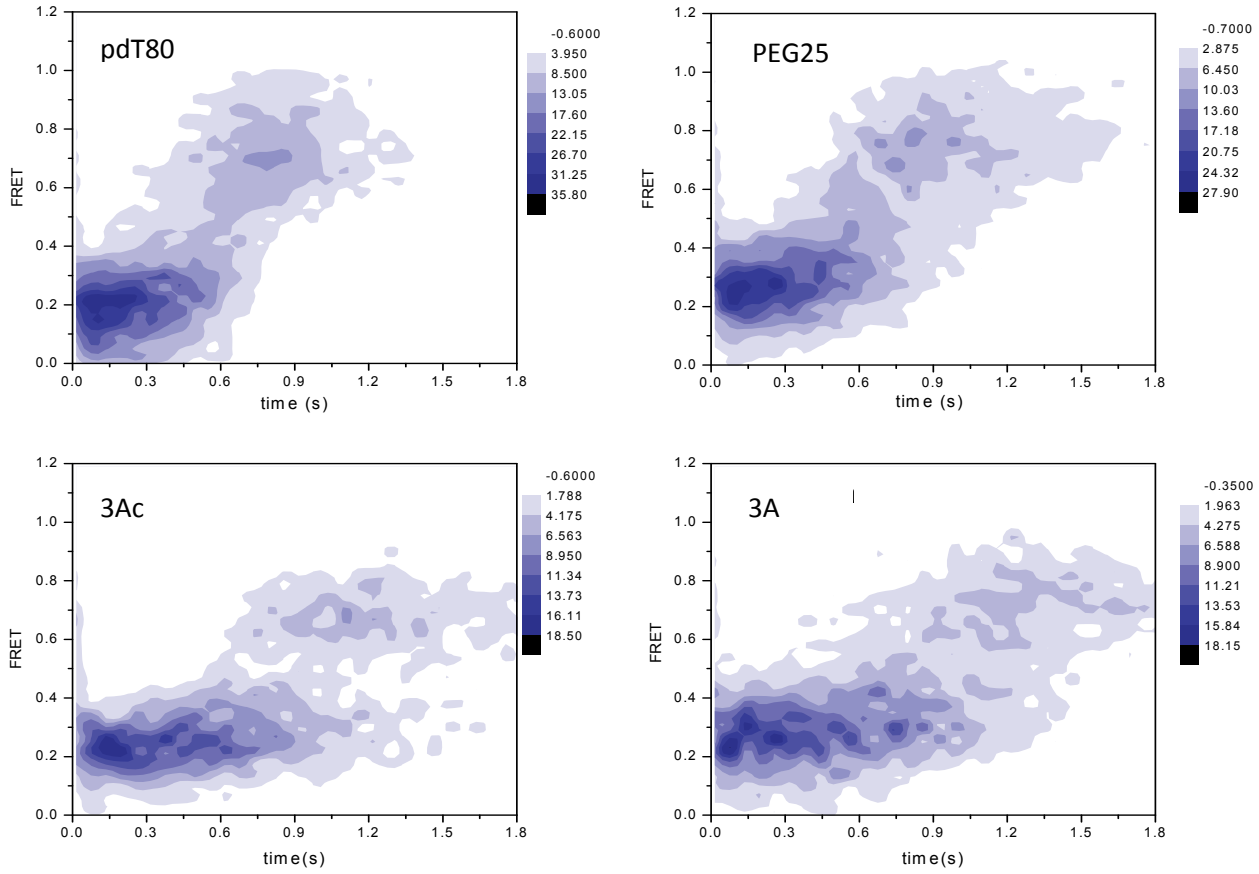


Figure 2.15 FRET traces of Rep translocation on modified DNAs. FRET efficiency traces from many translocation cycles superimposed and plotted as density maps for 4 different substrates. Start of each translocation cycle is synchronized. pdT80 is 18-bp DNA with non-modified dT_{80}-3' overhang. Modification nomenclature is listed in Figure 2.14.

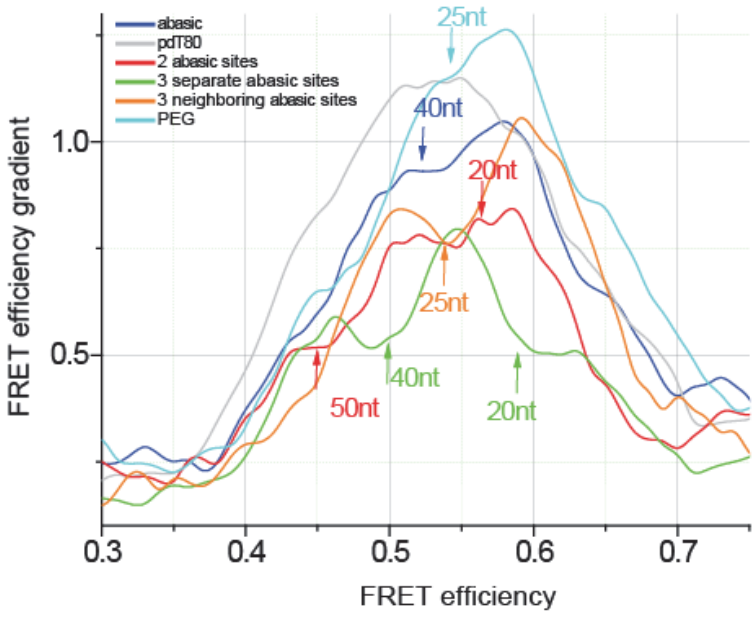


Figure 2.16 Gradient of FRET efficiency profiles. Ensemble average of time derivative of FRET profiles are plotted against the FRET efficiency. Local deviations from bell curve of non-modified DNA translocation (grey curve) means Rep stops at the deviation points momentarily, but continues afterwards.

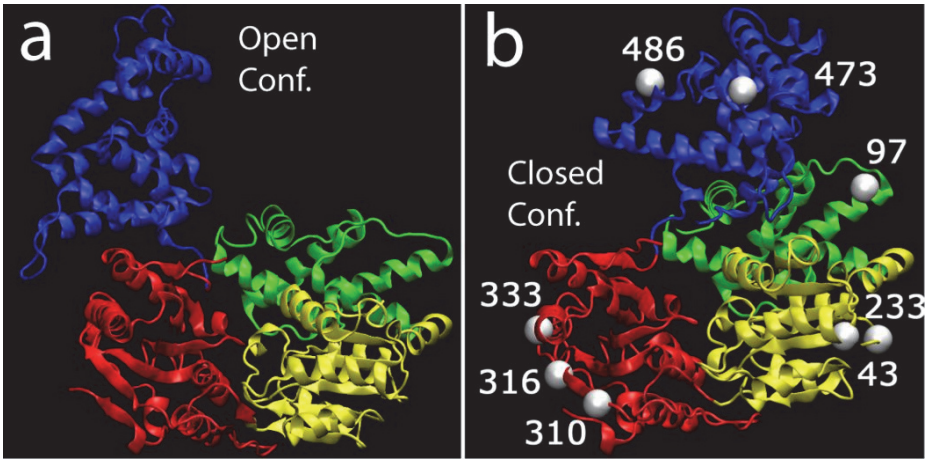


Figure 2.17 Rep structure (PDB ID: 1UAA and reference (11)) and labeling sites. Rep in (a) open conformation and (b) closed conformation with labeled residues shown as white spheres. Color coding of Rep subdomains: yellow 1A; red 2A; green 1B; blue 2B.

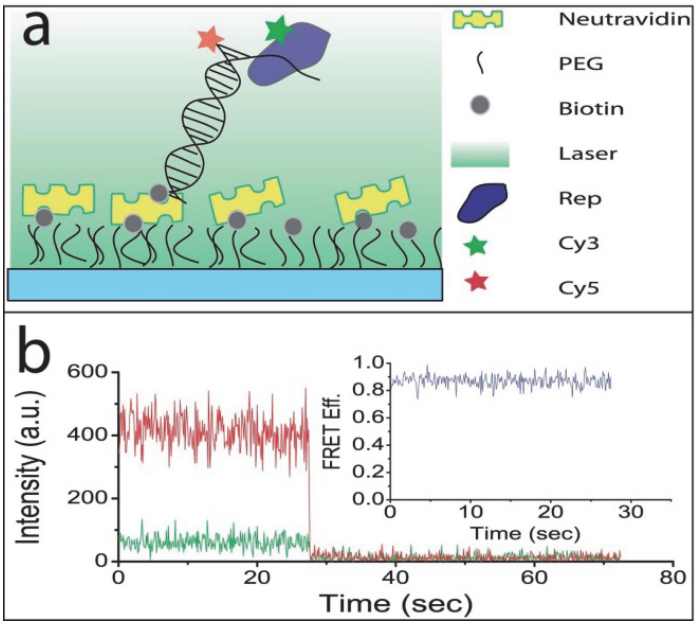


Figure 2.18 (a) Experimental configuration. (b) Representative time traces of a single Rep (donor-labeled) bound to a DNA (acceptor-labeled). Green line is the donor intensity, and the red line is the acceptor intensity. The donor photobleaches at time=28 sec. The inset shows the FRET efficiency.

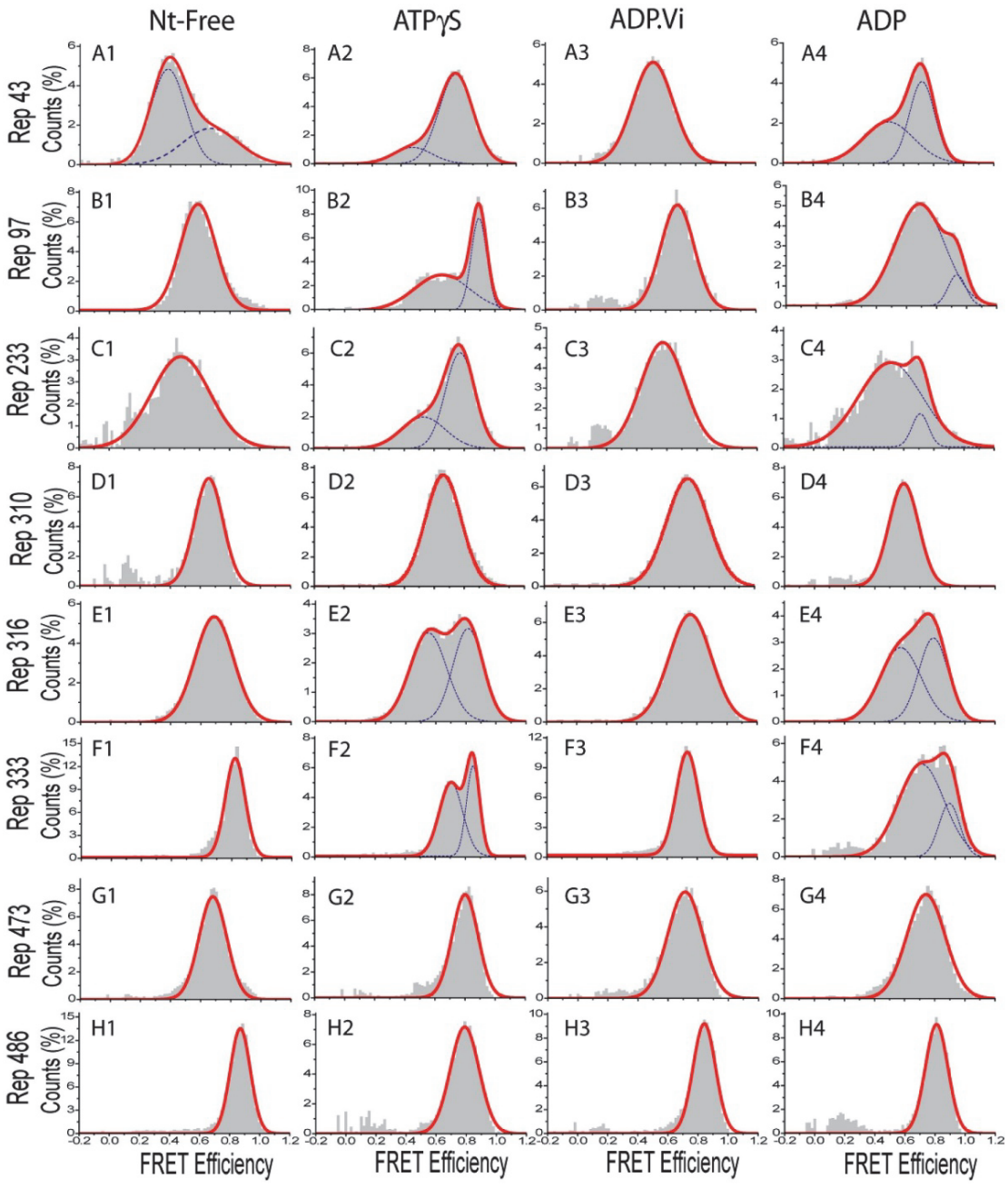


Figure 2.19 Histograms of single molecule FRET efficiency between Cy3-labeled Rep constructs, labeled at the indicated residue number, and Cy5-labeled DNA (label at the junction of a partial duplex DNA). Each row represents a Rep construct in the following order: Rep43, Rep97, Rep233, Rep310, Rep316, Rep 333, 473, Rep486 from row A to row H, respectively (see **Table 2.1** for the subdomain that each residue belongs to). The columns represent the nucleotide states in the following order: Nucleotide-free 1st column; ATPγS 2nd column; ADP.Vi 3rd column; ADP 4th column. Single or double Gaussian fits are also shown.

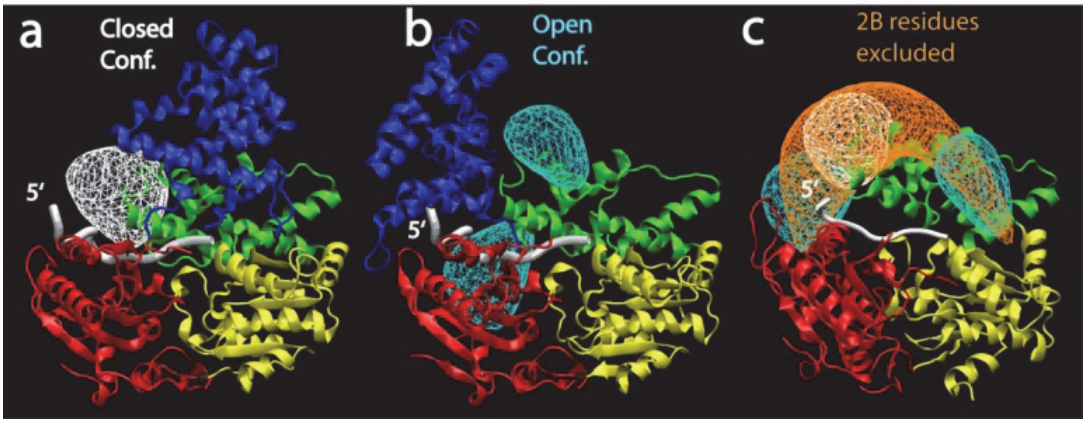


Figure 2.20 Probability distributions of optimal positions of the Cy5-labeled junction predicted by triangulation analysis, using FRET NPS, for the ADP state (Set 1). Color coding of Rep subdomains: yellow 1A; red 2A; green 1B; blue 2B. The predictions for (a) closed conformation (white mesh) and (b) open conformation (cyan mesh) are shown. In (c) 2B residues are excluded from the analysis (orange mesh) and the results of open (cyan) and closed (white) conformation results are also shown for comparison purposes.

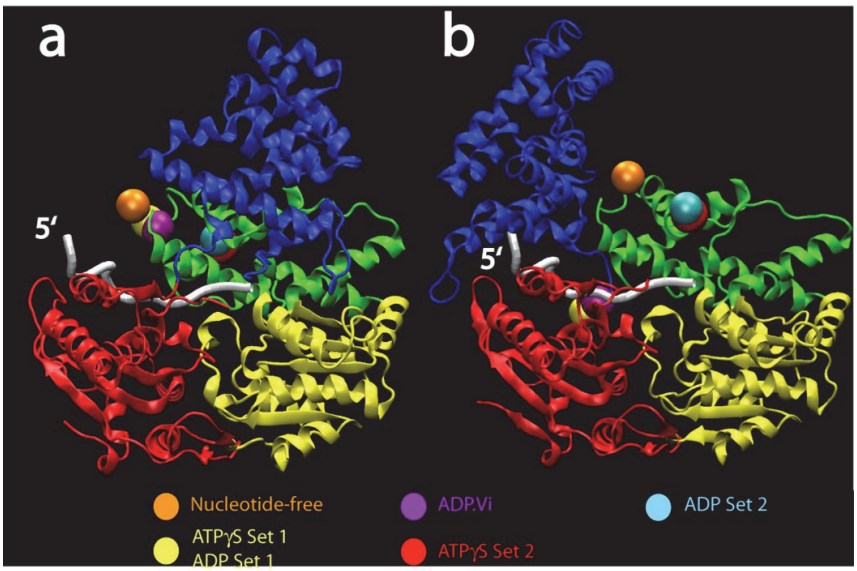


Figure 2.21 The most likely positions of the Cy5-labeled junction for (a) closed (b) open conformations for all nucleotide states are shown as colored spheres. In both figures most likely positions for Set1 of ATPγS and Set 1 of ADP are shown with a single yellow sphere as they are exactly the same. Color coding of spheres for both figures: nucleotide-free shown in orange, ATPgS Set 1 and ADP Set 1 shown in yellow, ADP.Vi shown in purple, ATPγS Set 2 shown in red, and ADP Set 2 shown in cyan. Color coding of Rep subdomains: yellow 1A; red 2A; green 1B; blue 2B.

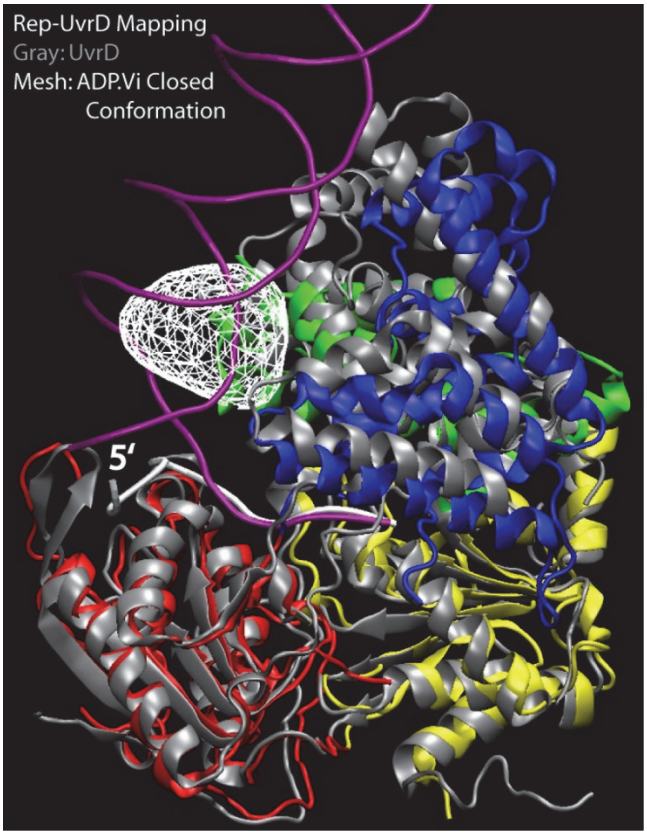


Figure 2.22 UvrD-pdDNA complex mapped onto Rep. The prediction of the junction position for the closed conformation is shown in white mesh for the ADP.Vi state. Same color coding is used for Rep subdomains.

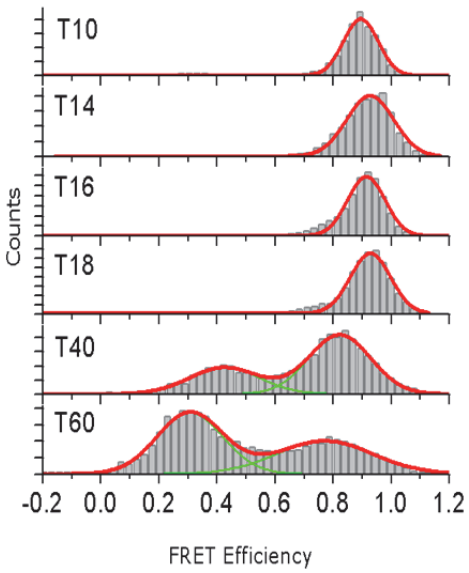


Figure 2.23 FRET efficiency between Cy3 on Rep, at Residue 333, and Cy5 at 5' junction of pdDNA that has varying tail lengths of 10T, 14T, 16T, 18T, 40T, and 60T. As the tail length gets

longer, Rep binds at locations away from the junction. However, even at tail lengths of 40T and 60T a significant population binds close to the junction forming a second peak in the distribution. We interpret this as a preference for Rep of binding to the vicinity of junction.

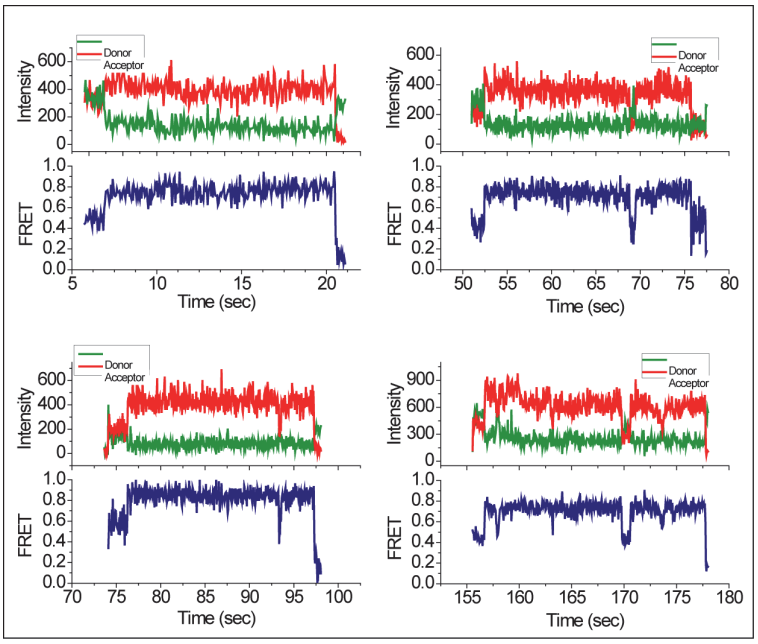


Figure 2.24 Sample single molecule traces demonstrating transitions between different FRET levels for ATPγS state.

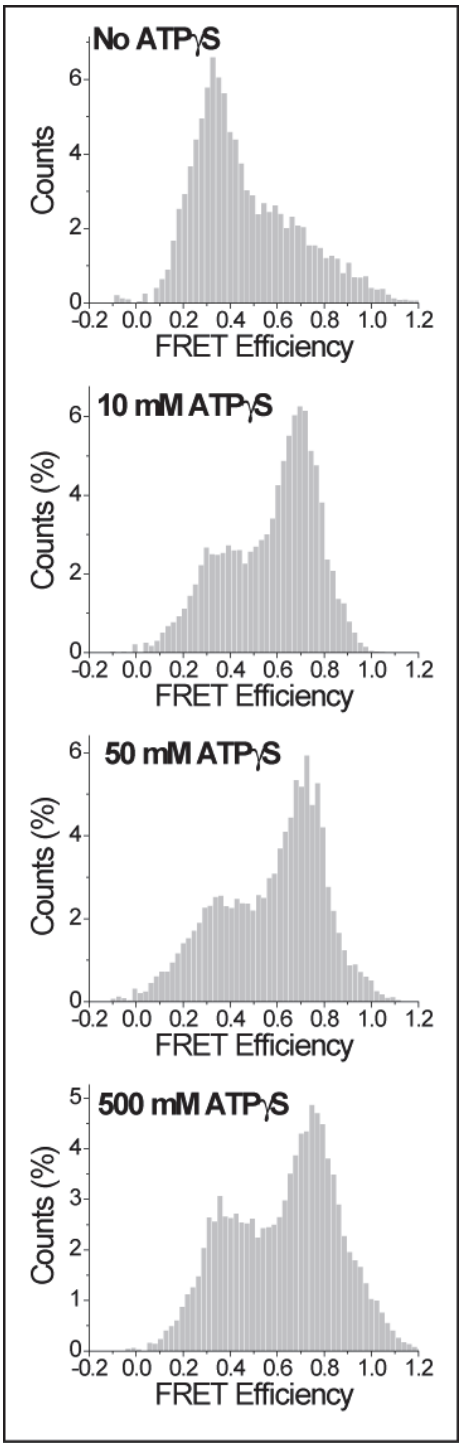


Figure 2.25 Rep-ATPγS binding saturates at ATPγS concentrations as low as 10 μM. Hence, at 500 μM ATPγS concentration, we expect all Rep molecules to be bound by ATPγS.

2.8 Tables

Summary of FRET Peaks

Residue Domain	Res 43 (1A)	Res 97 (1B)	Res 233 (1A)	Res 310 (2A)	Res 316 (2A)	Res 333 (2A)	Res 473 (2B)	Res 486 (2B)
Nt-free	0.38/0.67	0.59	0.48	0.66	0.69	0.82	0.68	0.87
ATPgS	0.48/0.78	0.65/0.89	0.53/0.77	0.66	0.55/0.81	0.70/0.85	0.80	0.79
ADPVi	0.52	0.68	0.57	0.76	0.76	0.73	0.71	0.84
ADP	0.50/0.72	0.69/0.94	0.51/0.71	0.59	0.57/0.79	0.71/0.89	0.74	0.81

Table 2.1 Summary of FRET Peaks. FRET efficiency peaks for all 32 measurements are shown. In the cases where two Gaussians were required to fit the data, the corresponding peaks are separated by "/", e.g. 0.38/0.67 means the data were fit by two Gaussians which were centered at FRET efficiency values of 0.38 and 0.67.

CLOSED CONFORMATION			
State	Most Likely Pos.	Dist. to 5'-end (Å)	Dist. to Junc. (Å)
Nt-free	(24,10,90)	17	14
ATPgS-1	(20,12,90)	20	16
ATPgS-2	(16,24,94)	27	26
ADPVi	(20,14,90)	19	16
ADP-1	(20,12,90)	20	16
ADP-2	(18,24,94)	25	25
OPEN CONFORMATION			
State	Most Likely Pos.	Dist. to 5'-end (Å)	Dist. to Junc. (Å)
Nt-free	(58,36,88)	32	44
ATPgS-1	(2,12,80)	34	25
ATPgS-2	(44,42,92)	31	42
ADPVi	(4,14,82)	32	24
ADP-1	(2,12,80)	34	25
ADP-2	(46,42,92)	32	43

Table 2.2 Comparison of Open and Closed Conformation Results. The most likely positions and the distances of these positions to the 5'-end of ssDNA for open and closed conformations. In the last column of the table, the distance of the most likely positions to the junction position estimated from UvrD-Rep mapping (**Figure 2.22**) is given

Chapter 3. Unwinding and regulation mechanism of Rep-like helicases

Rep, PcrA and UvrD are structurally homologous 3' to 5' single-stranded DNA (ssDNA) translocases and helicases that can unwind double-stranded DNA (dsDNA) utilizing energy from ATP binding and hydrolysis. However, *in vitro* studies show that monomers of these helicases have a very poor activity and cannot processively unwind DNA (*40, 59, 67, 78, 135*). They require oligomerization or association with cellular partner proteins to become unwinding-capable (*17, 136-143*). Although structural and biochemical studies have shown how these helicases can translocate on single-stranded nucleic acids directionally, consuming one ATP per nucleotide, the mechanism of nucleic acids unwinding and how the unwinding activity is regulated remain unclear and controversial (*17*). Since helicases that are free to unwind all nucleic acids encountered can be detrimental to genome integrity, it is critical to understand how their unwinding activities are regulated.

3.1 Regulation of helicase activity by protein conformation

Crystal structures of these helicases revealed a flexible domain (2B) that can rotate 130-160° in a swiveling motion between two conformations, referred to as *open* and *closed* orientations (*61, 62, 68, 144*) (Figure 3.1a). Two conflicting views posit that the 2B subdomain plays an inhibitory role or an active role in unwinding but no direct evidence has been found linking these conformations to the unwinding function (*55, 68, 70*).

Prior studies showed that when a Rep monomer translocates on the 3'-tail of a partial duplex DNA and approaches the duplex junction, the 2B subdomain gradually closes before the enzyme snaps back to the 3' end instead of unwinding the duplex ahead (*58*). This result led to our hypothesis that the *closed* conformation corresponds to the inactive form of Rep helicase. In this model, specific interactions between the 2B subdomain and the duplex DNA would instruct the protein to stop translocating as a way of maintaining the unwinding activity off until other molecular cues arrive. We aimed to test this model by engineering *E. coli* Rep mutants that are

intramolecularly crosslinked to constrain the 2B subdomain in *open* or *closed* conformations. Residues for the cysteine substitution mutagenesis and the length of the bis-maleimide crosslinkers were selected such that when crosslinked, the 2B subdomain cannot rotate appreciably, effectively locking the protein in one conformation (3.5.1, 3.5.2 and Figure 3.1a). We denote the *closed* form crosslinked with "-X", and the *open* form crosslinked with "-Y", Rep-X and Rep-Y. Contrary to our hypothesis Rep-X turned out to be a super active helicase, whereas Rep-Y was not distinguishable from wild type. Furthermore, a natural partner that enhances the helicase activity was shown to achieve its stimulating role also by selectively stabilizing the active conformation, i.e. Rep-X like closed form. Hence our work also provides insight into how the regulation of nucleic acids unwinding activity is achieved in nature.

Enzymatic activities of Rep-X and Rep-Y monomers were studied in single molecule and ensemble assays employing fluorescence resonance energy transfer (FRET), total internal reflection fluorescence (TIRF) microscopy, and optical tweezers force spectroscopy. Mutagenesis, purification, crosslinking procedures, and validation that crosslinking was intramolecular rather than intermolecular are described in 3.5.2: Intra-crosslinking, size exclusion chromatography and gel electrophoresis analysis.

3.2 Engineering of a super-helicase through conformational control

Regulation of protein function is an essential biological process; a major one among the myriad of ways it is achieved in nature, is by conformational control of regulatory and functional domains and residues. Artificial conformational control of biomolecular activities can reveal functional insights and enable the engineering of novel activities. For that purpose, many novel techniques have been developed for spatial and temporal control of protein function. They involve crosslinking between domains, cellular partner proteins, controlled effector molecule binding, caging moieties, affinity labels, fusion of voltage sensitive proteins that can be activated, cleaved or isomerized upon external stimuli such as light and pH (*123-126, 145-148*). However these techniques have been mainly based on inhibition of activity and removal of that inhibition in a controlled and sometimes reversible. In that regard, this works presents a unique case of a conformational control in which intramolecular crosslinking of a helicase monomer with undetectable unwinding activity was not inhibited but converted into a super-helicase that

can unwind thousands of base pairs processively even against a large opposing force. Control of protein function and tunable enzymatic activity has great potential in medicine and biotechnology (*146, 149-153*).We believe our bio-engineered monomeric super helicase without any nuclease activity can be useful in these areas, such as next generation DNA sequencing and isothermal DNA amplifications.

3.3 Results

If the closed form is inactive for unwinding as our initial hypothesis predicted, Rep-X would be inefficient in DNA unwinding even at high concentrations that make the wild type Rep active. Instead, in multiple turnover ensemble unwinding reactions using FRET-labeled DNA (see 3.5.3: Ensemble FRET unwinding assay), Rep-X unwound an 18-bp substrate with a 3'-$(dT)_{10}$ overhang at a much faster rate and higher reaction amplitude than the wild type Rep (Figure 3.1b-c). In contrast, Rep-Y unwinding rates were similar to that of Rep (Figure 3.2c), indicating that the dramatic unwinding enhancement is specifically achieved in the *closed* conformation. Because the large enhancement in unwinding activity observed in bulk solution can result from the activation of a monomer or from enhanced oligomerization, we performed single molecule FRET (smFRET) experiments to test if a single Rep-X can unwind DNA.

We immobilized Rep and Rep-X monomers to a polymer-passivated quartz surface using antibodies against the N-terminal hexa-histidine-tag on the protein (Figure 3.3a, 3.5.4: smFRET unwinding and RepD-PcrA interaction assays) to ensure that the observed activity belonged to monomers (*78*). For the unwinding substrate, we used a 18-bp duplex DNA with a 3'-$(dT)_{20}$ overhang labeled with a donor (Cy3) and an acceptor (Cy5) at two opposite ends of the DNA duplex, allowing us identify unwinding reactions as increases in FRET efficiency (E_{FRET})(Figure 3.3a) (*154*). When the DNA and ATP were added to the reaction chamber, we could observe the capture of a single DNA molecule by a single protein as the sudden appearance of fluorescence signal. Subsequent DNA unwinding generated ssDNA strands that coil up due to high flexibility and the E_{FRET} increased (*155*). Once the duplex was completely unwound, the acceptor–labeled strand was released, which was marked by sudden disappearance of the acceptor signal and recovery of the donor signal. The donor-labeled strand then dissociated, resulting in complete loss of fluorescence (Figure 3.3b). The mean duration of unwinding measured from the E_{FRET} increase to acceptor strand release was around 0.6 s, giving a lower limit on unwinding speed of 30 bp/s for the 18-bp substrate (Figure 3.3c). 82 % of the DNA molecules (661 out of 809) that

initially bound to Rep-X monomers were unwound (Figure 3.3d). In contrast, only 2 % of the DNA molecules (13 out of 847) that bound to Rep (i.e. without crosslinking) showed unwinding, and the unwinding yield for Rep-Y was 16 % (357 out of 2212), showing that constraining Rep into the *closed* form selectively activates the unwinding activity of a monomer. The nonzero amplitude of unwinding for Rep and Rep-Y may be due to conformational constraints caused by surface tethering in a small fraction of molecules.

In vitro studies have shown that the unwinding processivity of Rep and related helicases is limited even in their oligomeric forms, ranging from 30-50 bp (*59, 78, 156*). In order to investigate the processivity of Rep-X, we employed a dual optical tweezers assay (Figure 3.4, 3.5.5: Optical tweezers assay, (*5*)) that can monitor unwinding amplitudes and speeds over thousands of base pairs of DNA. The two traps held two streptavidin functionalized sub-micron sized polystyrene beads. The first was coated with 6-kbp dsDNA attached via a biotin on the blunt end and containing a 3' poly-dT ssDNA overhang on the other end ($(dT)_{10,15,75}$, as specified in figures, see 3.5.5 Optical tweezers assay). The other bead was coated with Rep-X molecules via biotinylated antibody against the hexa-histidine-tag. A laminar flow cell with two parallel streams of buffer was created for controlling the initiation of the unwinding reaction (Figure 3.4b inset, (*157*)). When we brought the two beads in close proximity in the first laminar stream (Buffer C with 100 μM ATP and 100 μM ATP-γS), a single Rep-X binding to the 3' overhang of the DNA formed a tether between the two beads without initiating unwinding. When we moved the tethered beads to the second laminar stream (Buffer C and 1mM ATP), the DNA tether between the beads progressively shortened as the Rep-X monomer unwound and pulled the DNA. Unless otherwise stated, as a precaution, SSB was added to second laminar stream, in order to prevent any subsequent interaction of unwound ssDNA with other Rep-X on the bead surface. The optical tweezers experiments that were performed without SSB yielded the same Rep-X behavior (see 3.5.5 Optical tweezers assay). By operating the trap under force feedback control, we kept the tension on the DNA constant at 10 pN. Additional controls and considerations ascertained that the observed activity stemmed from a single Rep-X regardless of the 3'-tail length and inclusion/omission of SSB (see 3.5.5 Optical tweezers assay). Remarkably, 97% (29 out of 30) of the Rep-X-DNA complexes tethered through a 3'-tail unwound the entire 6-kbp DNA in a processive manner (Figure 3.4b,d) and the average pause-free speed was 136 bp/s (Figure 3.4c). In comparison, only 3% (2 out of 61) of wild type Rep and 7% (5 of 70) of Rep-Y complexes displayed such processive unwinding events (Figure 3.4d). It is highly probable that

Rep-X has even greater processivity than 6-kbp, currently only limited by the length of the DNA used. The processive activity of a crosslinked Rep-X monomer shows the innate potential of these helicases that can be harnessed via conformational control.

We have also determined how much force Rep-X can generate during unwinding by performing measurements without the force feedback. Fixing trap positions (i.e. no force feedback) led to a rapid build-up of force on the Rep-X in the opposite direction of unwinding until the measurement was terminated due to the breakage of connection between the two beads (Figure 3.4e). The highest loads achieved in this experiment were not enough to stall the helicase permanently. More detailed analysis showed that the pause free unwinding rate of Rep-X was not impeded by increasing loads up to the limits of the linear regime of our trap (Figure 3.4f), approximately 60 pN (3.5.5 Optical tweezers assay). These results suggest that the engineered Rep-X is the strongest helicase known to date (*127, 158*).

In order to test if generation of a super active helicase can be reproduced for other helicases, thereby providing additional evidence of the conformational control mechanism, we engineered a PcrA-X helicase from *Bacillus stearothermophilus* PcrA. Mutations involved replacing two highly conserved Cys residues in this helicase (Table 3.1, Table 3.2 and Table 3.3) which reduced the apparent ssDNA-dependent ATPase activity from ~40 ATP/s (wild type) to 5 ATP/s. Nevertheless, upon crosslinking in the *closed* form, PcrA-X retained the low ATPase activity (4.3 ATP/s), but exhibited an enhanced helicase activity in comparison to both the wild type and non-crosslinked mutant in ensemble reactions (Figure 3.6a-b). smFRET experiments showed that PcrA-X monomers can unwind 39% (228 out of 578) of the 18-bp dsDNA they bind compared to only 4% (26 out of 617) for wild type PcrA (Figure 3.5a-b). In the optical tweezers assay, PcrA-X monomers, like Rep-X, were capable of processively unwinding of 1-6 kbp long DNA albeit at a much lower rate (2-15 bp/s, Figure 3.5c) whereas no PcrA molecule (0 out of 51) was able to do the same (Figure 3.5d). Despite the impaired activity levels of the PcrA mutant, conversion to PcrA-X made its monomers into highly processive helicase units, thus indicating a general mechanism of conformational control for this class of helicases.

Strong helicase activity of Rep-X and PcrA-X raises the possibility that the cellular partners of Rep or PcrA may switch on the powerful unwinding activity intrinsic to these enzymes by constraining them in the *closed* conformation. One such partner of PcrA is RepD, a plasmid replication initiator protein from *Staphylococcus aureus* that recognizes and forms an

covalent adduct with the *oriD* sequence of the plasmid, and then recruits PcrA for highly processive unwinding (*139, 140, 159, 160*). Based on our similar results with PcrA-X and the homologous *E. coli* counterpart Rep-X, but not the Rep-Y, we hypothesized that the RepD-induced PcrA activity enhancement is in fact the result of the conformational constraint of the helicase in the PcrA-X-like *closed* form. To test this prediction, we prepared an *oriD* DNA-RepD adduct, and measured the intramolecular conformation of PcrA bound to this adduct (3.5.4 smFRET unwinding and RepD-PcrA interaction assays). We used a double cysteine mutant of PcrA, PcrA-DM1, stochastically labeled with a mixture of donor and acceptor fluorophores that would be expected to generate high E_{FRET} in the *closed* form and low E_{FRET} in the *open* form (*1*) (Figure 3.5e, cartoons). The E_{FRET} distributions of PcrA-DM1 bound to the *oriD* DNA-RepD adduct and the *oriD* DNA alone are shown in Figure 3.5e. Only the PcrA-DM1 molecules with a fluorescence active Cy3-Cy5 pair were included in the analysis. The results revealed that the presence of RepD indeed biases PcrA toward the *closed* high E_{FRET} conformation. We therefore propose that the regulation mechanism of this class of helicases involves *in vivo* partner proteins that constrain the conformation of 2B subdomain to the *closed* form to activate its function.

3.4 Discussion

Why does constraining Rep and PcrA into the *closed* form convert an enzyme with undetectable unwinding activity to a super helicase? One possibility is that the intrinsic unwinding activity itself requires the *closed* form, for example via the torque-wrench mechanism proposed for UvrD (*68*). Another possibility is that the *open* form inhibits helicase function and crosslinking to the *closed* form prevents this inhibitory mechanism. We prefer the latter for the following reasons. First, Rep-Y crosslinked in the *open* form does unwind DNA as well as the wild type when the protein is at high concentrations in excess of DNA (Figure 3.2c). Therefore, the *closed* form *per se* is not absolutely required for unwinding activity. Second, using ultra-high resolution optical tweezers combined with smFRET capability, we found that UvrD assumes the *closed* conformation when it unwinds DNA but after it unwinds about 10 bp it switches to the *open* conformation and rewinds the DNA likely after strand switching (*161*). Therefore, we suggest that Rep-X becomes a highly processive super-helicase because the crosslinking prevents the *open* conformation required for strand-switching and rewinding that have been observed for UvrD (*161, 162*) and BLM (*163*).

Most conformational control of protein functions demonstrated so far first locks the naturally active protein to an artificially inhibited conformation so that additional controls imposed by researchers can be used to recover the original activity (*123, 125, 145, 148*). Our work is distinct in that we found a conformational control that activates a naturally inhibited unwinding function, and the resulting enzyme is a super-helicase that has unprecedentedly high processivity for a single motor helicase. RecBCD, another SF-1 helicase, has similarly high processivity but contains two motors and associated nucleases. Moreover it is known to backslide at opposing forces below 10 pN whereas Rep-X can be active against forces as high as 60 pN (*158*). This super helicase with high processivity and high tolerance against load without nuclease activities may also be useful for biotechnological applications such as single molecule nanopore sequencing (*164, 165*) and isothermal DNA amplification (*166*).

3.5 Materials and Methods

3.5.1 Mutagenesis and purification of proteins

Preparation of pET expression plasmids containing cysteine-less *rep* (C18L, C43S, C167V, C178A, C612A) and *pcrA* (C96A/C247A) with N-terminal hexa-histidine-tags were performed as described previously (*1, 72*). Site directed mutations to introduce two Cys residues for crosslinking (Rep-X: A178C/S400C, Cys178 is a native cysteine in the wild type, Rep-Y: D127C/S494C, PcrA-X: N187C/L409C) were done using QuikChange Lightning kit (Life Technologies, Inc.) and mutagenic primer oligonucleotides (Integrated DNA Technologies Inc., Coralville, IA). Protein purifications were performed as described previously (*1, 72*). Catalytic activity levels of purified proteins as well as the crosslinked samples were determined in ssDNA dependent ATPase activity assay using Invitrogen EnzChek phosphate assay kit (Life Technologies Inc.), oligonucleotide $(dT)_{45}$ and 1 mM ATP in buffer D (see ensemble FRET unwinding assay).

Wild type RepD from *Staphylooccus aureus*was purified as described in (*140, 159*) with the following differences. A wt-RepD encoding pET11m-RepD plasmid was constructed for expression in B834 (pLysS). The gene sequence contained silent mutations to introduce restriction sites for AgeI, PstI, SacI, and to modify the nick site (TCT'AAT to TCGAAT) to prevent premature cleavage by RepD during expression. An ammonium sulfate precipitated pellet (from

0.5 L culture) was resuspended and run through serially connected 5 ml Q-Sepharose (removed once the sample was through) and 5 ml heparin-Sepharose cartridges connected in series (GE Healthcare), and eluted on an ÄKTApurifier 10 FPLC system.

3.5.2 Intra-crosslinking, size exclusion chromatography and gel electrophoresis analysis

Dual-cysteine Rep mutants were incubated overnight at 4 °C with 2- to 100-fold excess of bis-maleimide crosslinkers DTME (13 Å) and BMOE (8 Å) purchased from Thermo Fisher Scientific, Rockford, IL. PcrA-X was crosslinked with DTME and $BM(PEG)_2$ (14.7 Å) from the same manufacturer. Excess crosslinkers were removed by Bio-Rad P-30 desalting column. Crosslinked Rep-X, Rep-Y and PcrA-X samples were stored at -20°C or -80°C as described (*1, 72*). Data presented in this manuscript used BMOE (8 Å), but other crosslinkers of various lengths gave similar results (data not shown). DTME is a di-sulfide containing crosslinker that we reduced with β-mercaptoethanol (β-ME) or tris(2-carboxyethyl)phosphine (TCEP) to revert the crosslinked helicase to the wild type form for control purposes.

Crosslinked Rep and PcrA samples were separated from multimeric byproducts using Superdex 200 grade 10/300GL or HiLoad 16/600 gel filtration columns on an ÄKTApurifier 10 FPLC system. The crosslinking efficiency was monitored by SDS-PAGE analysis on 7.5-10% Tris-glycine gels (Bio-Rad). As needed for gel analysis, reduction of samples crosslinked with DTME was achieved by 5% (v/v) β-ME during SDS denaturation step.

Crosslinking of the double Cys mutants with the bis-maleimide linkers has the potential of producing covalently attached multimeric species, in addition to the intended internally crosslinked monomeric species. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) can distinguish these species from the non-crosslinked monomers. Here we show a representative analysis of a crosslinked Rep-Y sample. On an SDS polyacrylamide gel, internally crosslinked monomers were visualized as a slightly higher molecular weight (MW) band (Figure 3.2a; middle band) with respect to the non-crosslinked monomer (bottom band) while the multimeric species showed up at much higher MW (top band) (*167*). Figure 3.2a Rep-Y lane shows three such bands of a Rep-Y sample crosslinked with a cleavable di-sulfide containing crosslinker (DTME). First we established that the dominant middle band and the fainter top band were the crosslinked species, since they disappeared upon cleavage of the crosslinker

(Rep-Y+β-ME lane). In order to ensure that the dominant middle band was not multimeric but the intramolecularly crosslinked monomeric species we were aiming for, the Rep-Y sample was fractionated according to the molecular size on a Superdex 200 size exclusion chromatography (SEC) column controlled by an FPLC apparatus. Elution profiles of Rep-Y and non-crosslinked Rep are shown in the Figure 3.2a, top panel. Eluted fractions were analyzed on the SDS polyacrylamide gel (lanes F1-F7). As expected, the only multimeric species were in the fainter top band that was eluted in the early SEC fractions (11-13 ml), and the dominant middle band was the intramolecularly crosslinked monomeric species that was eluted together with the non-crosslinked Rep monomer. Having established that the intra-crosslinked protein show up as a retarded band compared to the non-crosslinked form on SDS polyacrylamide gels (such as the Rep-Y data presented here), we routinely used it to check the efficiency of crosslinking reactions for Rep-X, Rep-Y and PcrA-X.

3.5.3 Ensemble FRET unwinding assay

Multiple turnover ensemble unwinding kinetics was used as a quick method to gauge the effect of the mutations and conformational modifications to the helicase activity. We used a 18-bp FRET labeled DNA substrate with 3'-$(dT)_{10}$ overhang (Figure 3.1b), constructed by annealing complementary oligonucleotides DNA7 (Cy5-GCC TCG CTG CCG TCG CCA) and amino-dT labeled DNA8 (TGG CGA CGG CAG CGA GGC-(T-Cy3)-T_{10}). Alternatively another similarly labeled 50-bp DNA with a 3'-$(dT)_{30}$ overhang was also used. It was made by annealing oligonucleotides DNA9 (Cy5-TCA ACT AGC AGT CAT AGG AGA AGT ATT AAC ATG CCT CGC TGC CGT CGC CA) and amino-dT labeled DNA10 (TG GCG ACG GCA GCG AGG CAT GTT AAT ACT TCT CCT ATG ACT GCT AGT TGA (T-Cy3) T_{29}). Unless otherwise stated, a 5 nM ensemble FRET DNA was mixed with 50 nM helicase in buffer D (10 mM Tris-HCl [pH 8.0], 15 mM NaCl, 10 mM $MgCl_2$, 10% (v/v) glycerol, 0.1 mg/ml BSA) and 1 mM ATP was added to start the unwinding reaction in a quartz cuvette. A Cary Eclipse fluorescence spectrophotometer was used to measure the donor (I_{555nm}) and the acceptor signal (I_{667nm}) under the 545 nm excitation (5 nm slit, 2-10 Hz acquisition rate and 600-900V photomultiplier voltage). Unwinding of the substrate was monitored by the decrease in ensemble E_{FRET} value, defined as $E_{FRET\text{-}ensemble} = I_{667nm}/(I_{555nm} - I_o + I_{667nm})$ where I_o was the baseline donor signal of unpaired Cy3 prior to addition of ATP.

3.5.4 smFRET unwinding and RepD-PcrA interaction assays

All smFRET experiments were conducted on a custom-built prism type TIRF microscopy stage with Andor EMCCD camera as described in (*83, 84*). Reaction chambers were formed by quartz slides and glass coverslips passivated with polyethyleneglycol (PEG) and 1% biotinylated PEG (mPEG-SC and bio-PEG-SC, Laysan Bio, Arab, AL), followed by 5 min incubation with Neutravidin (Thermo Scientific, Newington, NH) for immobilization of biotinylated molecules on the chamber surface as described below.

For the smFRET unwinding experiments, the reaction chamber was first incubated with biotinylated anti penta-histidine tag antibody (Qiagen, Valencia, CA), followed by 10-30 min incubation of His_6-tagged helicase sample (0.5-1 nM). The unwinding of the DNA was initiated by flowing 1 nM smFRET DNA and 1 mM ATP in the reaction buffer A (10 mM Tris-HCl [pH 8.0], 10 mM $MgCl_2$, 15 mM NaCl, 10% (v/v) glycerol, , 1% (v/v) gloxy and 0.2% (w/v) glucose, an oxygen scavenging system (*80*) and 3-4 mM Trolox, a fluorophore triplet state quencher (*81, 82*)). smFRET DNA substrate was constructed by annealing the oligonucleotides DNA3 (Cy5-GCC TCG CTG CCG TCG CCA) and DNA4 (Cy3-TGG CGA CGG CAG CGA GGC-T_{20}). PcrA-RepD interaction assay involved preparation of the RepD-*oriD* DNA adduct as described in (*140*). Biotinylated *oriD* DNA substrate was constructed by annealing oligonucleotides DNA1 (CTA ATA GCC GGT TAA GTG GTA ATT TTT TTA CCA CCC AAA GCC TGA AGA GCT AAT CGT TCG G) and DNA2 (biotin-CCG AAC GAT TAG CTC TTC AGG CTT TGG GTG GTA AAA AAA TTA CCA CTT T_{15}) In one chamber only *oriD* DNA (50-100 pM) was immobilized on the surface. In a second chamber RepD-*oriD* DNA adduct was immobilized. 100-500 pM dual labeled PcrA-DM1 was injected into the chambers in buffer B (10mM Tris [pH7.5], 10% glycerol, 15mM NaCl, 50mM KCl, 5mM $MgCl_2$, 3.4 mM Trolox, 1% (v/v) gloxy, 0.2% (w/v) glucose). Short movies of many different areas were recorded. Since the two Cys residues of PcrA-DM1 were randomly labeled with Cy3-Cy5 mixture, each movie contained a brief initial 633-nm laser excitation period to determine the molecules with a fluorescent Cy5, followed by turning on the 532-nm laser for Cy3 excitation. Only the PcrA-DM1 molecules with a colocalized donor-acceptor pair factored in the E_{FRET} histograms.

smFRET signal was acquired by an Andor EMCCD camera operated with a custom software at 16-100-ms time resolution. E_{FRET} was calculated as described in (*83*).

Unwinding period was measured as described in the text. Fraction of unwinding events calculated as the proportion of the all DNA binding events that displayed a E_{FRET} increase phase.

Error bars were calculated according to Clopper-Pearson binomial proportion confidence interval method (*168*).

3.5.5 Optical tweezers assay

The optical trap handle was a 6098-bp long DNA, amplified from λ-phage DNA and flanked by a 5'-biotin and a 3'-$(dT)_{10,15,75}$ overhang on the other end. First a 5'-tailed 6083-bp fragment was amplified by the auto-sticky PCR reaction (*169*) using primers P1 (biotin-GGC AGG GAT ATT CTG GCA) and P2 (GAT CAG TGG ACA GA-abasic-A AGC CTG AAG AGC TAA TCG TTC GG). Subsequently the amplicon was annealed and ligated with oligonucleotide DNA5 (TTC TGT CCA CTG ATC-$(T)_{10,15,75}$) to create the 3'-overhang for the initial helicase binding (10, 15 or 75-nt, as specified in figures). DNA beads were prepared by adding biotinylated 6-kbp DNA to the streptavidin-coated polystyrene beads (0.79 μm in diameter, Spherotech, Lake Forest, IL), and incubated at 25°C for 30 min. Protein samples were pre-incubated with biotinylated anti penta-histag antibody (Qiagen, Valencia, CA) on ice for 1 hour. 1 μl of this mixture, 1 μl of streptavidin beads, and 8 μl buffer (100 mM Tris-HCl [pH 7.5], 100 mM NaCl, 10% glycerol (v/v)) were mixed and incubated for 30 min on ice to make the protein coated beads. Reactions were performed in laminar flow chambers that were designed and assembled as described in (*170*). Reaction buffer C, consisted of 100 mM Tris pH 8.0, 15 mM NaCl, 10% (v/v) glycerol, 10 mM $MgCl_2$, and an oxygen scavenging system (100 μg/ml glucose oxidase, 20 μg/ml catalase, and 4 mg/ml glucose) to reduce photodamage to the sample (*171*). The reaction chamber contained two laminar streams of buffer C with different ATP, ATP-γS and SSB content as described in the text. The dual-trap optical tweezers were set up and calibrated as described in (*172, 173*). All measurements were recorded at 100 Hz with a custom LabView software (8.2; National Instruments, Austin, TX) and smoothed with a 100 Hz boxcar filter. In the "force–feedback" mode, unwinding was allowed to occur against a constant force of 10-22 pN (as specified). The contour length of DNA was calculated from the measured force and end-to-end extension of the molecule and using the worm-like chain model (persistence length of 53 nm, stretch modulus of 1,200 pN and distance per base-pair of 0.34 nm). The velocity of DNA unwinding in the force feedback mode was determined from a linear fit of the contour length of DNA in a sliding window of 0.2 s (twenty one data points), pauses longer than 0.2 s were removed, then the velocity was averaged in 1 s binning intervals. Error for the fraction of unwinding events per

tether formation was calculated with the Clopper-Pearson binomial proportion confidence interval method (*168*).

The force dependence of Rep-X unwinding activity was measured in the "fixed-trap" mode, by stopping the force feedback. The force data (100 Hz) was smoothed with a gaussian filter (by applying a 33-Hz moving average filter 10 times). Paused regions (velocity < 10 bp/s) were removed. The pause-free unwinding velocities were calculated and normalized by the velocity at 20 pN for all molecules, and binned against the dynamic force values up to 60 pN to create V/V_{20pN} vs. F plot (Figure 3.2f). We previously found that the force response of our trap is linear against the bead displacements up to 72 nm (determined in a separate experiment without force feedback and trap stiffness of 0.167 pN/nm, the force vs. extension curve of dsDNA started to deviate from the theoretical worm like chain model above 12 pN. Hence we calculated the maximum reliable force to be at least 59 pN (at a trap stiffness of 0.82 pN/nm).

We considered the possibility that the highly processive unwinding observed in our optical tweezers assay was caused by multiple Rep-X acting on the same DNA, and hence a single Rep-X may not be the highly processive helicase unit. If multimeric stoichiometry of Rep-X had been required for highly processive unwinding, then the majority of binding events (i.e. formation of a tether) would not have displayed unwinding activity, since single Rep-X binding is the statistically the most probable event. However, the majority of tethers formed displayed highly processive unwinding, indicative of single molecule activity.

To establish further that the unwinding of the 6-kbp DNA is achieved by single Rep-X molecule, we repeated the experiment using beads incubated in lower concentrations of Rep-X, thus decreasing the number of Rep-X molecules per bead. Consequently, Rep-X binding (tether formation) took longer and required more trials of bumping the two beads. However the subsequent unwinding was still the prevalent behavior.

A second test to ensure that the processive unwinding of 75-nt overhang DNA was due to single Rep-X molecule, but not multiple, we used shorter 3' overhangs (10 and 15-nt) that would allow initial binding of only one Rep-X molecule based on the crystal structure (*61*). Rep-X exhibited the same behavior on the short overhang DNA molecules, confirming that the high processivity of unwinding was indeed a monomer activity..

A third precautionary technique was to add 66 nM of *E. coli* ssDNA binding protein (SSB) in the unwinding reaction stream in order to render the unwound ssDNA inaccessible to other

Rep-X molecules. However, inclusion/omission of SSB did not change the Rep-X behavior indicating that SSB is not essential to ensure single molecule activity. This is probably due to the design of the dual optical tweezers assay in which the DNA is only under tension between the "front runner" Rep-X molecule and the streptavidin on the other bead. Binding of a secondary Rep-X to the already unwound ssDNA should not affect the measurements because the secondary Rep-X, which is also tethered to the bead, cannot interact with the front runner that is tethered elsewhere on the bead.

3.6 Enhanced unwinding activity by mutagenesis of 2B subdomain

Another SF1 helicase, UvrD shares close to 100 % structural and 40% sequence homology with Rep. A genetic screening study detected that a UvrD mutant *uvrD303* increased ATP hydrolysis and DNA unwinding activity (*174*). In that study, UvrD303 unwound DNA an order of magnitude more efficiently than the wild type UvrD in ensemble unwinding experiments. *uvrD303* is a double alanine replacement of Asp403 and Asp404 residues. These residues are located on the 2B subdomain surface that interacts with the 1B domain, in the closed form. Since we found that the closed form is the active form of Rep and PcrA, it is very likely that these mutations stabilize the closed form. This would further prevent dynamics of 2B subdomain and would increase the unwinding activity of UvrD. We also tested the unwinding activity of the same mutant variant of Rep helicase, Rep DD, has been purified and tested on optical trap for processivity and helicase unwinding activity.

We used same duplex DNA substrate as for Rep-X, Pcr-X and wt Rep. As soon as we formed tether we started a force feedback at 10pN. There was an apparent difference from wild type Rep and Rep-X measurements; this time we saw partial unwinding in 20% of formed tethers with subsequent dissociation. Worthwhile to mention that we still saw full unwinding in 3% of the traces consisted with wild type observation and the fact that Rep might be locked in "closed" form due to bead surface interaction. Interestingly we started to see slippages even at 10 pN forces.

Though whether latter mutations enhance interaction between two domains is not known, these results suggest that interaction surface between 1B and 2B subdomain is important and most likely drives Rep protein to "closed" conformation.

3.7 Figures

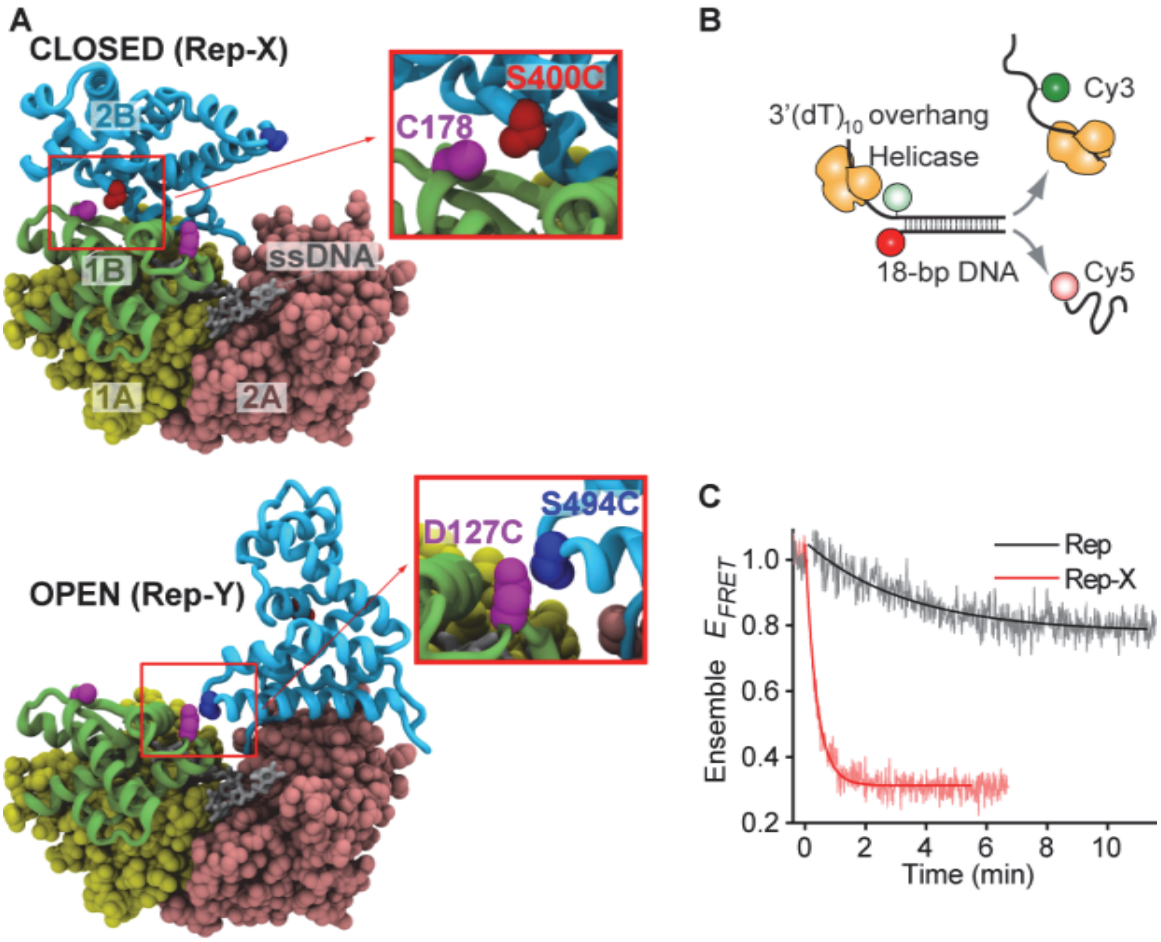


Figure 3.1 Crosslink mediated conformational control of the helicase activity of Rep. (A) Open and closed form Rep crystal structures (pdb entry 1UAA) are shown. Domains are colored and named accordingly. 3' end of the ssDNA is visible. Residues that were mutated to cysteine and crosslinked to lock the conformation are shown as pink, blue and red van der Waals spheres in both conformations as reference. Close-ups show the only two residues that were crosslinked for engineering Rep-X and Rep-Y. **(B)** Cartoon of FRET labeled ensemble unwinding DNA shows the conversion of high FRET efficiency (E_{FRET}) substrates to low E_{FRET} by the helicase. Lightness of the donor and acceptor color represents the relative intensity changes. . **(C)** Ensemble unwinding kinetics of DNA from **B** by Rep and Rep-X shows the enhanced helicase activity of Rep-X over Rep as measured via ensemble E_{FRET} (3.5.3 Ensemble FRET unwinding assay). Solid lines are fitted exponential decay curves for guiding the eye.

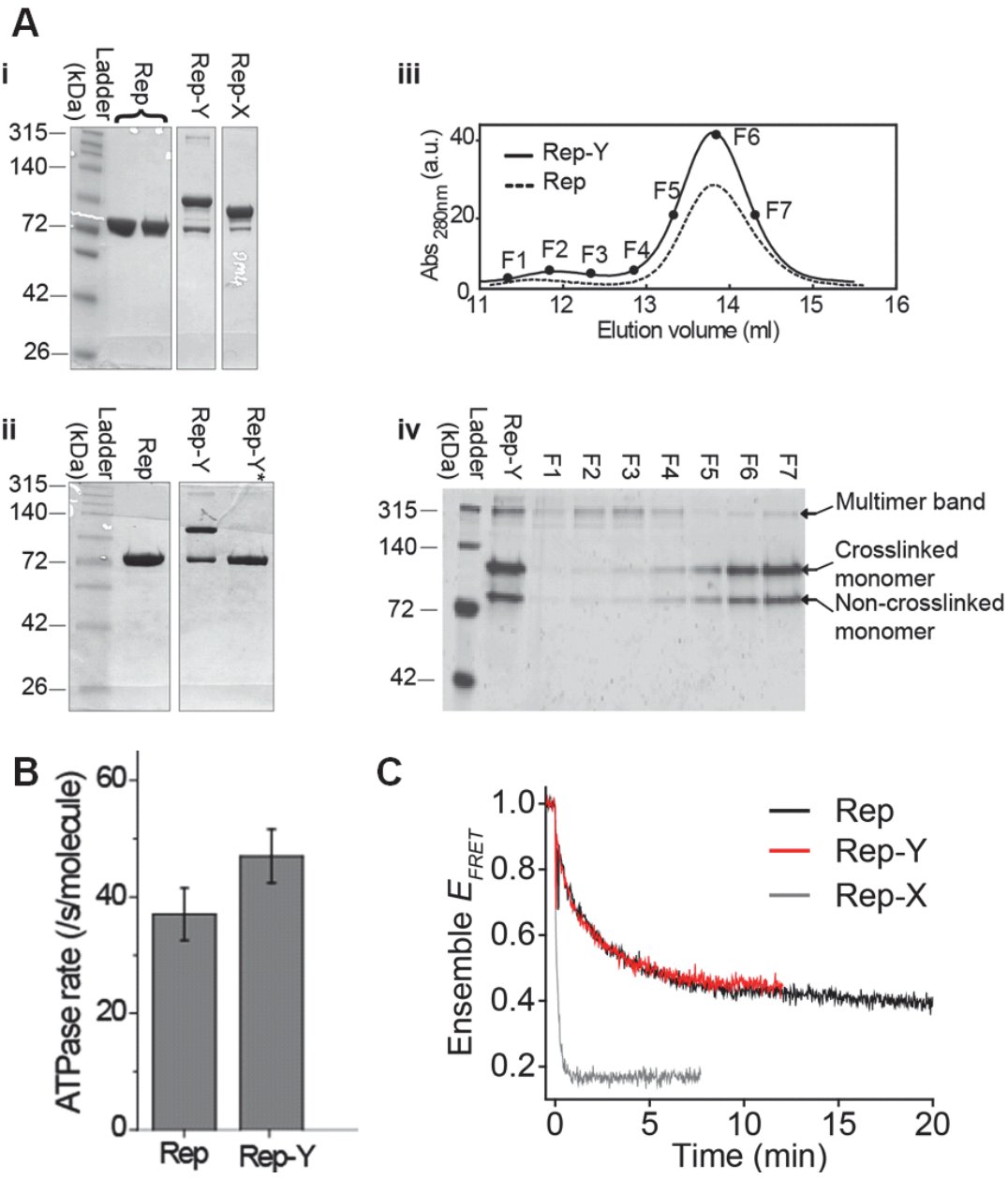


Figure 3.2 Analysis of intra-crosslinking reaction products and ensemble enzymatic activity tests for open form crosslinked Rep-Y. (A) Intra-crosslinking is validated by size exclusion chromatography (SEC) (iii, iv), and SDS-PAGE analysis (i, ii). Intra-crosslinked proteins always produced three bands on SDS polyacrylamide gels, a dominant middle band (location is slightly different for Rep-X and Rep-Y, see i), a faint bottom band and a much fainter top band as shown in the Rep-Y lanes (i-ii, iv). (iii) shows the SEC elution profile for Rep (dotted line) and the Rep-Y sample (solid line). Rep-Y fractions, F1-F7, collected from SEC were run on an SDS polyacrylamide gel along with Rep-Y (iv). β-ME reduced Rep-Y (crosslinked with a di-sulfide crosslinker DTME) showed a single band as Rep (ii). **(B)** ssDNA dependent ATPase level of Rep-Y is very similar to Rep. Error bars represent standard deviation over multiple preparations. **(C)** Ensemble unwinding kinetics of the Rep-Y does not show any enhancement over non-

crosslinked Rep (10 nM helicase, 5 nM 50-bp ensemble unwinding DNA with 3'-$(dT)_{30}$ overhang in buffer D and 1 mM ATP). In comparison, both helicases were much slower than Rep-X.

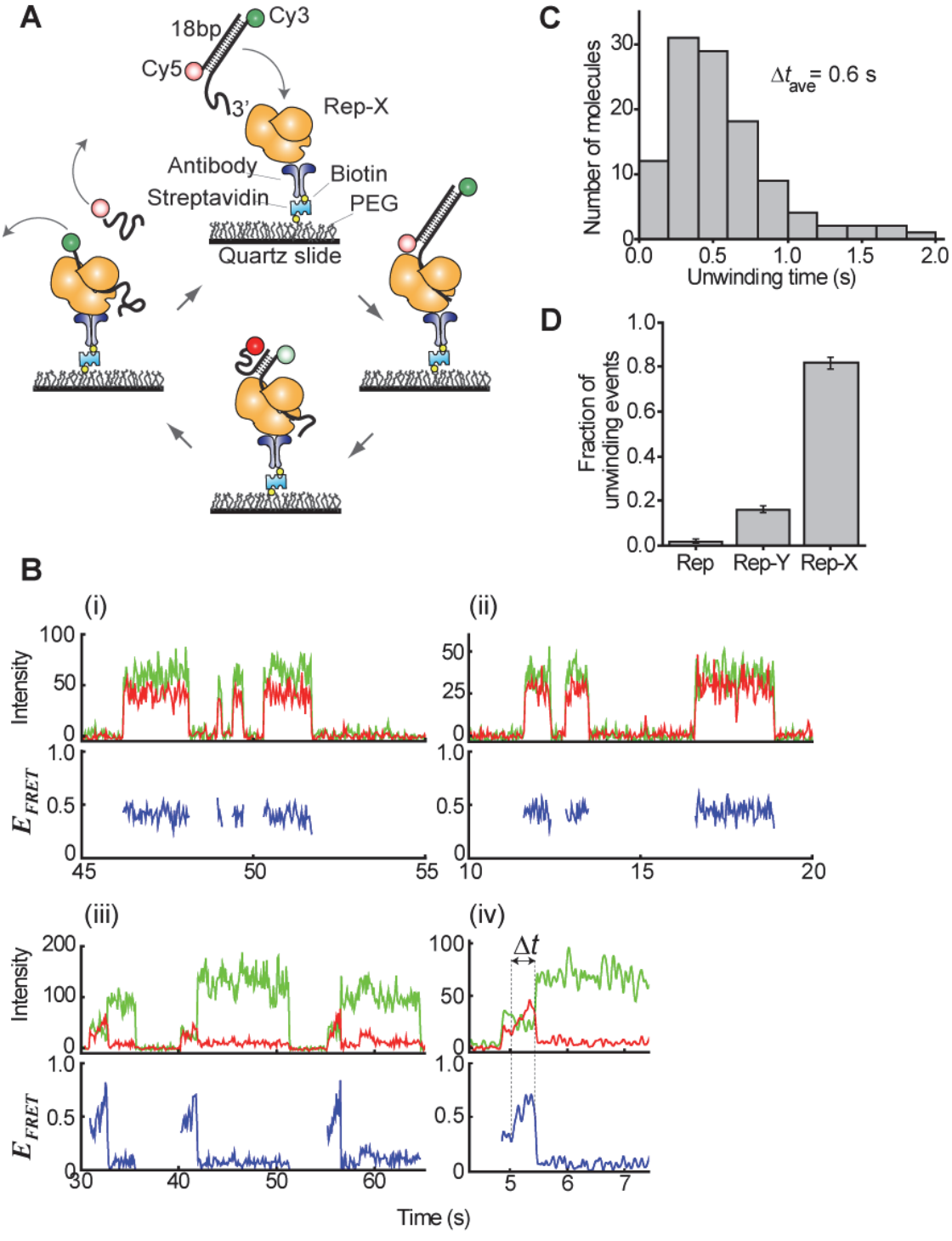


Figure 3.3 smFRET reveals DNA unwinding by Rep-X monomers. (A) A cartoon representation of unwinding stages of dual labeled DNA by a Rep-X monomer is shown. Color lightness of the donor (green) and acceptor (red) on the FRET labeled DNA represents the change in the

emission intensities from low to high E_{FRET} as the unwinding progresses. **(B)** Representative smFRET curves show the dominant behavior of DNA binding and dissociation for Rep (i) and Rep-Y (ii), and DNA binding, unwinding and dissociation (first the acceptor strand, then the donor strand) for Rep-X (iii, iv). The donor fluorescence signal is in green, acceptor in red and E_{FRET} in blue. Unwinding time is denoted by Δt. **(C)** Distribution of Rep-X unwinding periods measured from the smFRET assay. **(D)** Fractions of =DNA binding events that led to unwinding (i.e. exhibited an E_{FRET} increase phase) in smFRET experiments are shown for Rep, Rep-Y and Rep-X. Error bars represent 95% confidence bounds (3.5.4 smFRET unwinding and RepD-PcrA interaction assays).

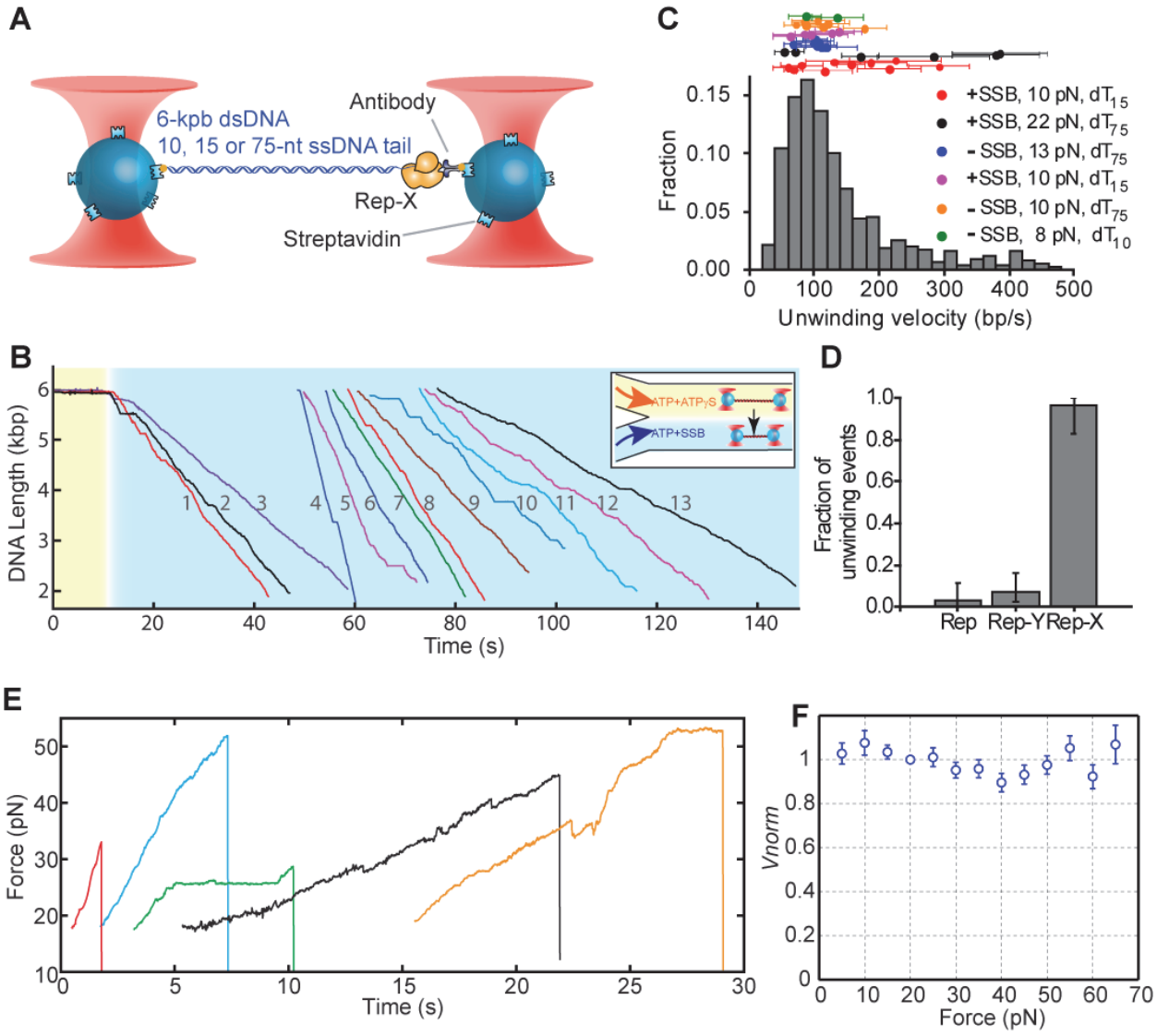


Figure 3.4 Optical tweezers reveal Rep-X is a highly processive and a very strong helicase. (A) Cartoon representation of the optical tweezers assay is shown. A single Rep-X molecule loads on the free ssDNA tail and unwinds the 6-kbp DNA, which shortens the tether and pulls the beads closer. The force feedback keeps the tension on the DNA and Rep-X constant at 10-22 pN. **(B)** Individual unwinding traces show the extent of processive unwinding by Rep-X on the 6-kbp

DNA offset for clarity. Background is color coordinated with the inset that shows the two laminar flows in the reaction chamber; the first stream (yellow) favors the stalled Rep-X-DNA complexes, and the second stream (blue) initiates the unwinding. Traces [1-3, 5, 7, 13] were obtained with a constant tension of F=10 pN, with SSB on 15-nt overhang DNA. Trace [4] was obtained with F=22 pN, with SSB on 75-nt overhang DNA. Traces [6, 8, 9] were obtained with F=10 pN, no SSB, on 75-nt tail DNA. Trace [10] was obtained with F=8 pN, no SSB on 10-nt overhang DNA. Traces [11, 12] were obtained with 13 pN, no SSB on 75-nt overhang DNA. **(C)** Distribution of Rep-X unwinding velocities measured from 38 Rep-X molecules is shown. Scatter plots represent the mean unwinding velocity and the standard deviation for each molecule from six different experiments (color coded). **(D)** Fraction of the enzyme-DNA binding events that led to unwinding of DNA for Rep, Rep-Y and Rep-X are compared (obtained with 75-nt overhang DNA and no SSB). Error bars represent the 95% confidence bounds (3.5.5 Optical tweezers assay). **(E)** Unwinding by five representative Rep-X molecules against pulling force in the fixed trap assay are plotted. Pulling force is linearly coupled to the extent of unwinding as the Rep-X pulls the beads closer, and the optical trap acts like a spring. Tether breaks appear as sudden drops. Fixed trap data was obtained with 15-nt overhang DNA and without SSB. **(F)** Average of normalized unwinding velocities of 58 Rep-X molecules is plotted against the opposing force from the fixed trap assay shown in E. Pause-free unwinding velocities of individual molecules were normalized by each molecule's velocity at 20 pN and averaged across all the molecules in 5 pN bins. Error bars represent standard error of the mean. This plot shows the high force tolerance of the engineered super-helicase Rep-X.

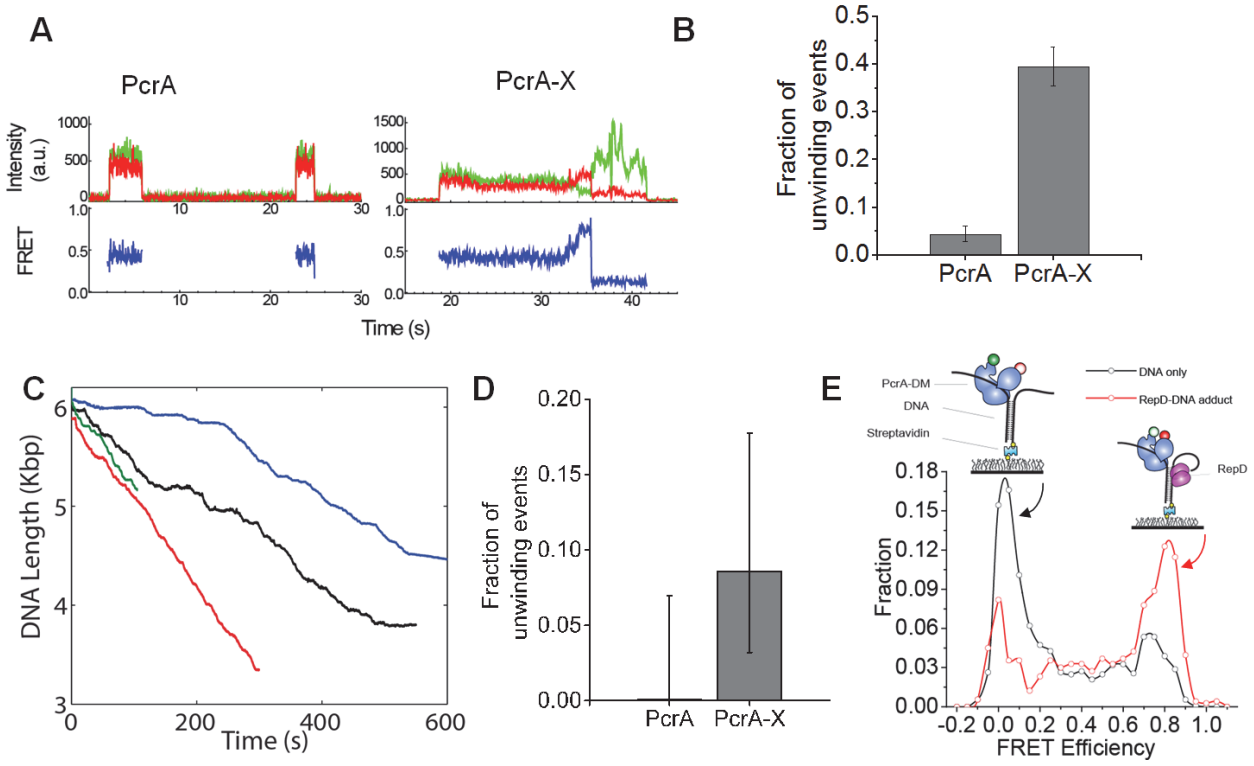


Figure 3.5 Demonstration of conformational control on PcrA helicase. Conformational control of protein function achieved on Rep-X was duplicated on PcrA helicase. **(A)** Sample smFRET DNA binding and unwinding traces for PcrA and PcrA-X shows that the PcrA-X monomers are too

active helicase units. **(B)** Fractions of enzyme-DNA binding events that led to an unwinding phase for PcrA and PcrA-X in the smFRET assay are shown. **(C)** Processive unwinding of 6-kbp DNA by four representative PcrA-X molecules on the optical tweezers assay. Blue and green traces are from 75-nt overhang DNA, with no SSB at 5 pN tension, red and black traces are from 15-nt overhang DNA, with SSB at 4 pN tension. **(D)** Fractions of enzyme-DNA binding that led to the unwinding of 6-kbp DNA in the optical tweezers assay (75-nt overhang DNA, no SSB, and a constant tension of 5 pN). **(E)** The conformational effect of RepD, a stimulatory partner of PcrA, on PcrA was measured in a smFRET assay (cartoons). E_{FRET} histograms show that the PcrA bound to RepD adduct is biased toward the closed form (high E_{FRET} population) compared to PcrA bound to the bare *ori-D* DNA. In B and D, error bars represent the 95% confidence bounds (3.5.4 smFRET unwinding and RepD-PcrA interaction assays, 3.5.5 Optical tweezers assay).

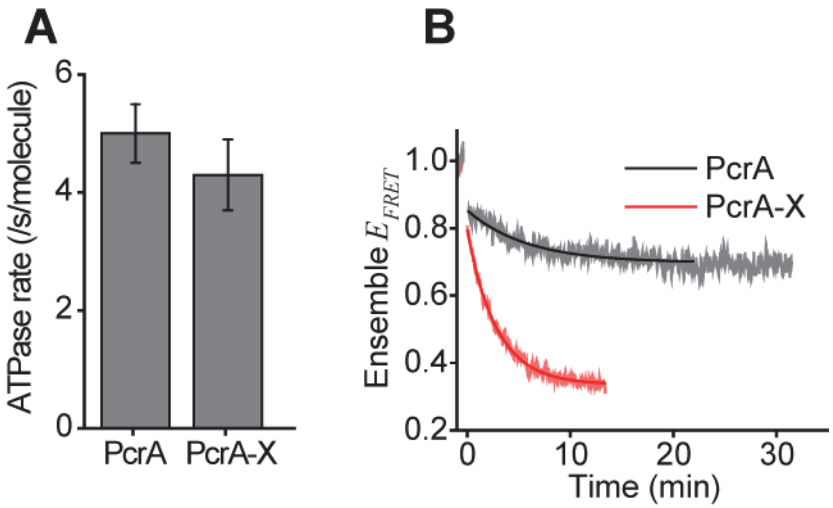


Figure 3.6 Ensemble enzymatic activity tests for PcrA and PcrA-X. (A) ATPase activity of our mutant PcrA sample before and after crosslinking into the PcrA-X form. Error bars represent standard deviation of multiple preparations. **(B)** Ensemble unwinding assay revealed an enhanced enzymatic activity of PcrA-X over wild type PcrA. Solid lines are fitted exponential decay curves for guiding the eye.

3.8 Tables

		Motif Ia	TxGx motif	
Rep	*Escherichia coli*	KIAHLIRGCGYQARHIAAVTFTNKAAREMKERVGQTLGRK-EARGLMISTFHTLGLDIIK		93
UvrD	*Escherichia coli*	RIAWLMSVENCSPYSIMAVTFTNKAAAEMRHRIGQLMGT--SQGGMWVGTFHGLAHRLLR		99
UvrD	*Mycobacterium tuberculosis*	RIAYLMAARGVGVGQILAITFTNKAAAEMRERVVGLVGE--KARYMWVSTFHSTCVRILR		112
UvrD	*Mycoplasma capricolum*	KIAYLIEKQNIDPSRILAVTFTNKAAKEMKERVLQITNN--SFKSPFISTFHSWCSKVLR		99
UvrD	*Deinococcus radiodurans*	RIAHLIGHYGVHPGEILAVTFTNKAAAEMRERAGHLVPG---AGDLWMSTFHSAGVRILR		101
PcrA	***Geobacillus stearothermophilus***	**RIAYLMAEKHVAPWNILAITFTNKAAREMRERVQSLLGG--AAEDVWISTFHSMCVRILR**		**101**
PcrA	*Bacillus subtilis*	RIAYLMAEKHVAPWNILAITFTNKAAREMKERVESILGP--GADDIWISTFHSMCVRILR		101
PcrA	*Staphylococcus aureus*	RIAYLLDEKDVSPYNVLAITFTNKAAREMKERVQKLVGD--QAEVIWMSTFHSMCVRILR		97
PcrA	*Leuconostoc citreum*	RIAHLVQDLNVFPWRILAITFTNKAAREMRERIAALLSED-VARDIWVSTFHALAVRILR		100
PcrA	*Fructobacillus fructosus*	RVAHLIEDLDVRPWRILAITFTNKAAREMRERIAKLVDPE-QAQAVWVSTFHALAVRILR		100
PcrA	*Staphylococcus epidermidis*	RIAYLLDEKDVSPYNILAITFTNKAAKEMKARVEHLVGE--EAQVIWMSTFHSMCVRILR		97
PcrA	*Carnobacterium maltaromaticum*	RIAYLIEEKQVNPWNILAITFTNKAAKEMKERVNRLNVR--GGNDVWVSTFHSMCVRILR		100
PcrA	*Alloiococcus otitis*	RIAYLIQEKGVNPWNILAITFTNKAAGEMKDRVQKLVSQ--GGSGVWVSTFHSMCVRILR		100
PcrA	*Mitsuokella multacida*	RIANLLA-QGVAPYSILAITFTNKAATEMRERVDRMIGD--AAKDVWLSTFHSFCARFLR		99
PcrA	*Alkaliphilus metalliredigens*	RIAYLVEELGVSPYHILSITFTNKAAREMKERIHQLLGD--GFRDLWVSTFHSSCVRILR		95
PcrA	*Desulfotomaculum reducens*	RIAQILS-QGVRPYNILAITFTNKAANEMRARVENLVPQ--AAKDLWVTTFHSACLRILR		95
PcrA	*Listeria fleischmannii*	RIAYLIKEREVNPYNILAITFTNKAAREMKARVTGLLGG--IAESIWMSTFHSMCVRILR		101
PcrA	*Sporosarcina newyorkensis*	RIAYLVVEKQVYPSNILAITFTNKAAREMRERIDGLLGAG-SGERMWVSTFHSMCVRILR		102
PcrA	*Kurthia massiliensis*	RIAYLVIEKAVNPWNILAITFTNKAAREMRDRIDGLLGNG-SGQQMWVSTFHSMCVRILR		102
PcrA	*Marinococcus halotolerans*	RIAYLIREKMIPPYAILAITFTNKAAKEMKERVGKLVGP--VAEDIWMSTFHSMCVRILR		102
PcrA	*Planococcus antarcticus*	RIAYLVLEKQVYPSNILAITFTNKAAREMRNRIDGLLGHG-TGQRMWASTFHSMCVRILR		102
PcrA	*Lysinibacillus fusiformis*	RIAYLVIEKEVYPSKILAITFTNKAAREMRDRIDGILGNG-TGDSMWVSTFHSMCVRILR		102
PcrA	*Oceanobacillus iheyensis*	RIAYLMGEKDVSPRNILAITFTNKAAREMKERVSKLVGP--QGEYMWVSTFHSMCVRILR		101
PcrA	*Virgibacillus sp*	RIAYLLEEKDVAARNILAITFTNKAAREMKERVRRLVGS--ESDQMWVSTFHSMCVRILR		101
PcrA	*Caldibacillus debilis*	RIAYLIAEKGVNPYNILAITFTNKAAREMKERVEKIIGG--TAAEVWISTFHAMCVKILR		118
PcrA	*Halobacillus halophilus*	RIAYLLSEKDVAPRNILAITFTNKAAREMKDRVESLVGK--DGEKIWMSTFHSMCVRILR		101
PcrA	*Gracilibacillus halophilus*	RIAYLLGEKEVSPRNILAITFTNKAAREMKERITDLVGS--RGADMWVSTFHSMCVRILR		101
PcrA	*Bacillus cereus*	RIAYLLGEKGVAPWNVLAITFTNKAAREMRERIDTLVGP--EAEDIWISTFHSMCVRILR		100
PcrA	*Macrococcus caseolyticus*	RIAYLIDEKDVSPYKILAITFTNKAAKEMKERVKKIIGD--ESDVIWISTFHSMCVRILR		97
PcrA	*Laceyella sacchari*	RVAYLLAEKGIHPWNVLAITFTNKAAREMKERIGQLVGP--MAEEIWISTFHSMCVRMLR		101
PcrA	*Brevibacillus laterosporus*	RISYLVGYKQAFPWSILAITFTNKAAREMKERVERMVGRDAGADDIWISTFHSLCVRILR		102
PcrA	*Paenibacillus sp*	RIAYLIATRKAPPWAILAITFTNKAAREMQDRVSRLVGGP-EGRDIWVSTFHSMCVRILR		109
PcrA	*Thermobacillus composti*	RIAYLIGKKRVAPWSILAITFTNKAAREMQERVEALV-GP-GASDIWVSTFHSMCVRILR		103
PcrA	*Amphibacillus xylanus*	RIAYLLDEKDVSPRNILAITFTNKAAREMKDRVTNLIGP--EADQMWVSTFHSMCVRILR		102

Table 3.1 Conserved Cys96 in TxGx motif of PcrA. Cysteine residue in the TxGx motif is conserved across 27 organisms within 10 out 11 families. Leuconostocaceae family species have an alanine at this position.

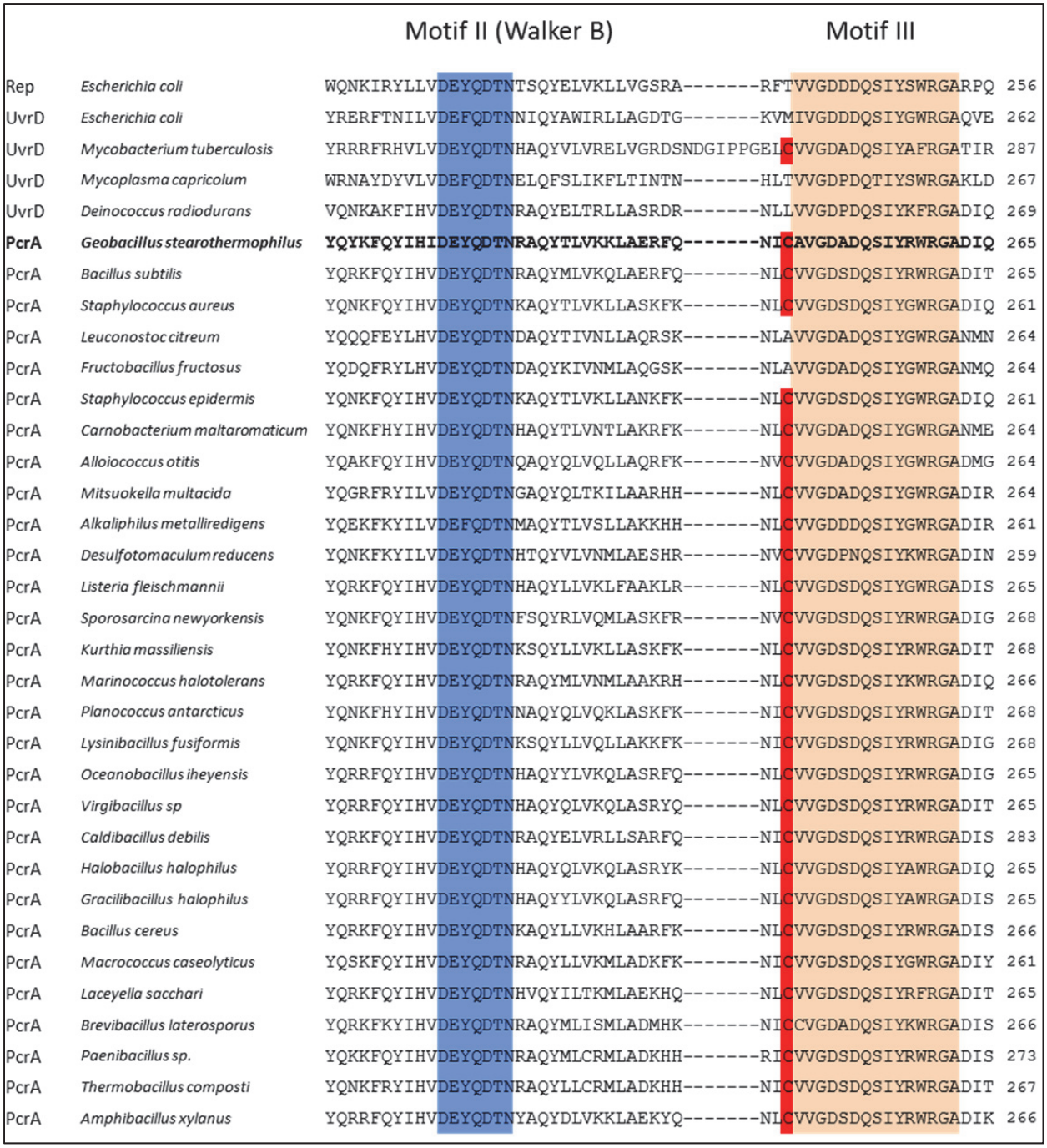


Table 3.2 Conserved Cys247 in motif III of PcrA. Cysteine residue in the motif III is also conserved across the 27 organisms shown here from 10 different families. Leuconostocaceae family species have an alanine at this position.

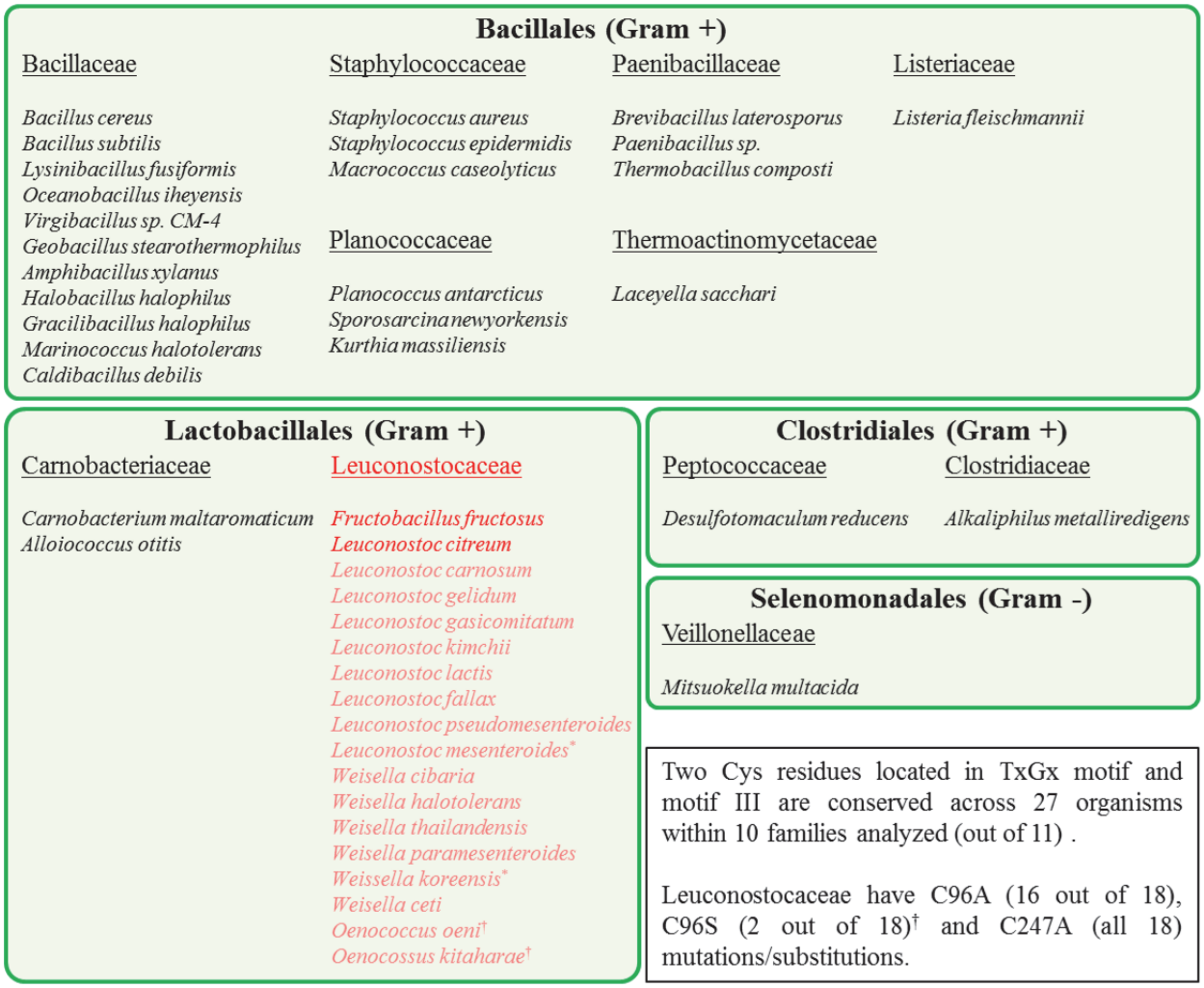


Table 3.3 Conservation of Cys96 and Cys247 across 11 families. 29 Organisms from 11 families used for PcrA sequence homology comparison. Only one family (red) had Ala instead of Cys, and ten families have conserved Cys residues. These 29 organisms and corresponding eleven families were selected randomly from an extensive PcrA library across ~250 species (~1100 including substrains) within the firmicutes phylum (mainly Gram +).

Chapter 4. Orientation dependence of FRET*

We have found that the efficiency of fluorescence resonance energy transfer between Cy3 and Cy5 terminally-attached to the 5'-ends of a DNA duplex is significantly affected by the relative orientation of the two fluorophores (*105*). The cyanine fluorophores are predominantly stacked upon the ends of the helix in the manner of an additional base pair, and thus their relative orientation depends on the length of the helix. Observed FRET efficiency is dependent upon the length of the helix, as well as its helical periodicity. By changing the helical geometry from B-form double-stranded DNA to A-form hybrid RNA/DNA, a marked phase shift occurs in the modulation of FRET efficiency with helix length. Both curves are well explained by the standard geometry of B- and A-form helices. The observed modulation for both polymers is less than calculated for a fully-rigid attachment of the fluorophores. However, a model involving lateral mobility of the fluorophores on the ends of the helix explains the observed experimental data. This has been further modified to take account of a minor fraction of unstacked fluorophore observed by fluorescent lifetime measurements.

4.1 Introduction

Fluorescence resonance energy transfer (FRET) has become widely used to report on distances over the macromolecular scale in biology (*175*), reviewed in references (*176, 177*). The method is highly sensitive, and consequently has been widely exploited in single molecule experiments in biological systems. Energy transfer results from dipolar coupling between the transition moments of two fluorophores, and the efficiency of the process (E_{FRET}) depends on the separation between the donor and acceptor fluorophores, raised to the sixth power. While such data are frequently interpreted on the assumption of a simple relationship between E_{FRET} and distance, E_{FRET} should also depend upon the relative orientation of the transition dipole vectors.

The orientation dependence is likely to be most significant where the fluorophores are constrained. This has been demonstrated experimentally using a fluorophore that was

* The content in this section was reproduced with permission from "Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids", Iqbal, A., Arslan, S., Okumus, B., Wilson, T. J., Giraud, G., Norman, D. G., Ha, T. & Lilley, D. M. J.. *Proc Natl Acad Sci USA*,**105**(32):11176-81 (2008). Copyright (2008) National Academy of Sciences, U.S.A

terminally affixed to duplex DNA by two points of covalent attachment (*178*), thereby seriously constraining its motion. This situation is not typical of most FRET studies involving nucleic acids. Fluorophores are normally tethered by a single point of attachment, and in theory would be significantly less constrained. But if the fluorophores adopt a rigid manner of attachment to the helix, orientation dependence could be observed.

Cy3 and Cy5 are a commonly used fluorophore pair, especially in single-molecule experiments. Our earlier NMR studies have shown that when these are attached to the 5'-termini of duplex DNA via a 3-carbon linker to the 5'-phosphate they are predominantly stacked onto the ends of the helix in the manner of an additional basepair (*179*). This would provide a favorable situation in which the orientation dependence of FRET could be observed, and we have therefore studied a series of DNA and DNA-RNA hybrid duplexes in order to seek this effect. We find that E_{FRET} values reduce with duplex length, but also exhibit a modulation with twice the periodicity of the helices that is consistent with the anticipated orientation effect. Thus it will be necessary to take fluorophore orientation into consideration when interpreting FRET data in terms of distances in some circumstances.

4.2 Background theory

The variation of E_{FRET} with the separation between donor and acceptor fluorophores (R) is given by (*2*) :

$$E_{FRET} = \frac{1}{1 + \left(R / R_0 \right)^6}$$

where R_0 is the distance at which energy transfer is 50% efficient. It depends upon the spectroscopic properties of the fluorophores and the medium, given by:

$$R_0^6 = \frac{0.529 \cdot \kappa^2 \cdot \Phi_D \cdot J(\lambda)}{N \cdot n^4}$$

where the units of R_0 and the wavelength λ are cm. κ^2 describes the relative orientation of the fluorophores (see below). Φ_D is the quantum yield of the donor, *N* is the Avogadro number and *n* is the index of refraction of the medium. $J(\lambda)$ is the spectral overlap integral, given by :

$$J(\lambda) = \frac{\int_0^\infty \phi_D(\lambda) \cdot \varepsilon_A(\lambda) \cdot \lambda^4 d\lambda}{\int_0^\infty \phi_D(\lambda) d\lambda}$$

where $\phi_D(\lambda)$ is the spectral shape of donor emission and $\varepsilon_A(\lambda)$ that for acceptor excitation (M^{-1} cm^{-1}). We have measured the overlap integral for Cy3 and Cy5 terminally attached to dsDNA as $J(\lambda) = 7.2 \times 10^{-13}\ M^{-1}\ cm^3$.

The largest potential uncertainty in the extraction of distance information from E_{FRET} lies in the orientation term κ^2, given by:

$$\kappa^2 = \left(\cos\Theta_T - 3.\cos\Theta_D.\cos\Theta_A\right)^2$$

where the angles Θ_T, Θ_D and Θ_A are defined in Figure 4.1 A, B. This can take values between 0 and 4, or between 0 and 1 if the transition moments are constrained to parallel planes. If the fluorophores undergo isotropic reorientation within the excited state lifetime of the donor then κ^2 = 2/3, and in most studies this is assumed to apply. While this is a good approximation for freely-mobile fluorophores like fluorescein that are terminally attached to double-stranded nucleic acids, it is rather less probable for cyanine dyes in this situation. Using NMR we have shown that Cy3 and Cy5 are predominantly stacked onto the end of dsDNA when coupled to the 5'-terminal phosphates via C_3 linkers (*179, 180*) (Figure 4.1 C, D).

The transition moments for the π to π^* transitions lie in the plane of the indole rings close to the polymethyne linker (*180*) and are therefore directed in planes that are approximately parallel to each other. This situation approximates to $\Theta_D = \Theta_A = 90°$, and where Θ_T depends on the length and geometry of the DNA helix and the angles between the terminal basepairs and the transition moments. This is simulated in Figure 4.2, on the assumption that the fluorophores are rigidly attached to the ends of the DNA. It can be seen that the efficiency of energy transfer is strongly modulated by the length of the helix in a periodic manner, falling close to zero twice per helix rotation when the transition moments are perpendicular (i.e. $\cos\Theta_T = 0$). This 'bouncing ball' type dependence is not significantly altered if the fluorophores are inclined relative to the terminal basepairs by 20°, as $3.\cos\Theta_D.\cos\Theta_A$ is small and the behavior remains dominated by the variation in Θ_T. The variation of E_{FRET} with length with a constant value of κ^2 = 2/3 is also shown.

4.3 Results and Discussion

We have investigated the variation of FRET efficiency experimentally for a series of terminally-labeled DNA duplexes of lengths 10 – 24-bp, thus covering more than one complete helical turn. Each species was generated by hybridizing extensively-purified complementary strands with Cy3 or Cy5 separately attached to their 5'-termini via a three-carbon linker (Figure 4.8). E_{FRET} was measured in both ensemble and single-molecule experiments.

4.3.1 Modulation of energy transfer efficiency in a DNA duplex series

Each member of the series was studied in free solution under steady-state conditions, and E_{FRET} was calculated using the acceptor normalization method (*181*). The resulting efficiencies are plotted as a function of helix length in Figure 4.3 A (filled circles). Overall the E_{FRET} values decrease with helix length, yet there is a clear periodic modulation in phase with that anticipated for simple orientational dependence based on B-form helical geometry (Figure 4.2). Minima are observed at 13 and 18-bp, and maxima at 14 and 19-bp. The period is therefore ≈5-bp, which is half the structural periodicity of a B-form double helix; this is consistent with the anticipated orientational dependence of dipolar coupling. It should be noted that the observed modulation has twice the frequency of that observed where the inter-fluorophore separation varies with helix length due to off-axial positioning of freely mobile fluorophores (*182*).

The values of E_{FRET} for the DNA duplex series was measured by an alternative approach based on single-molecule methods. This was done for two main reasons. First, we were concerned that the ensemble results might be distorted by the presence of molecules in which Cy3 was not active. Second, the majority of experiments using the Cy3-Cy5 combination are performed with single-molecule methodology, and so this is directly relevant. E_{FRET} was measured from single DNA molecules encapsulated in phospholipid vesicles (*98*) that were tethered to a quartz slide, using total internal reflection microscopy (*78*). This avoided any perturbation of DNA structure that might otherwise have arisen if there was a direct tether to the surface, and only molecules with active donor and acceptor fluorophores were analyzed. E_{FRET} values of active species were calculated from the donor and acceptor wavelength channels, and histograms generated for many single molecules. The plot of E_{FRET} vs helix length (Figure 4.3)

is very similar to that generated from the ensemble data, with modulation of almost identical period and phase.

Thus we observe the same modulation of FRET efficiency with helix length irrespective of the method of measurement. The period and phase of the modulation are fully consistent with the expected dependence upon the relative orientation of the transition moments of the fluorophores, i.e. the variation of κ^2.

4.3.2 A dynamic models can account for the extent of modulation

While the positions of the maxima and minima of the experimental profiles are fully consistent with modulation due to fluorophore orientation, the peaks are clearly less sharply defined than that calculated for rigid fluorophores (Figure 4.2), suggesting that the data are averaged by some dynamic process. Motions in which the fluorophores tilt away from their stacked position on the end of the DNA cannot explain this effect. We therefore simulated the data using an alternative model in which the fluorophores are allowed to move laterally on the end of the helix, with a Gaussian distribution of probabilities of chosen half-width – this corresponds to a Boltzmann population of conformations where both fluorophores and the intervening DNA helix collectively act as a torsional spring. Using standard helical parameters for B-form DNA with 10.5-bp/turn and a helical rise of 3.6 Å, together with mean angles between the terminal basepair and the transition moments of Cy3 and Cy5 of 32 and 30° respectively (measured from the NMR structures), we generated simulations of the E_{FRET} vs helical length curves. Setting a half-width at half-maximum (HWHM) of 55° provided excellent agreement with the shape of the experimental curve (Figure 4.3), reproducing the maxima and minima in closely similar positions. The only parameter adjusted was the refractive index; n = 1.38 gave the best agreement to the absolute values of E_{FRET}.

4.3.3 A fraction of Cy3 is unstacked from the DNA

While the lateral motion can account for the observed data very well, it does not provide a unique solution. A model in which one fraction of the fluorophores are stacked while the remaining fluorophores are unstacked and freely mobile with $\kappa^2 = 2/3$ provided equally good agreement (see below). While our NMR data indicate that most of the Cy3 will be stacked on the end of the DNA (*179*), a minor fraction that is unstacked would not be detected in these

experiments. Levitus and co-workers (*183*) have provided time-resolved fluorescence data that indicate that there is a fraction of Cy3 that is unstacked when attached to the 5'-terminus of double-stranded DNA. We therefore measured lifetime distributions for Cy3 attached to the 5'-terminus of DNA duplexes such that the fluorophore was in exactly the same environment as in the FRET experiments.

Singly Cy3-labelled duplexes of 16 and 22-bp were analyzed using time-correlated photon counting, and the decay curves obtained fitted to a number of exponential functions (Figure 4.4 A). For each sample our data were significantly better fitted using three exponentials compared to two (Table 4.1). Two species had relatively long lifetimes of 1.04 and 1.91 ns, with relative amplitudes of 61.6 and 21.4 % respectively, while the third species had a lifetime of 390 ps and an intensity of 16.9 % (Table 4.1). We assign the short lifetime species to unstacked Cy3, where the excited state is relaxed by *cis-trans* photoisomerization occurring in the polymethyne linker (*184*). The fluorescent lifetime becomes longer if rotation about the polymethyne linkage is prevented, for example by steric constraints in Cy3B (*183*), or due to stacking. Thus the two species we observe with longer lifetimes are likely to be stacked onto the end of the helix in a manner similar to that we have observed by NMR (*179*). Indeed, the longer lifetime is very close to the lifetime of 2 ns expected for a Cy3 species that is unable to photoisomerize, calculated from the radiative fluorescence rate for Cy3 of 5×10^8 s^{-1} (*183*). The intermediate lifetime probably arises from Cy3 in a series of environments in which rotation about the polymethyne linkage of the fluorophore is constrained but not prevented. These are unlikely to be freely mobile, because such a high proportion of unstacked fluorophore is not consistent with the NMR data. However, the lateral motion could partially expose the fluorophore such that some segmental rotation within the polymethyne linker becomes possible, shortening the fluorescent lifetime.

Assuming that the dynamic properties of the chemically-similar Cy5 are comparable to Cy3, up to 31 % of the fluorescent emission could be due to molecules in which one or both fluorophores are in an unstacked conformation. We therefore simulated our data assuming that 31 % of the molecules had freely mobile fluorophore such that $\kappa^2 = 2/3$ (Figure 4.4 B). This provides an equally good fit to the experimental data (both ensemble and steady-state) compared to the fully-stacked model, giving a HWHM of 42° for the fluorophores remaining stacked on the helix. However, simulations show that using a free fraction of fluorophore plus a

stacked fraction with no lateral flexibility cannot explain the data. First, the minima and maxima are no longer in the correct positions. Second, we require a free fraction of ≈75% to approximate the experimental data, and this is plainly in contradiction with our earlier NMR data (*179*).

4.3.4 An altered periodicity of modulation in a DNA-RNA hybrid duplex series

The modulation of E_{FRET} with twice the frequency of the DNA helix provides strong evidence for the orientation dependence of the dipolar coupling. And this suggests a further test of the model. On the basis of this model we expect that an A-form helix should produce a similar modulation, but with an altered period and phase since this helical conformation is less tightly wound than the B-form helix of double-stranded DNA. We therefore tested the prediction by the construction of a new series of terminally Cy3-Cy5-labeled duplex species with one DNA and one RNA strand. These hybrid duplexes should adopt an A-form helix (*185-187*). FRET efficiency was measured using both steady-state ensemble spectroscopy and single encapsulated molecules by TIR microscopy as before. The data are presented in Figure 4.5, with the data for the DNA duplexes reproduced in grey for comparison. The profiles have the same appearance as those for the DNA duplexes, with a clear modulation superimposed onto a generally reducing E_{FRET} as duplex length increases. However, the modulation is clearly not in phase with the DNA data, with maxima observed at 11, 17 and 22-bp. The efficiency of energy transfer is modulated with twice the periodicity of a helix corresponding to ≈12-bp/turn, that is, in good agreement with that expected for an A-form helix.

We also carried out time-resolved fluorescence lifetime measurements of Cy3 attached to DNA/RNA duplexes of 16 and 22-bpin length (Table 4.1, Figure 4.9). As with the DNA, the data were best fitted using three species, with lifetimes (amplitudes) of 330 ps (6.1 %), 990 ps (46.4 %) and 1.84 ns (47.5 %). As with the DNA, we attribute the short lifetime species to unstacked Cy3 molecules, and the remaining species to stacked fluorophores. These data indicate that the Cy3 is stacked more fully on the A-form helix compared to the DNA duplex.

On the basis of these results we simulated our FRET efficiencies as a function of DNA/RNA duplex length based on an A-form helix with 12-bp/turn and a rise of 3 Å and an unstacked fraction of fluorophores of 12 % with $\kappa^2 = 2/3$ (Figure 4.5). Good agreement with the

experimental data was obtained by inclusion of lateral fluorophore mobility with HWHM = 42° for the stacked fluorophores, and a refractive index n = 1.33. To obtain the best absolute agreement, a value of fluorescence quantum yield of 0.35 was used for the DNA-RNA duplexes, consistent with the greater proportion of the long-lifetime species.

4.4 Conclusions

The observed modulation of the distance dependence of FRET for the DNA and hybrid DNA-RNA series, and the agreement with simulations based on B and A form helices, provide strong evidence for the orientational dependence (i.e., κ^2) of energy transfer efficiency by using the commonly employed Cy3-Cy5 donor–acceptor pair. Our experiments unequivocally establish that Förster transfer obeys the orientation dependence as expected for a dipole–dipole interaction. In many situations a simple inverse-distance interpretation of FRET efficiency will provide an adequate qualitative interpretation, but our data show that the common assumption that FRET is a monotonic function of distance can actually fail under certain circumstances. For the extraction of precise distance information, it will be necessary to take account of the orientation dependence. From our data we calculate that the assumption that $\kappa^2 = 2/3$ could result in an error of up to 12 Å in distance estimation in some circumstances, notably when the transition moments are close to perpendicular. The discrepancy might be reduced if significant flexibility can be introduced into the linker connecting the two fluorophores. For example, changing the length and characteristics of the covalent tether might result in a flexible fluorophore, although this cannot be assumed *a priori*. If the fluorophores remain stacked on the helix, the orientation effect could lead to misassignment of states in single-molecule experiments if the assignment is made based only on the FRET efficiencies, especially in more complex systems with multiple states. Therefore, additional control experiments should be performed to provide independent support for the assignments. On the positive side, the orientation effects in FRET could be a valuable tool in structural biology. A full understanding of the orientation dependence could greatly extend the use of FRET measurements to provide both accurate distance and angular information.

4.5 Materials and Methods

4.5.1 Synthesis and preparation of duplex species

Deoxyribooligonucleotides were synthesized using standard phosphoramidite chemistry, and ribooligonucleotides were synthesized using 2'-*t*-BDMS ribonucleoside β-cyanoethyl phosphoramidites (*188*), as described in ref. (*189*). Cy3 and C5 were added to 5'-termini as phosphoramidites at the end of synthesis as required. Fully deprotected oligonucleotides were purified by electrophoresis in 20% polyacrylamide gels containing 7 M urea, and recovered by electroelution. The cyanine-conjugated strands were further purified by reversed-phase chromatography using a C_{18} column eluted with a gradient of acetonitrile in 100mM triethylammonium acetate. Duplex species were assembled by mixing stoichiometric quantities of purified DNA (Cy3 or Cy5 labeled) or RNA (Cy3 labeled) in 90 mM Tris-borate (pH 8.3), 25mM NaCl, cooling slowly from 95°C to 4°C. Duplexes were then purified by electrophoresis in 20% polyacrylamide under non-denaturing conditions in 90 mM Tris-borate (pH 8.3), 25 mM NaCl. Gel fragments containing the required duplexes were excised, and the double-stranded nucleic acids recovered by electroelution. The purified DNA or DNA/RNA species were dissolved in 90 mM Tris-borate (pH 8.3) and absorption spectra recorded from 220 to 800 nm using a Cary 1E UV-visible spectrophotometer. The full sequences of all the duplex species used in this study are listed in Table 4.2.

4.5.2 Steady-state ensemble fluorimetry

Fluorescence emission spectra were recorded at 4°C using an SLM-Aminco 8100 fluorimeter in 90 mM Tris-borate (pH 8.3). Spectra were corrected for lamp fluctuations and instrumental variations, and polarization artifacts were avoided by crossing excitation and emission polarizers at 54.7°. E_{FRET} were measured with the acceptor normalization method (*181*), using emission spectra (550-720 nm) excited at 535 and 600 nm.

4.5.3 Modeling the dependence of FRET efficiency on duplex length for rigidly stacked fluorophores

We calculate the inter-fluorophore distance for each duplex R as

$$R = ((L - 1) * H) + D$$

where L is the length of the helix (bp), H is the helical rise per-bp step, and D is the additional axial separation for the two fluorophores. The angle between the transition moments A is calculated as

$$A = ((L - 1) * T) + C3A + C5A$$

where T is the twist angle for each base pair, and $C3A$ and $C5A$ are the rotations of Cy3 and Cy5 relative to the terminal base pairs. For B form DNA,

H = 3.6 Å and T = 36°, and D = 8 Å, $C3A + C5A$ = 62°.

The orientation factor is calculated as

$$\kappa^2 = (sin\Theta_D * sin\Theta_A * cosA - 2 * cos\Theta_D * cos\Theta_A)^2$$

For the case where the fluorophore planes are parallel and close to coaxial (i.e., Θ_D, Θ_A ~90°), this reduces to

$$\kappa^2 = cos^2 A$$

κ^2 and hence R_0 are calculated for each length, and thus the value for E_{FRET}. This calculation has been performed over 141 steps from 10 to 24-bp in Kalaidagraph, and interpolated (Figure 4.2).

4.5.4 Modeling the dependence of FRET efficiency on duplex length with lateral mobility of fluorophores

As before, the inter-fluorophore distance for each duplex is given by

$$R = ((L - 1) * H) + D$$

and the mean angle between the transition moments A is given by

$$A = ((L - 1) * T) + C3A + C5A$$

For each species, we set up an array of angles (AA) over a ±100° range about the mean, and calculate κ^2 and hence R_0 for each. The value of E_{FRET} is calculated for each angular position (E_{AA}), and the resulting distribution summed, weighted by its distance from the mean angle by using a Gaussian distribution, that is,

$$E_{FRET} = \Sigma (E_{AA} * P)$$

where

$$P = exp - (AA^2/1.44H^2)$$

where H is the half-width. The sum of P is normalized to unity. This simulation procedure was implemented in a MATLAB program.

Where a fraction of unstacked fluorophores is allowed, that fraction is calculated by using $\kappa^2 = 2/3$, together with an increment to the inter-fluorophore distance (7 Å was found to give the best agreement with experimental data). The resulting E_{FRET} is calculated as a linear combination of the contributions from the stacked fluorophore (calculated as above) and the freely mobile fluorophore.

4.5.5 Single-molecule measurements of encapsulated duplex species

Single duplex molecules were studied trapped within phospholipid vesicles in 10mM Tris·HCl (pH 8.0), 50mM NaCl. Individual DNA or RNA/DNA duplex species were encapsulated in phospholipid vesicles comprising a 100:1 mixture of either L-a-phosphatidylcholine (EggPC) or 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) with 1,2-dipalmitoyl-*sn*-glycero-3-phosphoethanolamine-N-(cap biotinyl) (DPPE-biotin)(Avanti Polar Lipids) by repeated extrusion through a polycarbonate membrane containing 200 nm pores (Osmonics) (*97*) as described (*98*). This generated 200 nm diameter vesicles, with a 1:1 molar ratio of duplex to vesicles. The vesicles were conjugated via NeutrAvidin (Pierce) to biotin-functionalized polyethylene glycol (PEG)-coated quartz slide surfaces. Encapsulated molecules were excited at 532 nm by prism-based total internal reflection microscope at 12°C (*78*). Up to 100 images of the surface at Cy3 and Cy5 emission wavelengths were obtained by using an Andor intensified CCD camera, with an integration time of 100 ms for durations of 2–3 s. Dual-view design of the emission pathway allowed us to image the Cy3 and Cy5 signals simultaneously. E_{FRET} values were calculated from averaged Cy3 and Cy5 intensities over 300- to 500-ms periods for each molecule in the images. Data were corrected for crosstalk between Cy3 and Cy5 channels, backreflection from the dichroic mirror surface in the dual-view emission pathway, and the background (see 4.5.6). Histograms of E_{FRET} values from 5,000 to 32,000 molecules are plotted in Figure 4.6. Each histogram contained two major peaks; one at $E_{FRET} \approx 0$ corresponding to duplexes with active Cy3 only, and another at higher E_{FRET} resulting from duplex molecules with an active Cy3-Cy5 pair. However, for some constructs, an additional, broad peak was observed at intermediate FRET values because of some vesicles containing multiple DNA molecules with at least one

photobleached Cy5. In such cases, we fit the data with three Gaussian peaks to separate the contribution of the important species. Errors on the reported E_{FRET} values represent the standard deviation of randomly sampled subsets of the data, each comprising ≈1,000 molecules.

4.5.6 Data analysis of single molecule experiments

From the short time traces average donor (I_D) and acceptor (I_A) intensities are calculated over a period of 300 to 500 ms. These values are recorded for up to 40,000 molecules for further processing. Using a scatter plot of total intensity vs. E_{FRET}, lower and higher thresholds are determined so that E_{FRET} populations of Cy3-Cy5 pairs are selected. All of the molecules outside the threshold limits are discarded. Because the time traces are short we usually don't see the photobleached flat lines in time traces to be used as the background value. We employ two techniques to determine the background values (I_{Do} and I_{Ao}) as shown in Figure 4.7. In the first method, we make histograms of I_D and I_A that are measured from the last 3–5 data points of each time trace so that fluorophores are more likely to photobleach (Figure 4.7 C). The flat lined I_D and I_A will appear as a peak near 0. Peak positions are then used as the background values. The second method involves the dual histogram of I_D and I_A pairs from all of the molecules of the duplex specie in question. A color density plot of such a distribution is shown in Figure 4.4 B. Because, in theory, Cy3 and Cy5 intensity ratio is constant, E_{FRET} populations in the dual histogram can be fitted to a line. The E_{FRET} of Cy3-Cy5 pairs shows up at a higher I_A value, whereas Cy3 only duplexes populate lower I_A values. Fitting each of these populations to a line and finding the intersection point provides us with the background values, I_{Do} and I_{Ao}. Sometimes some I_D and I_A population is found near the origin because of non-active Cy3-Cy5 pairs, the center of which should also coincide with the intersection. By using these guidelines, we determine the background values and subtract them from the mean intensities to calculate the apparent E_{FRET} value for each molecule.

$$E_{FRET} = (1 + [I_D - I_{Do}] / [I_A - I_{Ao}])^{-1}$$

From these values we prepare the initial FRET histograms, which usually comprise two dominant peaks. The first peak is the donor-only molecules with non-active Cy5. It is not at zero, because some of the donor emission leaks into the acceptor channel, yielding a donor-only peak. The peak position of this value, p, is then used to perform crosstalk correction. Another factor in our apparatus arises from the long-pass dichroic mirror in the dual-view emission pathway, which reflects a fraction of the Cy5 signal into the Cy3 channel. We measured the ratio

of this backreflection by using a construct where the Cy3 and Cy5 are so close to each other that all of the emission comes from Cy5, and Cy3 emission should be zero. E_{FRET} from that substrate should be 1, but we only measure ≈0.88 indicating a 12% reflection ratio. The experiment was repeated with a number of substrates where the fluorophores are at very close proximity; the average value for the reflection ratio is calculated to be r = 0.12 with a very narrow distribution.

True FRET efficiency (E_{FRET}) of each molecule is then calculated from averaged donor and acceptor emissions along with the crosstalk correction factor, p, and backreflection ratio, r, using the formula:

$$E^*_{FRET} = (E_{FRET} - p) / (1 - p - r)$$

The gamma factor in our system (*190*) was measured by looking at Cy5 photobleaching or blinking traces, where the decrease in the acceptor intensity should be equal to the increase in donor intensity in theory. If they are not equal, then the ratio of donor-to-acceptor emissions should be scaled accordingly. This is called gamma correction. This value was measured and found to be 1 in our system, and therefore we did not apply any additional correction. E^*_{FRET} values from each molecule were then used to generate the histograms for DNA and RNA/DNA duplexes. The peak positions of Gaussian fits were insensitive to the bin size in a range from ≈0.005 to ≈0.05; we chose 0.02. Histograms clearly show the peaks for Cy3-Cy5 pairs and the donor-only signal E^*_{FRET}=0. E^*_{FRET} values less than that of the single pair populate the intermediate E_{FRET} region; these arise because we use all of the molecules indiscriminately within the threshold boundaries, and because of multiple duplexes in single vesicles with at least one photobleached Cy5. The magnitude of intermediate peaks depends on the relative population of the vesicles with multiple duplexes. We accounted for the intermediate peaks by adding an additional Gaussian peak in the fits. As the E^*_{FRET} value gets lower, and the two major peaks starts to overlap, the presence or absence of the intermediate Gaussian peak does not change the higher E^*_{FRET} peak; so, for the longest duplex species, the intermediate peak is omitted.

4.5.7 Time-resolved fluorimetry

Time resolved fluorescence intensity measurements were performed by time-correlated single-photon counting (TCSPC), using an FL 920 spectrometer (Edinburgh Instruments) (*191*). The excitation source was a titanium sapphire (Coherent) laser with 200-fs pulse duration operated at 76 MHz. Its fundamental output was sent through a pulse picker (Coherent 9200)

and a harmonic generator (Coherent 5-050) to obtain 450-nm pulses at 9.5 MHz. The excitation beam was attenuated as needed to avoid pile-up effects. Fluorescence emission was detected using a monochromator and a cooled microchannel plate photomultiplier tube (Hamamatsu C4878). The instrument response function, measured by scattering the excitation beam from a dilute suspension of colloidal silica (Ludox), was less than 100 ps FWHM. Fluorescence decay curves were recorded on a time scale of 20 ns, resolved into 4096 channels, to a total of 10,000 counts in the peak channel. Decay curves were analyzed using a standard iterative reconvolution method in the F900 (Edinburgh Instruments) software packages, on the basis of a multiexponential decay function. The quality of fit was judged on the basis of the reduced chi-square statistic, and the randomness of residuals.

4.6 Figures

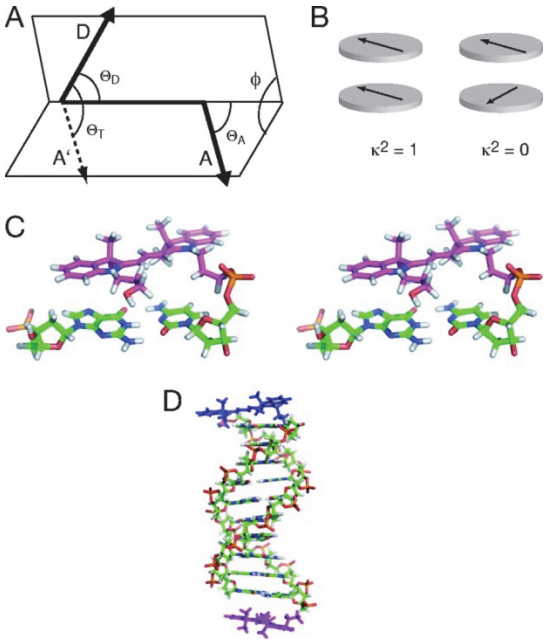


Figure 4.1 Orientation of transition moments of cyanine fluorophores terminally attached to double-stranded DNA. (*A*) The orientation parameter κ^2. The transition dipole vectors for the coupled donor and acceptor fluorophores are indicated by the arrows, labeled D and A. Vector A' is generated by the in-plane translation of vector A to share its origin with vector D. The definition of κ^2, given in 4.2, is based on the angles shown. (*B*) If the fluorophores lie in parallel planes, the orientation parameter simplifies to $\kappa^2 = \cos^2\theta_T$ and varies between 0 and 1. The schematic shows the limiting cases, where the transition moments are parallel ($\kappa^2 = 1$) and crossed ($\kappa^2 = 0$). Note that these conditions will occur twice per complete rotation of one fluorophore relative to the other around their common axis. (*C*) Parallel-eye stereo image showing the structure of Cy3 stacked onto the terminal base pair of a DNA duplex, as determined by NMR. (*D*) A molecular graphics model of a duplex with Cy3 and Cy5 fluorophores attached to the 5' termini via C_3 linkers. This was generated by using our NMR structures of Cy3 and Cy5 attached to duplex DNA. Note that the fluorophores lie in approximately parallel planes, and that the angular relationship between them (and thus their transition moments) will depend on the length of the DNA helix and its helical periodicity.

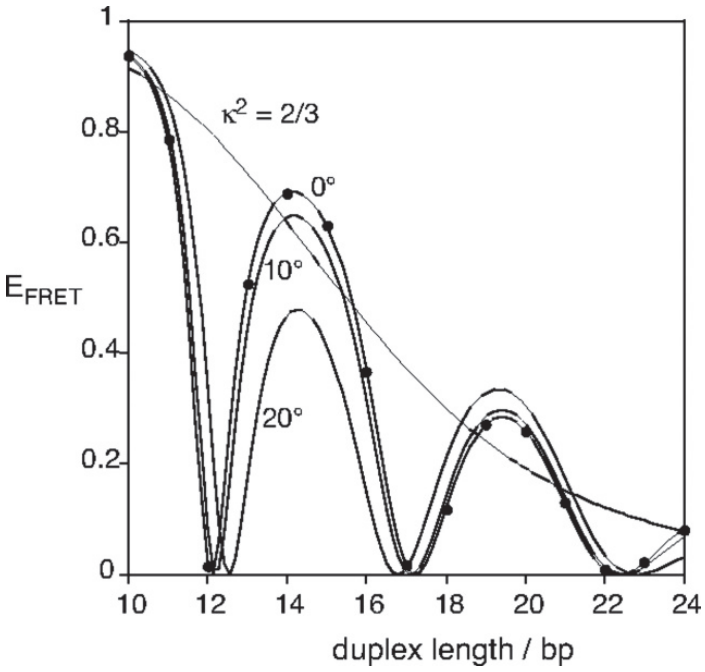


Figure 4.2 Simulation of the dependence of calculated efficiency of energy transfer between Cy3 and Cy5 terminally attached to duplex DNA as a function of the length of the helix. The calculations are based on the extremes of complete fluorophore mobility ($\kappa^2 = 2/3$) and total rigidity of the fluorophore stacked on B form DNA. The latter calculations were performed for fluorophores lying in planes that are perpendicular to the helical axis (0°), or with the planes inclined to the axis by 10 or 20°. These calculations were based on a measured value of the spectral overlap integral $J(\lambda) = 7.2 \times 10^{-13}$ M^{-1} cm^3, giving $R_0 = 60.1$ Å when $\kappa^2 = 2/3$. For the case of rigid and parallel fluorophores, the points plotted are the values for integral numbers of base pairs.

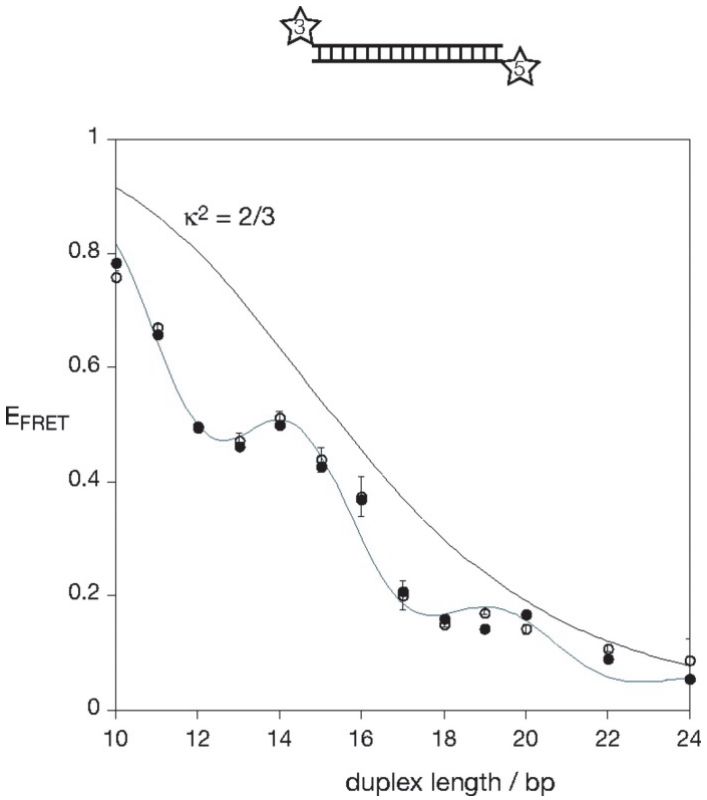


Figure 4.3 Efficiency of energy transfer for Cy3, Cy5-labeled DNA duplexes as a function of duplex length. E_{FRET} was measured for each duplex species, from an ensemble of molecules in the steady state (filled circles) or from phospholipid vesicle-encapsulated single molecules (open circles). Estimated errors are shown on the single-molecule data. The data have been simulated by using a model in which the fluorophores undergo a lateral motion within the plane perpendicular to the helical axis, giving a Gaussian distribution of probabilities about the mean relative angle set by the NMR structures. The fluorescent quantum yield for Cy3 was 0.30 (*183*), and the refractive index 1.33. The half-width giving the best agreement to the experimental data was 55°. Standard B form geometry of the DNA helix was used, with 10.5-bp per turn and a helical rise of 3.6 Å per-bp step. The dependence for fully flexible fluorophore ($\kappa^2 = 2/3$) is also plotted.

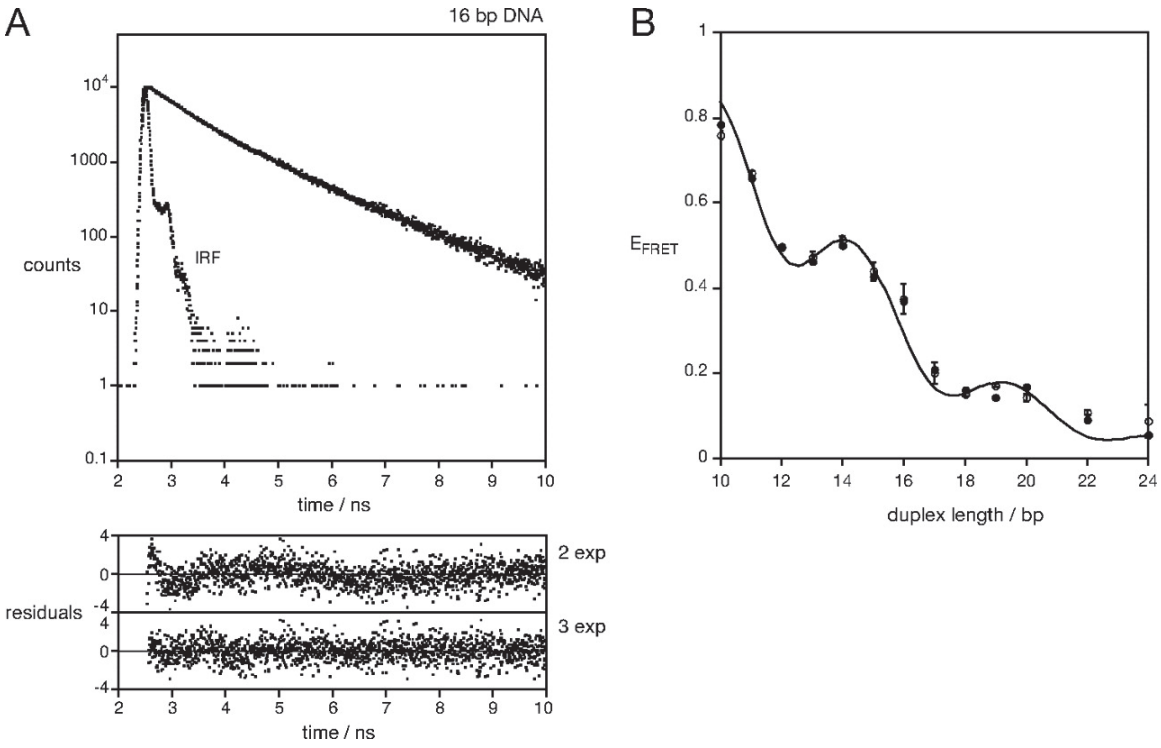


Figure 4.4 The existence of an unstacked fraction of fluorophore. (*A*) Time-resolved fluorescent lifetime analysis of Cy3 attached to double-stranded DNA. Fluorescent decay curve for Cy3 attached to a 16-bp DNA duplex, showing the experimental data and the instrument response function (IRF), and the fit to three exponential functions (line). Plots of weighted residuals for fitting the data to two (χ^2 = 1.33) or three (χ^2 = 1.06) exponential functions are shown below. (*B*) On the basis of this analysis, the E_{FRET} data for the DNA duplexes were simulated by using a model in which the fraction of the fluorophores indicated by the fluorescent lifetime analysis was given free mobility with κ^2 = 2/3, whereas the remaining fluorophore was constrained to undergo lateral motion with a distribution of species as before. The simulation (line) has been plotted with the experimental data (points) obtained from ensemble (filled circles) and single-molecule (open circles) experiments. Both datasets were found to be well described by a simulation in which 31% of the molecules had a mobile fluorophore extended by a further 7Å, and the remaining fraction of stacked fluorophore underwent lateral motion with a Gaussian half-width of 42°. The fluorescent quantum yield for Cy3 was 0.30, and the refractive index 1.33. Standard B form geometry of the DNA helix was used.

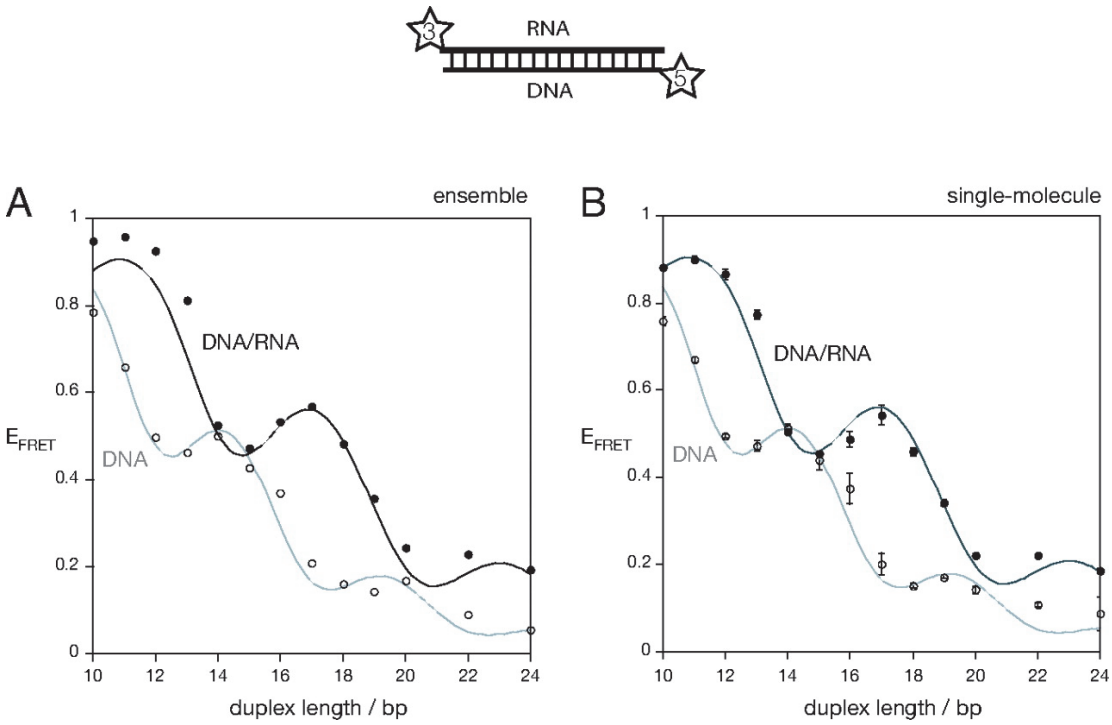


Figure 4.5 Efficiency of energy transfer for Cy3, Cy5-labeled hybrid RNA/DNA duplexes as a function of duplex length. The duplexes were constructed by using the same Cy5-labeled DNA strands as before, hybridized to complementary Cy3-labeled RNA strands. E_{FRET} was measured for each duplex species, from an ensemble of molecules in the steady state (*A*) or from phospholipid vesicle-encapsulated single molecules (*B*). The E_{FRET} values are plotted (filled circles) as a function of helix length. The estimated errors are shown on the single-molecule data. The black lines show simulation of the data based on the geometry of an A form helix with a periodicity of 12-bp per turn and a rise of 3 Å/bp step. A fraction of freely mobile fluorophore of 12% was assumed from the time-resolved fluorescence measurements of Cy3 attached to an RNA/DNA duplex (see text). A refractive index of n = 1.33 and quantum yield of 0.35 were used for the efficiency calculations. For both datasets, a half-width for the distribution of lateral motion of 42 degrees gave the optimal agreement with the experimental data. The data and simulations for the DNA duplexes are shown for comparison by using open circles and gray lines. The phase shift between the two helical forms is very clear.

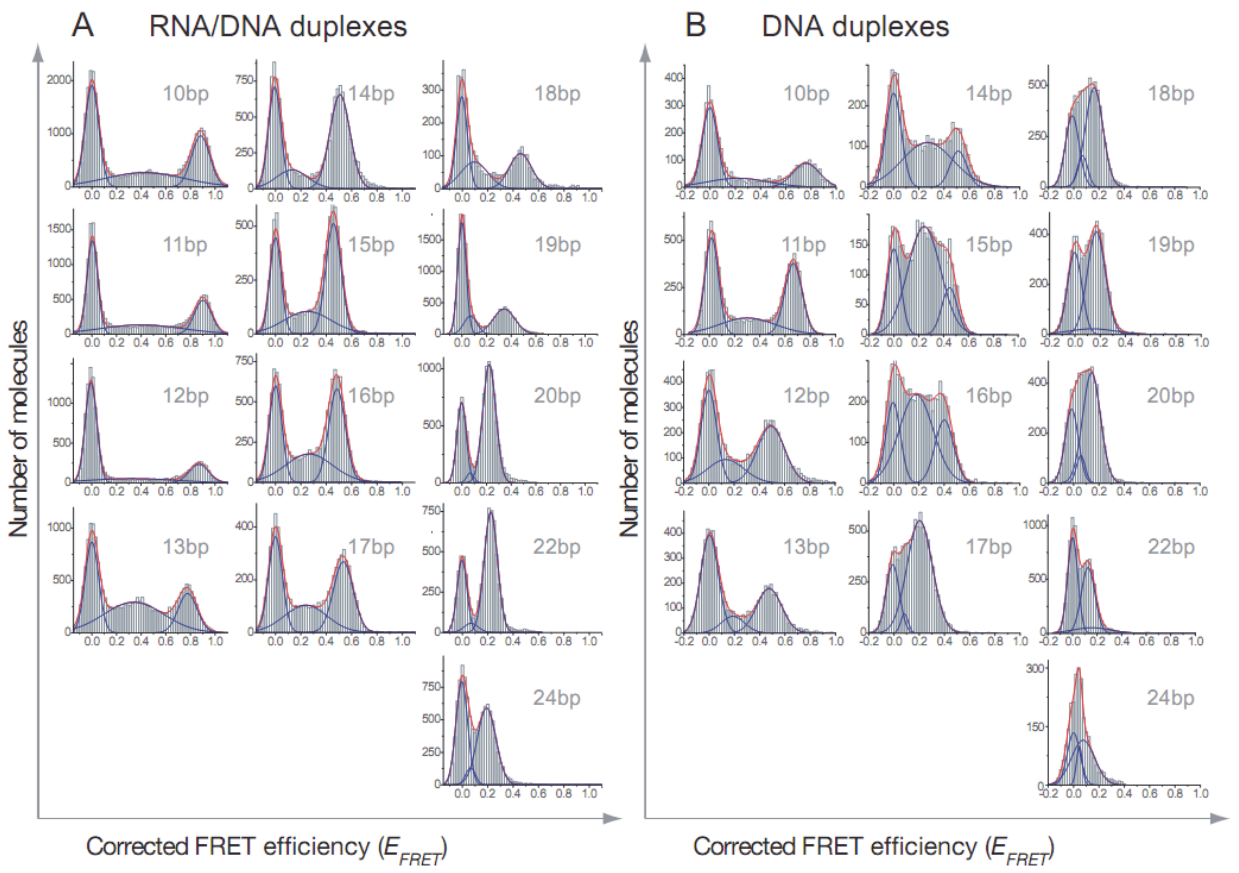


Figure 4.6 Histograms of FRET efficiency (E_{FRET}) for phospholipid vesicle-encapsulated DNA and RNA/DNA duplex species. The y axes represent the number of molecules in each FRET efficiency bin. Mean E_{FRET} value for each molecule have been calculated from Cy3 and Cy5 fluorescent intensities that are corrected for the crosstalk between Cy3 and Cy5 channels, back reflection from the dichroic mirror surface in the dual-view emission pathway, and the background. Each histogram contains two major peaks (thick blue curves) and an intermediate peak (thin blue curves). The first major peak at E_{FRET} =0 corresponds to duplexes with active Cy3 only, and the second major peak at higher E_{FRET} resulting from duplex molecules with an active Cy3-Cy5 pair. The latter was used to calculate the E_{FRET} value for each species. Intermediate peaks are due to the multiple Cy3-Cy5 pairs in vesicles, some of which have photobleached Cy5, yielding average E_{FRET} values that are less than E_{FRET} of the active Cy3-Cy5 pairs.

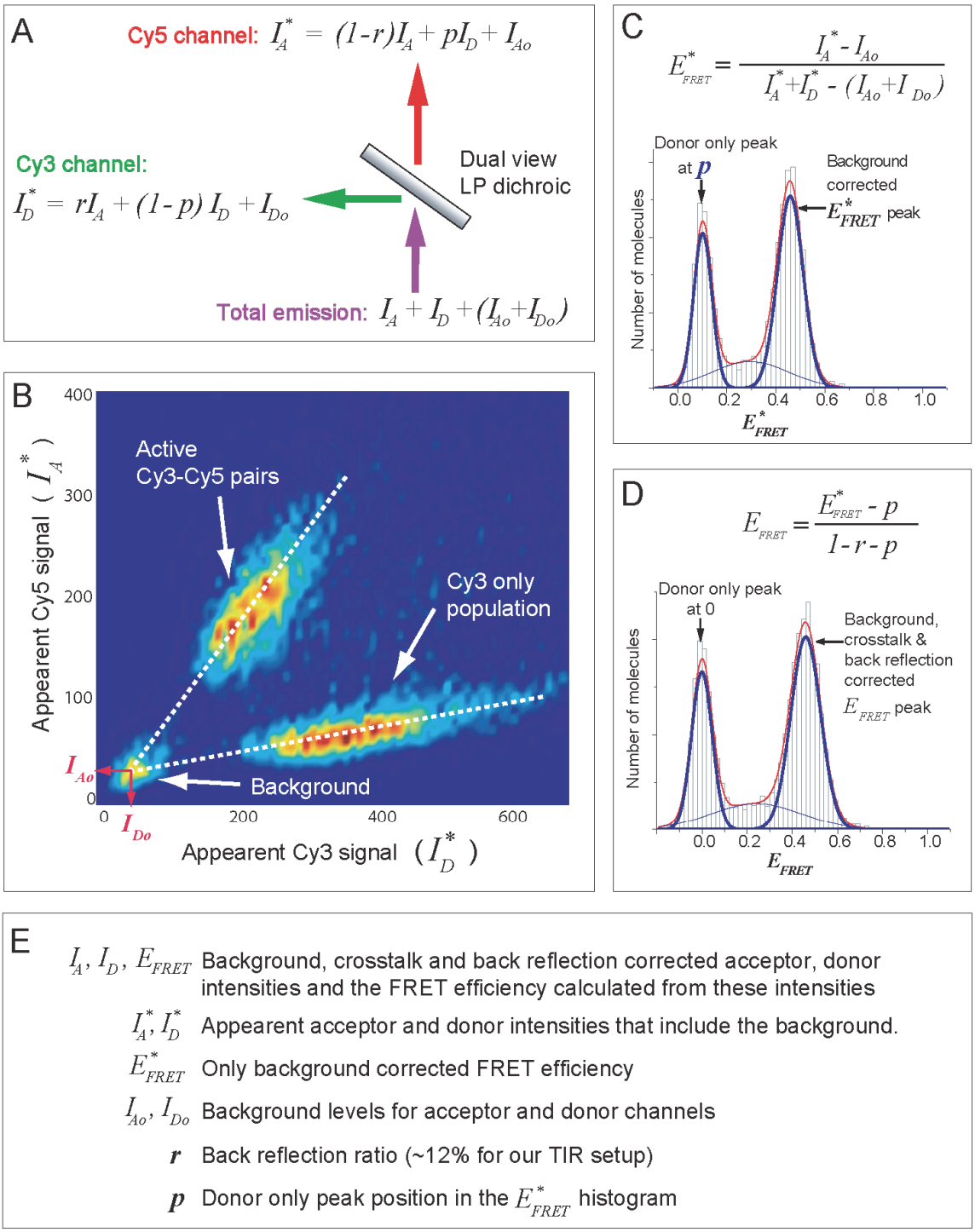


Figure 4.7 Data analysis method for single-molecule experiments. (*A*) How the total emission from the Cy3 and Cy5 molecules is separated to their respective channels at the 650-nm long-pass dichroic mirror. The explanation for all of the terms in the formulas is provided in *E*. (*B*) The background for the donor and acceptor channel is determined by using a dual histogram of apparent Cy3 and Cy5 signal pairs. The particular density map was generated from 8,800 molecules of 15-bp DNA/RNA duplexes. The color represents the density of the molecules, red

being the highest density and the blue being the lowest. Because E_{FRET} is determined by the ratio of the acceptor signal-to-donor signal, a population of certain E_{FRET} pairs should be distributed along a straight line in this plot depending on the total intensity of the pair. Two major populations belong to active Cy3-Cy5 pairs and only Cy3 active duplexes as shown. A small third population appears close to origin and arises from duplexes with inactive Cy3 or Cy5. Because we have a background signal in both channels, this third population is not at the origin, so the center of this population gives the background values in the respective channels as shown with red arrows. The intersection point of the lines that are fitted to the two major populations should also give this background value. Both techniques are used. (C) E_{FRET} histogram after the background correction is shown for the 8,800 15-mer RNA/DNA molecules. The donor-only peak position (crosstalk correction parameter) p is determined at this stage via Gaussian fitting. The intermediate broad peak arises mainly from vesicles with multiple duplexes of active and only Cy3 active pairs. (D) The same histogram after the crosstalk (p) and the back reflection (r=0.12) corrections are performed. The true E_{FRET} value of the Cy3-Cy5 pairs on this duplex is given by the center of the second major peak. (E) The explanation for the symbols used in the formulae used above.

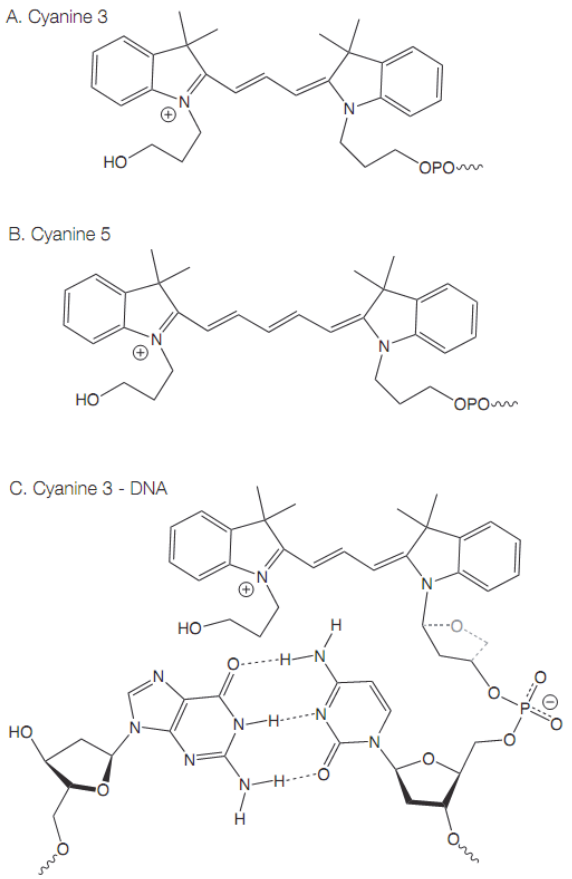


Figure 4.8 The chemical structures of the cyanine fluorophores. (A and B) The structures of Cy3 and Cy5, respectively. (C) The structure of Cy3 attached to a terminal GC base pair of DNA. The broken line shows how the 3-carbon tether attaching Cy3 to the 5'-phosphate can be drawn equivalently to an additional deoxyribose group.

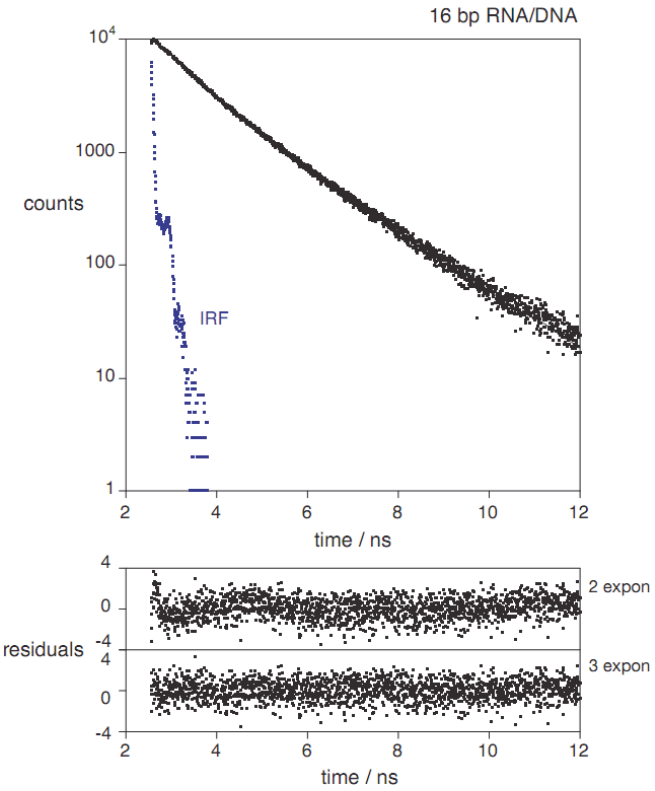


Figure 4.9 Time-resolved fluorescent lifetime analysis of Cy3 attached to double-stranded RNA/DNA hybrid duplex. Fluorescent decay curve for Cy3 attached to a 16-bp RNA/DNA duplex, showing the experimental data and the instrument response function (IRF), and the fit to three exponential functions (line). Plots of weighted residuals for fitting the data to two (X^2=1.113) or three (X^2=0.986) exponential functions are shown below.

4.7 Tables

	Three components				Two components			
	τ	α	f	χ^2	τ	α	f	χ^2
DNA 22 bp								
	0.39	0.019	16.5	0.957	0.61	0.032	42.6	1.238
	1.05	0.026	61.3		1.53	0.017	57.4	
	1.97	0.005	22.2					
DNA 16 bp								
	0.39	0.018	16.9	1.062	0.59	0.029	40.6	1.326
	1.04	0.025	61.6		1.47	0.017	59.4	
	1.91	0.005	21.4					
RNA/DNA 22 bp								
	0.36	0.011	7.0	1.060	0.71	0.026	33.0	1.493
	1.03	0.026	49.1		1.64	0.023	67.0	
	1.84	0.013	43.9					
RNA/DNA 16 bp								
	0.31	0.009	5.3	0.986	0.69	0.025	31.2	1.113
	0.96	0.025	43.7		1.63	0.023	68.8	
	1.77	0.016	51.1					

Table 4.1 Fluorescent lifetime data for 5-Cy3-labeled DNA and RNA/DNA duplexes. The results of fitting to two and three exponential functions are presented. τ (ns), α, and *f* are the fluorescent lifetime, relative amplitude, and steady-state fractional intensity for each component, and χ^2 is the chi-squared statistic.

10bp
Cy3-5’-CCACTCTAGG-3’
Cy5-5’-CCTAGAGTGG-3’
11bp
Cy3-5’-CCACTGCTAGG-3’
Cy5-5’-CCTAGCAGTGG-3’
12bp
Cy3-5’-CCACTGGCTAGG-3’
Cy5-5’-CCTAGCCAGTGG-3’
13bp
Cy3-5’-CCACTGCGCTAGG-3’
Cy5-5’-CCTAGCGCAGTGG-3’
14bp
Cy3-5’-CCACTGCTGCTAGG-3’
Cy5-5’-CCTAGCAGCAGTGG-3’
15bp
Cy3-5’-CCACTGCCTGCTAGG-3’
Cy5-5’-CCTAGCAGGCAGTGG-3’
16bp
Cy3-5’-CCACTGCACTGCTAGG-3’
Cy5-5’-CCTAGCAGTGCAGTGG-3’
17bp
Cy3-5’-CCACTGCACCTGCTAGG-3’
Cy5-5’-CCTAGCAGGTGCAGTGG-3’
18bp
Cy3-5’-CCACTGCACGCTGCTAGG-3’
Cy5-5’-CCTAGCAGCGTGCAGTGG-3’
19bp
Cy3-5’-CCACTGCACCGCTGCTAGG-3’
Cy5-5’-CCTAGCAGCGGTGCAGTGG-3’
20bp

Table 4.2 DNA Sequences Used to Construct the Duplex Series. The RNA/DNA duplexes had the same sequence, where the Cy3-labeled strand was exchanged for RNA, for example, the 10-nt strand was Cy3–5’-CCACUCUAGG-3’. These were hybridized with the same Cy5-labeled DNA strands.

References

1. J. Park *et al.*, PcrA helicase dismantles RecA filaments by reeling in DNA in uniform steps. *Cell* **142**, 544-555 (2010).
2. T. Förster, Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der Physik* **437**, 55-75 (1948).
3. E. J. Ambrose, A surface contact microscope for the study of cell movements. *Nature* **178**, 1194 (1956).
4. D. Axelrod, Cell-substrate contacts illuminated by total internal reflection fluorescence. *J. Cell. Biol.* **89**, 141-145 (1981).
5. J. R. Moffitt *et al.*, Intersubunit coordination in a homomeric ring ATPase. *Nature* **457**, 446-450 (2009).
6. A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, S. Chu, Observation of a single-beam gradient force optical trap for dielectric particles. *Opt Lett* **11**, 288 (1986).
7. M. T. Woodside *et al.*, Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science* **314**, 1001-1004 (2006).
8. B. Onoa *et al.*, Identifying kinetic barriers to mechanical unfolding of the T. thermophila ribozyme. *Science* **299**, 1892-1895 (2003).
9. J. Mameren *et al.*, Dissecting elastic heterogeneity along DNA molecules coated partly with Rad51 using concurrent fluorescence microscopy and optical tweezers. *Biophys. J.* **91**, L78-80 (2006).
10. R. Galletto, I. Amitani, R. J. Baskin, S. C. Kowalczykowski, Direct observation of individual RecA filaments assembling on single DNA molecules. *Nature* **443**, 875-878 (2006).
11. G. V. Shivashankar, M. Feingold, O. Krichevsky, A. Libchaber, RecA polymerization on double-stranded DNA by using single-molecule manipulation: the role of ATP hydrolysis. *Proc Natl Acad Sci U S A* **96**, 7916-7921 (1999).
12. K. Hatch, C. Danilowicz, V. Coljee, M. Prentiss, Measurement of the salt-dependent stabilization of partially open DNA by Escherichia coli SSB protein. *Nucleic Acids Res* **36**, 294-299 (2008).
13. B. M. Ali *et al.*, Compaction of single DNA molecules induced by binding of integration host factor (IHF). *Proc Natl Acad Sci U S A* **98**, 10658-10663 (2001).
14. T. M. Lohman, K. P. Bjornson, Mechanisms of helicase-catalyzed DNA unwinding. *Annu. Rev. Biochem.* **65**, 169-214 (1996).
15. S. S. Patel, K. M. Picha, Structure and function of hexameric helicases. *Annu. Rev. Biochem.* **69**, 651-697 (2000).
16. M. R. Singleton, M. S. Dillingham, D. B. Wigley, Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* **76**, 23-50 (2007).
17. T. M. Lohman, E. J. Tomko, C. G. Wu, Non-hexameric DNA helicases and translocases: mechanisms and regulation. *Nature Rev. Mol. Cell Biol.* **9**, 391-401 (2008).
18. O. Cordin, J. Banroques, N. K. Tanner, P. Linder, The DEAD-box protein family of RNA helicases. *Gene* **367**, 17-37 (2006).
19. E. Delagoutte, P. H. von Hippel, Helicase mechanisms and the coupling of helicases within macromolecular machines. Part II: Integration of helicases into cellular processes. *Q Rev Biophys* **36**, 1-69 (2003).
20. S. W. Matson, D. W. Bean, J. W. George, DNA helicases: enzymes with essential roles in all aspects of DNA metabolism. *Bioessays* **16**, 13-22 (1994).

21. A. Saha, J. Wittmeyer, B. R. Cairns, Chromatin remodelling: the industrial revolution of DNA around histones. *Nat Rev Mol Cell Biol* **7**, 437-447 (2006).
22. X. Veaute *et al.*, UvrD helicase, unlike Rep helicase, dismantles RecA nucleoprotein filaments in Escherichia coli. *EMBO J.* **24**, 180-189 (2005).
23. M. C. Hall, S. W. Matson, Helicase motifs: the engine that powers DNA unwinding. *Mol Microbiol* **34**, 867-877 (1999).
24. R. L. Eoff, K. D. Raney, Helicase-catalysed translocation and strand separation. *Biochem. Soc. Trans.* **33**, 1474-1478 (2005).
25. P. H. von Hippel, E. Delagoutte, A general model for nucleic acid helicases and their "coupling" within macromolecular machines. *Cell* **104**, 177-190 (2001).
26. M. Abdel-Monem, H. Hoffmann-Berling, Enzymic unwinding of DNA. 1. Purification and characterization of a DNA-dependent ATPase from Escherichia coli. *Eur. J. Biochem.* **65**, 431-440 (1976).
27. C. A. Brennan, A. J. Dombroski, T. Platt, Transcription termination factor rho is an RNA-DNA helicase. *Cell* **48**, 945-952 (1987).
28. A. Lahaye, S. Leterme, F. Foury, PIF1 DNA helicase from Saccharomyces cerevisiae. Biochemical characterization of the enzyme. *J. Biol. Chem.* **268**, 26155-26161 (1993).
29. N. A. Ellis, DNA helicases in inherited human disorders. *Curr. Opin. Genet. Dev.* **7**, 354-363 (1997).
30. N. A. Ellis *et al.*, The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell* **83**, 655-666 (1995).
31. I. D. Hickson, RecQ helicases: caretakers of the genome. *Nat Rev Cancer* **3**, 169-178 (2003).
32. M. D. Gray *et al.*, The Werner syndrome protein is a DNA helicase. *Nat Genet* **17**, 100-103 (1997).
33. J. C. Shen *et al.*, Werner syndrome protein. I. DNA helicase and dna exonuclease reside on the same polypeptide. *J. Biol. Chem.* **273**, 34139-34144 (1998).
34. C. E. Yu *et al.*, Positional cloning of the Werner's syndrome gene. *Science* **272**, 258-262 (1996).
35. K. Geider, H. Hoffmann-Berling, Proteins controlling the helical structure of DNA. *Annu. Rev. Biochem.* **50**, 233-260 (1981).
36. T. L. Hill, T. Tsuchiya, Theoretical aspects of translocation on DNA: adenosine triphosphatases and treadmilling binding proteins. *Proc Natl Acad Sci U S A* **78**, 4796-4800 (1981).
37. T. M. Lohman, Escherichia coli DNA helicases: mechanisms of DNA unwinding. *Mol Microbiol* **6**, 5-14 (1992).
38. T. M. Lohman, Helicase-catalyzed DNA unwinding. *J. Biol. Chem.* **268**, 2269-2272 (1993).
39. I. Wong, T. M. Lohman, A double-filter method for nitrocellulose-filter binding: application to protein-nucleic acid interactions. *Proc Natl Acad Sci U S A* **90**, 5428-5432 (1993).
40. W. Cheng, J. Hsieh, K. M. Brendza, T. M. Lohman, E. coli Rep oligomers are required to initiate DNA unwinding in vitro. *J. Mol. Biol.* **310**, 327-350 (2001).
41. T. Ha *et al.*, Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase. *Nature* **419**, 638-641 (2002).
42. A. E. Gorbalenya, E. V. Koonin, Helicases - Amino-Acid-Sequence Comparisons and Structure-Function-Relationships. *Curr. Opin. Struct. Biol.* **3**, 419-429 (1993).
43. H. Bialkowskahobrzanska, D. T. Denhardt, The Rep Mutation .7. Cloning and Analysis of the Functional Rep Gene of Escherichia-Coli-K-12. *Gene* **28**, 93-102 (1984).

44. H. Bialkowskahobrzanska, C. A. Gilchrist, D. T. Denhardt, Rep Mutation .8. Escherichia-Coli-Rep Gene - Identification of the Promoter and N-Terminus of the Rep Protein. *J. Bacteriol.* **164**, 1004-1010 (1985).
45. J. Colasanti, D. T. Denhardt, The Escherichia-Coli Rep Mutation .10. Consequences of Increased and Decreased Rep Protein-Levels. *Molecular & General Genetics* **209**, 382-390 (1987).
46. A. Kornberg, J. F. Scott, L. L. Bertsch, Atp Utilization by Rep Protein in Catalytic Separation of DNA Strands at a Replicating Fork. *J. Biol. Chem.* **253**, 3298-3304 (1978).
47. H. E. D. Lane, D. T. Denhardt, Rep Mutation .4. Slower Movement of Replication Forks in Escherichia-Coli Rep Strains. *J. Mol. Biol.* **97**, 99-& (1975).
48. H. E. D. Lane, D. T. Denhardt, Rep Mutation .3. Altered Structure of Replicating Escherichia-Coli Chromosome. *J. Bacteriol.* **120**, 805-814 (1974).
49. P. McGlynn, Helicases that underpin replication of protein-bound DNA in Escherichia coli. *Biochem. Soc. Trans.* **39**, 606-610 (2011).
50. J. Atkinson *et al.*, Localization of an accessory helicase at the replisome is critical in sustaining efficient genome duplication. *Nucleic Acids Res* **39**, 949-957 (2011).
51. C. P. Guy *et al.*, Rep provides a second motor at the replisome to promote duplication of protein-bound DNA. *Mol Cell* **36**, 654-666 (2009).
52. J. E. Yanceywrona, S. W. Matson, Bound Lac Repressor Protein Differentially Inhibits the Unwinding Reactions Catalyzed by DNA Helicases. *Nucleic Acids Res.* **20**, 6713-6721 (1992).
53. S. W. Matson, D. W. Bean, J. W. George, DNA Helicases - Enzymes with Essential Roles in All Aspects of DNA Metabolism. *Bioessays* **16**, 13-22 (1994).
54. B. Michel, M. Uzest, Lethality of Double Mutations Rep-Recb and Rep-Recc in Escherichia-Coli. *J. Cell. Biochem.*, 114-114 (1995).
55. K. M. Brendza *et al.*, Autoinhibition of Escherichia coli Rep monomer helicase activity by its 2B subdomain. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10076-10081 (2005).
56. M. S. Dillingham, D. B. Wigley, M. R. Webb, Direct measurement of single-stranded DNA translocation by PcrA helicase using the fluorescent base analogue 2-aminopurine. *Biochemistry* **41**, 643-651 (2002).
57. C. J. Fischer, N. K. Maluf, T. M. Lohman, Mechanism of ATP-dependent translocation of E.coli UvrD monomers along single-stranded DNA. *J. Mol. Biol.* **344**, 1287-1309 (2004).
58. S. Myong, I. Rasnik, C. Joo, T. M. Lohman, T. Ha, Repetitive shuttling of a motor protein on DNA. *Nature* **437**, 1321-1325 (2005).
59. A. Niedziela-Majka, M. A. Chesnik, E. J. Tomko, T. M. Lohman, Bacillus stearothermophilus PcrA monomer is a single-stranded DNA translocase but not a processive helicase in vitro. *J. Biol. Chem.* **282**, 27076-27085 (2007).
60. E. J. Tomko, C. J. Fischer, A. Niedziela-Majka, T. M. Lohman, A nonuniform stepping mechanism for E. coli UvrD monomer translocation along single-stranded DNA. *Mol Cell* **26**, 335-347 (2007).
61. S. Korolev, J. Hsieh, G. H. Gauss, T. M. Lohman, G. Waksman, Major domain swiveling revealed by the crystal structures of complexes of E. coli Rep helicase bound to single-stranded DNA and ADP. *Cell* **90**, 635-647 (1997).
62. H. S. Subramanya, L. E. Bird, J. A. Brannigan, D. B. Wigley, Crystal structure of a DExx box DNA helicase. *Nature* **384**, 379-383 (1996).
63. S. S. Velankar, P. Soultanas, M. S. Dillingham, H. S. Subramanya, D. B. Wigley, Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* **97**, 75-84 (1999).

64. K. J. Marians, Mechanisms of replication fork restart in Escherichia coli. *Philos Trans R Soc Lond B Biol Sci* **359**, 71-77 (2004).
65. S. J. Sandler, Multiple genetic pathways for restarting DNA replication forks in Escherichia coli K-12. *Genetics* **155**, 487-497 (2000).
66. R. C. Heller, K. J. Marians, Unwinding of the nascent lagging strand by Rep and PriA enables the direct restart of stalled replication forks. *J. Biol. Chem.* **280**, 34143-34151 (2005).
67. N. K. Maluf, C. J. Fischer, T. M. Lohman, A Dimer of Escherichia coli UvrD is the active form of the helicase in vitro. *J. Mol. Biol.* **325**, 913-935 (2003).
68. J. Y. Lee, W. Yang, UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. *Cell* **127**, 1349-1360 (2006).
69. H. F. Jia *et al.*, Rotations of the 2B Sub-domain of E-coli UvrD Helicase/Translocase Coupled to Nucleotide and DNA Binding. *J. Mol. Biol.* **411**, 633-648 (2011).
70. W. Cheng *et al.*, The 2B domain of the Escherichia coli Rep protein is not required for DNA helicase activity. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16006-16011 (2002).
71. K. M. Brendza *et al.*, The 2B domain of E-coli rep protein is not required for helicase activity. *Biophys. J.* **84**, 365a-365a (2003).
72. I. Rasnik, S. Myong, W. Cheng, T. M. Lohman, T. Ha, DNA-binding orientation and domain conformation of the E. coli rep helicase monomer bound to a partial duplex junction: single-molecule studies of fluorescently labeled enzymes. *J. Mol. Biol.* **336**, 395-408 (2004).
73. E. J. Tomko *et al.*, 5'-Single-stranded/duplex DNA junctions are loading sites for E. coli UvrD translocase. *EMBO J.* **29**, 3826-3839 (2010).
74. H. Hwang, H. Kim, S. Myong, Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity. *Proc Natl Acad Sci U S A* **108**, 7414-7418 (2011).
75. B. Okumus, S. Arslan, S. M. Fengler, S. Myong, T. Ha, Single molecule nanocontainers made porous using a bacterial toxin. *J. Am. Chem. Soc.* **131**, 14844-14849 (2009).
76. K. Sasaki *et al.*, Structural basis of the 3'-end recognition of a leading strand in stalled replication forks by PriA. *EMBO J.* **26**, 2584-2593 (2007).
77. I. R. Lehman, A. L. Nussbaum, The Deoxyribonucleases of Escherichia coli : V. ON THE SPECIFICITY OF EXONUCLEASE I (PHOSPHODIESTERASE). *J. Biol. Chem.* **239**, 2628-2636 (1964).
78. T. Ha *et al.*, Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase. *Nature* **419**, 638-641 (2002).
79. P. R. Selvin, T. Ha, *Single-molecule techniques: a laboratory manual* (Cold Spring Harbor Laboratory Press, New York, ed. 1st).
80. Y. Harada, K. Sakurada, T. Aoki, D. D. Thomas, T. Yanagida, Mechanochemical coupling in actomyosin energy transduction studied by in vitro movement assay. *J. Mol. Biol.* **216**, 49-68 (1990).
81. T. Yanagida, M. Nakase, K. Nishiyama, F. Oosawa, Direct observation of motion of single F-actin filaments in the presence of myosin. *Nature* **307**, 58-60 (1984).
82. I. Rasnik, S. A. McKinney, T. Ha, Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat. Methods* **3**, 891-893 (2006).
83. R. Roy, S. Hohng, T. Ha, A practical guide to single-molecule FRET. *Nat. Methods* **5**, 507-516 (2008).
84. C. Joo, T. Ha, in *Cold Spring Harb. Protoc.* (2012), vol. 2012.

85. M. Sorokina, H. R. Koh, S. S. Patel, T. Ha, Fluorescent lifetime trajectories of a single fluorophore reveal reaction intermediates during transcription initiation. *J. Am. Chem. Soc.* **131**, 9630-9631 (2009).
86. C. Joo, H. Balci, Y. Ishitsuka, C. Buranachai, T. Ha, Advances in single-molecule fluorescence methods for molecular biology. *Annu. Rev. Biochem.* **77**, 51-76 (2008).
87. C. Joo *et al.*, Real-time observation of RecA filament dynamics with single monomer resolution. *Cell* **126**, 515-527 (2006).
88. X. S. Xie, P. J. Choi, G. W. Li, N. K. Lee, G. Lia, Single-molecule approach to molecular biology in living bacterial cells. *Annual review of biophysics* **37**, 417-444 (2008).
89. W. E. Moerner, New directions in single-molecule imaging and analysis. *Proceedings of the National Academy of Sciences* **104**, 12596-12602 (2007).
90. A. Yildiz *et al.*, Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science* **300**, 2061-2065 (2003).
91. A. Yildiz, M. Tomishige, R. Vale, P. Selvin, Kinesin walks hand-over-hand. *Science (New York, N.Y.)* **303**, 676-684 (2004).
92. Lommerse, *J. Cell Sci.* **118**, 1799 (2005).
93. J. Y. Lee, B. Okumus, D. S. Kim, T. Ha, Extreme conformational diversity in human telomeric DNA. *Proc Natl Acad Sci U S A* **102**, 18938-18943 (2005).
94. E. V. Amirgoulova *et al.*, Biofunctionalized polymer surfaces exhibiting minimal interaction towards immobilized proteins. *Chemphyschem* **5**, 552-555 (2004).
95. R. M. Dickson, D. J. Norris, Y. L. Tzeng, W. E. Moerner, Three-dimensional imaging of single molecules solvated in pores of poly(acrylamide) gels. *Science* **274**, 966-969 (1996).
96. A. E. Cohen, W. E. Moerner, Suppressing Brownian motion of individual biomolecules in solution. *Proc Natl Acad Sci U S A* **103**, 4362-4365 (2006).
97. E. Boukobza, A. Sonnenfeld, G. Haran, Immobilization in surface-tethered lipid vesicles as a new tool for single biomolecule spectroscopy. *Journal of Physical Chemistry B* **105**, 12165-12170 (2001).
98. B. Okumus, T. J. Wilson, D. M. Lilley, T. Ha, Vesicle encapsulation studies reveal that single molecule ribozyme heterogeneities are intrinsic. *Biophys. J.* **87**, 2798-2806 (2004).
99. J. J. Benítez *et al.*, Probing Transient Copper Chaperone−Wilson Disease Protein Interactions at the Single-Molecule Level with Nanovesicle Trapping. *J. Am. Chem. Soc.* **130**, 2446-2447 (2008).
100. E. Rhoades, E. Gussakovsky, G. Haran, Watching proteins fold one molecule at a time. *Proc Natl Acad Sci U S A* **100**, 3197-3202 (2003).
101. I. Cisse, B. Okumus, C. Joo, T. Ha, Fueling protein DNA interactions inside porous nanocontainers. *Proc Natl Acad Sci U S A* **104**, 12646-12650 (2007).
102. M. Mickler, M. Hessling, C. Ratzke, J. Buchner, T. Hugel, The large conformational changes of Hsp90 are only weakly coupled to ATP hydrolysis. *Nat. Struct. Mol. Biol.* **16**, 281-287 (2009).
103. D. T. Chiu *et al.*, Chemical transformations in individual ultrasmall biomimetic containers. *Science* **283**, 1892-1895 (1999).
104. D. S. Talaga *et al.*, Dynamics and folding of single two-stranded coiled-coil peptides studied by fluorescent energy transfer confocal microscopy. *Proc Natl Acad Sci U S A* **97**, 13021-13026 (2000).
105. A. Iqbal *et al.*, Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids. *Proc Natl Acad Sci U S A* **105**, 11176-11181 (2008).

106. V. Noireaux, A. Libchaber, A vesicle bioreactor as a step toward an artificial cell assembly. *Proc Natl Acad Sci U S A* **101**, 17669-17674 (2004).
107. M. Winterhalter *et al.*, Controlling membrane permeability with bacterial porins: application to encapsulated enzymes. *Talanta* **55**, 965-971 (2001).
108. W. Meier, C. Nardin, M. Winterhalter, Reconstitution of Channel Proteins in (Polymerized) ABA Triblock Copolymer Membranes. *Angewandte Chemie International Edition* **39**, 4599-4602 (2000).
109. P. Y. Bolinger, D. Stamou, H. Vogel, Integrated nanoreactor systems: triggering the release and mixing of compounds inside single vesicles. *Journal of American Chemical Society* **126**, 8594-8595 (2004).
110. S. Bhakdi *et al.*, Staphylococcal alpha-toxin, streptolysin-O, and Escherichia coli hemolysin: prototypes of pore-forming bacterial cytolysins. *Arch. Microbiol.* **165**, 73-79 (1996).
111. C. Y. Chang, B. Niblack, B. Walker, H. Bayley, A Photogenerated Pore-Forming Protein. *Chem. Biol.* **2**, 391-400 (1995).
112. T. Ha *et al.*, Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc Natl Acad Sci U S A* **93**, 6264-6268 (1996).
113. E. Tan *et al.*, A four-way junction accelerates hairpin ribozyme folding via a discrete intermediate. *Proc Natl Acad Sci U S A* **100**, 9308-9313 (2003).
114. S. Hohng *et al.*, Conformational Flexibility of Four-way Junctions in RNA. *J. Mol. Biol.* **336**, 69-79 (2004).
115. H. Bayley, Building doors into cells. *Sci. Am.* **277**, 62-67 (1997).
116. M. J. Russo, H. Bayley, M. Toner, Reversible permeabilization of plasma membranes with an engineered switchable pore. *Nat Biotech* **15**, 278-282 (1997).
117. Y. Ishitsuka, B. Okumus, S. Arslan, K. H. Chen, T. Ha, Temperature-independent porous nanocontainers for single-molecule fluorescence studies. *Anal. Chem.* **82**, 9694-9701 (2010).
118. V. Vamvakaki, N. A. Chaniotakis, Immobilization of enzymes into nanocavities for the improvement of biosensor stability. *Biosensors Bioelectron.* **22**, 2650-2655 (2007).
119. K. Saikrishnan, B. Powell, N. J. Cook, M. R. Webb, D. B. Wigley, Mechanistic basis of 5'-3' translocation in SF1B helicases. *Cell* **137**, 849-859 (2009).
120. X. He *et al.*, The T4 phage SF1B helicase Dda is structurally optimized to perform DNA strand separation. *Structure* **20**, 1189-1200 (2012).
121. J. Yu, T. Ha, K. Schulten, Structure-based model of the stepping motor of PcrA helicase. *Biophys. J.* **91**, 2097-2114 (2006).
122. J. Yu, T. Ha, K. Schulten, How directional translocation is regulated in a DNA helicase motor. *Biophys. J.* **93**, 3783-3797 (2007).
123. M. Tomishige, R. D. Vale, Controlling kinesin by reversible disulfide cross-linking. Identifying the motility-producing conformational change. *J. Cell. Biol.* **151**, 1081-1092 (2000).
124. A. M. Taylor, Q. Zhu, J. R. Casey, Cysteine-directed cross-linking localizes regions of the human erythrocyte anion-exchange protein (AE1) relative to the dimeric interface. *The Biochemical journal* **359**, 661-668 (2001).
125. D. M. Veine, K. Ohnishi, C. H. Williams, Jr., Thioredoxin reductase from Escherichia coli: evidence of restriction to a single conformation upon formation of a crosslink between engineered cysteines. *Protein Sci.* **7**, 369-375 (1998).

126. A. Watty, C. Methfessel, F. Hucho, Fixation of allosteric states of the nicotinic acetylcholine receptor by chemical cross-linking. *Proc Natl Acad Sci U S A* **94**, 8202-8207 (1997).
127. J. G. Yodh, M. Schlierf, T. Ha, Insight into helicase mechanism and function revealed through single-molecule approaches. *Q. Rev. Biophys.* **43**, 185-217 (2010).
128. H. Balci, S. Arslan, S. Myong, Timothy M. Lohman, T. Ha, Single-Molecule Nanopositioning: Structural Transitions of a Helicase-DNA Complex during ATP Hydrolysis. *Biophys. J.* **101**, 976-984 (2011).
129. R. N. Lindquist, J. L. Lynn, Jr., G. E. Lienhard, Possible transition-state analogs for ribonuclease. The complexes of uridine with oxovanadium(IV) ion and vanadium(V) ion. *Journal of American Chemical Society* **95**, 8762-8768 (1973).
130. F. H. Westheimer, Why nature chose phosphates. *Science* **235**, 1173-1178 (1987).
131. U. B. Choi *et al.*, Single-molecule FRET-derived model of the synaptotagmin 1-SNARE fusion complex. *Nat Struct Mol Biol* **17**, 318-324 (2010).
132. A. Muschielok *et al.*, A nano-positioning system for macromolecular structural analysis. *Nat Methods*, (2008).
133. J. Andrecka *et al.*, Single-molecule tracking of mRNA exiting from RNA polymerase II. *Proc Natl Acad Sci U S A* **105**, 135-140 (2008).
134. X. Sun *et al.*, Architecture of the 99 bp DNA-six-protein regulatory complex of the lambda att site. *Mol Cell* **24**, 569-580 (2006).
135. B. Sun *et al.*, Impediment of E. coli UvrD by DNA-destabilizing force reveals a strained-inchworm mechanism of DNA unwinding. *EMBO J.* **27**, 3279-3287 (2008).
136. M. S. Dillingham, Superfamily I helicases as modular components of DNA-processing machines. *Biochem. Soc. Trans.* **39**, 413-423 (2011).
137. M. Yamaguchi, V. Dao, P. Modrich, MutS and MutL activate DNA helicase II in a mismatch-dependent manner. *J. Biol. Chem.* **273**, 9197-9201 (1998).
138. L. E. Mechanic, B. A. Frankel, S. W. Matson, Escherichia coli MutL loads DNA helicase II onto DNA. *J. Biol. Chem.* **275**, 38337-38346 (2000).
139. P. Soultanas *et al.*, Plasmid replication initiator protein RepD increases the processivity of PcrA DNA helicase. *Nucleic Acids Res.* **27**, 1421-1428 (1999).
140. A. F. Slatter, C. D. Thomas, M. R. Webb, PcrA helicase tightly couples ATP hydrolysis to unwinding double-stranded DNA, modulated by the initiator protein for plasmid replication, RepD. *Biochemistry* **48**, 6326-6334 (2009).
141. P. Soultanas, M. S. Dillingham, D. B. Wigley, Escherichia coli ribosomal protein L3 stimulates the helicase activity of the Bacillus stearothermophilus PcrA helicase. *Nucleic Acids Res.* **26**, 2374-2379 (1998).
142. J. E. Yancey, S. W. Matson, The DNA unwinding reaction catalyzed by Rep protein is facilitated by an RHSP-DNA interaction. *Nucleic Acids Res.* **19**, 3943-3951 (1991).
143. S. Barranco-Medina, R. Galletto, DNA binding induces dimerization of Saccharomyces cerevisiae Pif1. *Biochemistry* **49**, 8445-8454 (2010).
144. H. Jia *et al.*, Rotations of the 2B sub-domain of E. coli UvrD helicase/translocase coupled to nucleotide and DNA binding. *J. Mol. Biol.* **411**, 633-648 (2011).
145. B. Choi, G. Zocchi, Y. Wu, S. Chan, L. J. Perry, Allosteric control through mechanical tension. *Phys. Rev. Lett.* **95**, 078102 (2005).
146. M. Volgraf *et al.*, Allosteric control of an ionotropic glutamate receptor with an optical switch. *Nat. Chem. Biol.* **2**, 47-52 (2006).
147. B. Schierling *et al.*, Controlling the enzymatic activity of a restriction enzyme by light. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 1361-1366 (2010).

148. B. X. Huang, H. Y. Kim, Interdomain conformational changes in Akt activation revealed by chemical cross-linking and tandem mass spectrometry. *Mol. Cell. Proteomics* **5**, 1045-1053 (2006).
149. M. Banghart, K. Borges, E. Isacoff, D. Trauner, R. H. Kramer, Light-activated ion channels for remote control of neuronal firing. *Nat. Neurosci.* **7**, 1381-1386 (2004).
150. D. L. Fortin *et al.*, Optogenetic photochemical control of designer K+ channels in mammalian neurons. *J. Neurophysiol.* **106**, 488-496 (2011).
151. F. Zhang, K. A. Timm, K. M. Arndt, G. A. Woolley, Photocontrol of coiled-coil proteins in living cells. *Angewandte Chemie* **49**, 3943-3946 (2010).
152. P. Gorostiza, E. Y. Isacoff, Optical switches for remote and noninvasive control of cell signaling. *Science* **322**, 395-399 (2008).
153. N. Caporale *et al.*, LiGluR restores visual responses in rodent models of inherited blindness. *Molecular therapy : the journal of the American Society of Gene Therapy* **19**, 1212-1219 (2011).
154. G. Lee, M. A. Bratkowski, F. Ding, A. Ke, T. Ha, Elastic coupling Between RNA Degradation and Unwinding by an Exoribonuclease. *Science* **336**, 1726-1729 (2012).
155. M. C. Murphy, I. Rasnik, W. Cheng, T. M. Lohman, T. Ha, Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. *Biophys. J.* **86**, 2530-2537 (2004).
156. J. A. Ali, T. M. Lohman, Kinetic measurement of the step size of DNA unwinding by Escherichia coli UvrD helicase. *Science* **275**, 377-380 (1997).
157. L. R. Brewer, P. R. Bianco, Laminar flow cells for single-molecule studies of DNA-protein interactions. *Nat. Methods* **5**, 517-525 (2008).
158. T. T. Perkins, H. W. Li, R. V. Dalal, J. Gelles, S. M. Block, Forward and reverse motion of single RecBCD molecules on DNA. *Biophys. J.* **86**, 1640-1648 (2004).
159. W. Zhang *et al.*, Directional loading and stimulation of PcrA helicase by the replication initiator protein RepD. *J. Mol. Biol.* **371**, 336-348 (2007).
160. C. Machon *et al.*, RepD-mediated recruitment of PcrA helicase at the Staphylococcus aureus pC221 plasmid replication origin, oriD. *Nucleic Acids Res.* **38**, 1874-1888 (2010).
161. M. J. Comstock, H. Jia, T. M. Lohman, T. Ha, Y. R. Chemla, Direct observation of structure-function relationship in a nucleic acid processing enzyme. *Concurrent submission with this manuscript*, (2014).
162. M. N. Dessinges, T. Lionnet, X. G. Xi, D. Bensimon, V. Croquette, Single-molecule assay reveals strand switching and enhanced processivity of UvrD. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 6439-6444 (2004).
163. J. G. Yodh, B. C. Stevens, R. Kanagaraj, P. Janscak, T. Ha, BLM helicase measures DNA unwound before switching strands and hRPA promotes unwinding reinitiation. *EMBO J.* **28**, 405-416 (2009).
164. D. Branton *et al.*, The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* **26**, 1146-1153 (2008).
165. A. H. Laszlo *et al.*, Decoding long nanopore sequencing reads of natural DNA. *Nat. Biotechnol.*, (2014).
166. M. Vincent, Y. Xu, H. Kong, Helicase-dependent isothermal DNA amplification. *EMBO Rep.* **5**, 795-800 (2004).
167. I. L. Urbatsch *et al.*, Cysteines 431 and 1074 are responsible for inhibitory disulfide cross-linking between the two nucleotide-binding sites in human P-glycoprotein. *J. Biol. Chem.* **276**, 26980-26987 (2001).

168. C. J. Clopper, E. S. Pearson, The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26**, 404-413 (1934).
169. J. Gal, R. Schnell, S. Szekeres, M. Kalman, Directional cloning of native PCR products with preformed sticky ends (autosticky PCR). *Mol. Gen. Genet.* **260**, 569-573 (1999).
170. Z. Qi, R. A. Pugh, M. Spies, Y. R. Chemla, Sequence-dependent base pair stepping dynamics in XPD helicase unwinding. *eLife* **2**, e00334 (2013), doi:10.7554/eLife.00334.
171. M. P. Landry, P. M. McCall, Z. Qi, Y. R. Chemla, Characterization of photoactivated singlet oxygen damage in single-molecule optical trap experiments. *Biophys. J.* **97**, 2128-2136 (2009).
172. C. Bustamante, Y. R. Chemla, J. R. Moffitt, *High-resolution dual-trap optical tweezers with differential detection*. Single-molecule techniques: a laboratory manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2008).
173. K. Berg-Sørensen, H. Flyvbjerg, Power spectrum analysis for optical tweezers. *Rev. Sci. Instrum.* **75**, 594-612 (2004).
174. G. Zhang, E. Deng, L. Baugh, S. R. Kushner, Identification and characterization of Escherichia coli DNA helicase II mutants that exhibit increased unwinding efficiency. *J. Bacteriol.* **180**, 377-387 (1998).
175. L. Stryer, R. P. Haugland, Energy transfer: a spectroscopic ruler. *Proc Natl Acad Sci U S A* **58**, 719-726 (1967).
176. D. M. J. Lilley, T. J. Wilson, Fluorescence resonance energy transfer as a structural tool for nucleic acids. *Curr. Opin. Chem. Biol.* **4**, 507-517 (2000).
177. R. M. Clegg, FRET tells us about proximities, distances, orientations and dynamic properties. *J. Biotechnol.* **82**, 177-179 (2002).
178. F. D. Lewis, L. G. Zhang, X. B. Zuo, Orientation control of fluorescence resonance energy transfer using DNA as a helical scaffold. *J. Am. Chem. Soc.* **127**, 10002-10003 (2005).
179. D. G. Norman, R. J. Grainger, D. Uhrin, D. M. J. Lilley, Location of cyanine-3 on double-stranded DNA: Importance for fluorescence resonance energy transfer studies. *Biochemistry* **39**, 6317-6324 (2000).
180. A. Iqbal, L. Wang, K. C. Thompson, D. M. Lilley, D. G. Norman, The structure of cyanine 5 terminally attached to double-stranded DNA: implications for FRET studies. *Biochemistry* **47**, 7857-7862 (2008).
181. R. M. Clegg, FLUORESCENCE RESONANCE ENERGY-TRANSFER AND NUCLEIC-ACIDS. *Methods Enzymol.* **211**, 353-388 (1992).
182. R. M. Clegg, A. I. H. Murchie, A. Zechel, D. M. J. Lilley, OBSERVING THE HELICAL GEOMETRY OF DOUBLE-STRANDED DNA IN SOLUTION BY FLUORESCENCE RESONANCE ENERGY-TRANSFER. *Proc Natl Acad Sci U S A* **90**, 2994-2998 (1993).
183. M. E. Sanborn, B. K. Connolly, K. Gurunathan, M. Levitus, Fluorescence properties and photophysics of the sulfoindocyanine Cy3 linked covalently to DNA. *Journal of Physical Chemistry B* **111**, 11064-11074 (2007).
184. A. K. Chibisov *et al.*, PHOTORELAXATION PROCESSES IN COVALENTLY-LINKED INDOCARBOCYANINE AND THIACARBOCYANINE DYES. *Journal of Physical Chemistry* **99**, 886-893 (1995).
185. N. C. Horton, B. C. Finzel, The structure of an RNA/DNA hybrid: A substrate of the ribonuclease activity of HIV-1 reverse transcriptase. *J. Mol. Biol.* **264**, 521-533 (1996).
186. Y. Xiong, M. Sundaralingam, Crystal structure of a DNA center dot RNA hybrid duplex with a polypurine RNA r(gaagaagag) and a complementary polypyrimidine DNA d(CTCTTCTTC). *Nucleic Acids Res* **28**, 2171-2176 (2000).

187. M. L. Kopka, L. Lavelle, G. W. Han, H. L. Ng, R. E. Dickerson, An unusual sugar conformation in the structure of an RNA/DNA decamer of the polypurine tract may affect recognition by RNase H. *J. Mol. Biol.* **334**, 653-665 (2003).
188. S. L. Beaucage, M. H. Caruthers, DEOXYNUCLEOSIDE PHOSPHORAMIDITES - A NEW CLASS OF KEY INTERMEDIATES FOR DEOXYPOLYNUCLEOTIDE SYNTHESIS. *Tetrahedron Lett.* **22**, 1859-1862 (1981).
189. T. J. Wilson, Z. Y. Zhao, K. Maxwell, L. Kontogiannis, D. M. J. Lilley, Importance of specific nucleotides in the folding of the natural form of the hairpin ribozyme. *Biochemistry* **40**, 2291-2302 (2001).
190. T. Ha *et al.*, Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proc Natl Acad Sci U S A* **96**, 893-898 (1999).
191. R. K. Neely, S. W. Magennis, D. T. F. Dryden, A. C. Jones, Evidence of tautomerism in 2-aminopurine from fluorescence lifetime measurements. *Journal of Physical Chemistry B* **108**, 17606-17610 (2004).

APPENDIX: List of publications

Publications* that I co-authored are listed below at the request of the doctoral committee chair.

1. “Activated GTPase movement on an RNA scaffold drives co-translational protein targeting”, K. Shen, S. Arslan, D. Akopian, T. Ha, S. O. Shan. **Nature** 492: 271-275 (2012)
2. “Single-Molecule Nanopositioning: Structural Transitions of a Helicase-DNA Complex during ATP Hydrolysis”, H. Balci, S. Arslan, S. Myong, T. M. Lohman, T. Ha, **Biophys. J.** 101: 976-984 (2011)
3. “Temperature-independent porous nanocontainers for single-molecule fluorescence studies”, Y. Ishitsuka, B. Okumus, S. Arslan, K. H. Chen, T. Ha, **Anal. Chem.** 82: 9694-9701 (2010)
4. “Single molecule nanocontainers made porous using a bacterial toxin”, B. Okumus, S. Arslan, S. M. Fengler, S. Myong, T. Ha, **J. Am. Chem. Soc.** 131: 14844-14849 (2009)
5. "Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids", A. Iqbal†, S. Arslan†, B. Okumus†, T. J. Wilson, G. Giraud, D. G. Noman, T. Ha, D. M. J. Lilley, **Proc. Natl. Acad. Sci. U.S.A.** 105: 11176-11181 (2008)

* Current as of October 18, 2014. Submitted and in preparation manuscripts, and pending patent application(s) are not included.

† Equally contributing authors.