Single Molecule FRET Studies of Reverse Transcription and Chromatin Remodeling

Single molecule FRET studies of reverse transcription and chromatin remodeling

A dissertation presented

by

Bryan T. Harada

to

The Committee on Higher Degrees in Biophysics

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of

Biophysics

Harvard University
Cambridge, Massachusetts

August 2015

Single molecule FRET studies of reverse transcription and chromatin remodeling

Abstract

The measurement of Förster resonance energy transfer (FRET) at the single-molecule level provides a powerful method for monitoring the structural dynamics of biomolecular systems in real time. These single-molecule FRET (smFRET) assays enable the characterization of the transient intermediates that form during enzymatic processes, providing information about the mechanism and regulation of the enzymes involved. In this dissertation, I develop and use smFRET assays to study two processes driven by motor proteins—reverse transcription and chromatin remodeling—and reveal novel features of their mechanism and regulation.

Reverse transcription of the human immunodeficiency virus genome initiates from a cellular tRNA primer that is bound to a specific sequence on the viral RNA (vRNA). During initiation, reverse transcriptase (RT) exhibits a slow mode of synthesis characterized by pauses at specific locations, and RT transitions to a faster mode of synthesis after the extension of the tRNA primer by six nucleotides. By using smFRET to examine how RT interacts with the tRNA-vRNA substrate, we found that RT binds to its substrate in either an active or inactive orientation and samples the two orientations during a single binding event. The equilibrium between these two orientations is a major factor influencing the activity and pausing of the enzyme, and a specific RNA secondary structure in the vRNA substrate modulates the binding mode of RT, determining the locations of the pauses and the transition to the faster mode of synthesis. These results provide a mechanistic explanation for the changes in RT activity observed during initiation and show how the dynamics of a ribonucleoprotein complex can regulate enzymatic activity.

ISWI family chromatin remodelers are another family of motor enzymes regulated by nucleic acid structures. These enzymes are involved in creating evenly spaced nucleosome arrays, and this nucleosome spacing activity arises from the regulation of the enzymes' catalytic activity by the amount of linker DNA present on the nucleosome. We use smFRET and other biochemical assays to monitor intermediates of the remodeling reaction and examine various remodeler mutants in order to elucidate the mechanism of this regulation. These experiments led to the discovery of an allosteric mechanism by which one subunit of the ISWI remodeling complex communicates the presence of linker DNA to the the catalytic subunit by modulating the availability of the histone H4 tail. These results provide a mechanistic explanation for the nucleosome spacing activity of the ISWI chromatin remodelers.

Like the ISWI chromatin remodelers, the SWI/SNF family chromatin remodelers can also reposition nucleosomes, but they may do so by a different mechanism. To investigate the mechanism by which these remodelers move DNA around the nucleosome, we used smFRET to monitor the structural dynamics of nucleosomes during remodeling by the SWI/SNF enzymes. Our results are consistent with movement of the DNA along its canonical path without substantial lifting of DNA off the edges of the nucleosome or displacement of the H2A-H2B dimer. We observe DNA translocation in 1-2 bp increments at both edges of the nucleosome, which suggests that the motion of DNA at the edges of the nucleosome is driven directly by the action of the ATPase near the dyad of the nucleosome.

Contents

Title i
Copyright ii
Abstract iii
Table of Contents v
Citation to previously published work vii
Acknowledgments viii

1 Introduction **1**
1.1 The structural dynamics of biomolecules 1
1.2 smFRET as a tool for monitoring structural dynamics 3
1.3 smFRET reveals the orientational and translational dynamics of HIV RT . . 8
1.4 Monitoring nucleosome sliding in real time with smFRET 13
1.5 Outline of the dissertation 20

2 Initiation complex dynamics direct the transitions between distinct phases of early HIV reverse transcription **23**
2.1 Introduction 23
2.2 Results 25
2.2.1 Single-molecule FRET assay for studying initiation complex dynamics 25
2.2.2 RT adopts two opposite orientations in initiation complexes 26
2.2.3 Polymerase activity correlates with RT binding orientation 32
2.2.4 The stem-loop near the PBS causes the early pausing 34
2.2.5 Disruption of stem-loop triggers transition to elongation 39
2.2.6 Role of NC in the initiation of DNA synthesis 40
2.3 Discussion 43
2.4 Materials and methods 45
2.4.1 Protein preparation 45
2.4.2 Nucleic acid preparation 46
2.4.3 Single-molecule FRET measurement 46
2.4.4 Single-turnover nucleotide incorporation kinetics assay 48
2.4.5 RT pausing assays 50

3 Histone H4 tail mediates allosteric regulation of nucleosome remodleing by linker DNA **51**
3.1 Introduction 51
3.2 Results 53

3.2.1 Linker DNA length and the histone H4 tail regulate the remodeling pause phases . . . 53
3.2.2 AutoN and the histone H4 tail are involved in linker length sensing . 55
3.2.3 The N-terminal region of Acf1 regulates H4 tail accessibility . . . 58
3.3 Discussion . . . 62
3.4 Materials and methods . . . 63
3.4.1 DNA constructs . . . 63
3.4.2 Histone octamers . . . 64
3.4.3 Nucleosomes . . . 65
3.4.4 ATP-dependent chromatin assembly and remodelling factor (ACF) . 66
3.4.5 Ensemble FRET assay for nucleosome remodelling . . . 67
3.4.6 Single-molecule FRET assay for nucleosome remodelling . . . 69
3.4.7 Fluorescence anisotropy binding assay . . . 70
3.4.8 Protein crosslinking . . . 71
3.4.9 Yeast genetics experiments . . . 72
3.4.10 Data analysis . . . 73

4 Stepwise nucleosome translocation by RSC remodeling complexes 74
4.1 Introduction . . . 74
4.2 Results . . . 76
4.2.1 Single-molecule FRET assay for monitoring nucleosome remodeling . 76
4.2.2 Probing for potential H2A-H2B dimer movement during remodeling . 80
4.2.3 Probing for potential DNA unwrapping at the edge of the nucleosome 85
4.2.4 Characterization of the step size of DNA translocation . . . 90
4.3 Discussion . . . 98
4.4 Materials and methods . . . 100
4.4.1 Histone purification and labeling . . . 100
4.4.2 Preparation of DNA constructs . . . 101
4.4.3 Preparation of nucleosome constructs . . . 101
4.4.4 RSC purification . . . 101
4.4.5 Single-molecule FRET measurements . . . 101
4.4.6 Data analysis . . . 102
4.4.7 Ensemble FRET measurements . . . 103

5 Conclusions and future directions 104

A Unpublished data 110
A.1 Termination of reverse transcription . . . 110
A.2 Effect of DNA sequence on the remodeling direction of RSC . . . 114
A.3 Measuring the path of DNA translocation during nucleosome remodeling by RSC . . . 116
A.4 RSC can reverse directions during remodeling . . . 121
A.5 The kinetic step size of RSC . . . 124

Bibliography 126

Citation to previously published work

Chapter 2 was adapted from the following publication:

Liu S.*, Harada B.T.*, Miller J.T., Le Grice S.F., and Zhuang X. (2010) Initiation complex dynamics direct the transitions between distinct phases of early HIV reverse transcription. *Nature Structural & Molecular Biology 17*, 1453-60.

*denotes co-first authors

Chapter 3 was adapted from the following publication:

Hwang W.L., Deindl S., Harada B.T., and Zhuang X. (2014) Histone H4 tail mediates allosteric regulation of nucleosome remodeling by linker DNA. *Nature 412*, 213-7.

Acknowledgments

Science is a team sport, and the work I present in this dissertation would not have been possible without a number of individuals who have taken the time to guide and support me throughout my graduate career. Because of them, I have grown as a scientist and as a person.

First, I would like to thank my advisor Xiaowei Zhuang. She has been an excellent mentor whose passion and dedication to science provides an excellent model for all of us to follow. She taught me much about how to identify interesting scientific problems, develop the right approaches to study them, and effectively communicate the results. I would also like to thank my past advisors—Marilynn Etzler at the University of California, Davis and Jim Bowie at the University of California, Los Angeles—as I would not be where I am now without them.

Next, I would like to thank Jim Hogle and Michele Jakoulov for all that they do to make the Biophysics Program such a unique and wonderful program. I would also like to thank the faculty members who have volunteered their time to serve on my various committees and provide feedback on my work: David Liu, Eugene Shakhnovich, Shamil Sunyaev, Bob Kingston, Sunney Xie, Timur Yusufzai, Joe Loparo, and Cigall Kadoch.

I would also thank all of the members of the Zhuang laboratory with whom I have worked. Xiaowei has assembled a talented group of scientists from a number of different scientific backgrounds, and it has been a pleasure to work with and learn from all of them. In particular, Shixin Liu helped train me when I first joined the lab, and it was wonderful to be able to work together with him on our studies of reverse transcriptase. I would also like to thank William Hwang for being a great friend and co-worker, for all our discussions about

science or sports, and for teaching me how to play volleyball. I would also like to thank Sebastian Deindl for his helpful insights into chromatin remodeling; John Wu, Elio Abbondanzieri, Tim Blosser, Graham Dempsey, Alistair Boettiger, Ben Altheimer, and Pallav Kosuri for many helpful discussions and suggestions throughout the years; Josh Vaughan and Hazen Babcock for their willingness to always help troubleshoot microscopy problems; Jeff Moffitt and Eileen Sun for help with various experimental techniques; and Alec Goodman and Matt Kilroy for all they do administratively around the lab to make our lives easier.

I thank Jennifer Miller and Stuart Le Grice at the National Cancer Institute, with whom we collaborated for our work on reverse transcriptase, and Robert Gorelick also at the National Cancer Institute for providing NC protein for our studies. I thank Toshio Tsukiyama at the University of Washington and Shankar Mukherji at Harvard for helpful discussions about working with yeast. I also thank Arjan Hada and Blaine Bartholomew at the MD Anderson Cancer Center, with whom we collaborated with on our studies of RSC, and Geeta Narlikar at the University of California, San Francisco for providing the yeast strain for expression of RSC.

I have had a number of wonderful classmates who have been good friends and have helped support me throughout graduate school, including Alice Wang, Igor Lovchinsky, Shan Lou, Kevin Takasaki, Lia Min, Andrew Leifer, Franziska Graf, and Naomi Rajapaksa. I also thank Brian and Samantha Madrid and Kevin and Emily Varee, Peter Aldous, TJ Rakitan, and Nathan Hannon for being such great and supportive friends. I would also like to thank the Zhuang lab's Foosball league and Six Pack Volleyball team as well as the Dudley House intermural volleyball team for helping provide fun distractions outside of lab.

Finally, I would like to thank my parents, John and Susan Harada, and my siblings, Mac and Jaye, for all of their love and support. From a young age, my parents nurtured my interests in science and have always encouraged and pushed me to succeed. They have served as great role models to me and my siblings. This thesis is dedicated to them.

Chapter 1

Introduction

1.1 The structural dynamics of biomolecules

The structures of biological molecules define their function. Proteins and other biomolecules, however, typically exist not as one single static structure but rather an ensemble of conformations that interconvert along an energy landscape defined by the sequence of the biomolecule. These structural dynamics play important roles in a number of biological processes. For example, in biomolecular recognition, biomolecules can often exist in equilibrium between binding-competent and binding-incompetent conformations that can regulate the binding of a ligand to its receptor (a process termed conformational selection), and binding of a ligand to its receptor can further alter the energy landscape of the molecule to induce additional conformational changes (a process termed induced fit) [1]. These ligand-induced changes to the structure or dynamics of molecules underlie the ability of ligands to regulate the activity of biomolecules at sites distant to the ligand binding site (a process termed allostery) [2, 3]. Biomolecular dynamics also play important roles in catalysis, and the rate limiting steps of many enzymes involve protein conformational changes [4, 5]. Finally, the conformational dynamics of a biomolecule may influence the evolutionary pathways available to that molecule [6]. Thus, characterizing the structural dynamics of biomolecules is crucial for understanding their function, regulation, and evolution.

Given the importance of the structural dynamics of molecules, a great deal of effort has been made to develop techniques able to study the structural dynamics of enzymes and their substrates [7, 8, 9, 10]. In particular, single-molecule methods have enabled researchers to watch the movements of proteins and nucleic acids in real time, providing a powerful means to understand the mechanism of action and regulation of the molecules under study [11, 12, 13, 14]. In this dissertation, I develop and use a single-molecule Förster resonance energy transfer (FRET) [15] as a tool to monitor the dynamics of biomolecules. FRET is a phenomena that can report on the distance between two dye molecules [16], and because the readout (fluorescence) is non-perturbative, these distance changes can be observed in real time. Performing measurements at the single-molecule level offers a number of advantages over techniques that rely on ensemble averaging [17]. First, many biomolecular systems exhibit significant amounts of heterogeneity, for example, a protein can interact with its partner in a number of different conformations. Such heterogeneity confounds traditional methods that rely on measuring the average of a quantity over a heterogeneous population of molecules. By tracking individual molecules, one can identify the different populations that exist within a sample and characterize each population separately. Second, many biomolecular processes involve the formation of transient, short-lived intermediates, and characterizing the identities of these intermediates can aid in obtaining a mechanistic understanding of the process under study. Because they exist only transiently, the intermediates compose a small fraction of an ensemble of molecules at any single time, making their detection difficult in experiments involving ensemble averaging. Using single-molecule FRET (smFRET) to track individual molecules allows one to directly observe the formation and dissipation of these intermediates in real time. The FRET value of these intermediate states reports on their structures while the distribution of lifetimes of these states provides information about the kinetics and mechanisms of the reactions that govern interconversion between the states [18]. Examining the lifetime of these intermediates under different conditions can reveal which factors regulate each substep of a multi-step enzymatic process. Thus, smFRET provides a

powerful means to study heterogeneous systems, characterize the structures of intermediates, and dissect the mechanism and regulation of biomolecular processes.

In this dissertation, I apply smFRET techniques to the study of two processes driven by molecular motor proteins: reverse transcription and chromatin remodeling. The motor proteins involved in these process couple the dissipation of energy (the hydrolysis of nucleoside triphosphates) to directed motion along a nucleic acid template. Because these systems operate under non-equilibrium, steady-state conditions; catalyze complex, multi-step processes; and exhibit heterogeneous behaviors, they are excellent candidates for study using smFRET. In this introduction, I will first briefly describe smFRET as a tool to study structural dynamics (I refer readers to refs [19] and [20] for a more comprehensive description of the technique), then I will review how previous smFRET studies of these systems have contributed to our understanding of reverse transcription and chromatin remodeling.

1.2 smFRET as a tool for monitoring structural dynamics

FRET occurs when two fluorophores, a donor fluorophore in its excited state and an acceptor fluorophore in its ground state, are in close proximity. Dipole-dipole interactions between the two fluorophores can result in the transfer of energy from the donor to the acceptor without the intermediate emission of a photon, resulting in an excited acceptor and a ground-state donor. The rate at which energy transfer occurs (k_T) depends on the spectral overlap of the donor and acceptor (represented by the overlap integral $J(\lambda)$), the quantum yield of the donor (Q_D), the relative orientation of the donor and acceptor (represented by the orientation factor κ^2), and, importantly, the distance between the donor and acceptor fluorophores (r) [21]:

$$k_T = \frac{Q_D \kappa^2}{\tau_d r^6} \left(\frac{9000 \ln(10)}{128 \pi^6 N_A n^4} \right) J(\lambda) \tag{1.2.1}$$

$$J(\lambda) = \frac{\int_0^\infty F_D(\lambda) \epsilon_A(\lambda) \lambda^4 d\lambda}{\int_0^\infty F_D(\lambda) d\lambda} \tag{1.2.2}$$

where τ_d is the lifetime of the donor in the absence of acceptor, N_A is Avogadro's number, n is the index of refraction of the medium, $F_D(\lambda)$ is the emission spectrum of the donor, and $\epsilon_A(\lambda)$ is the extinction coefficient of the acceptor. Under the assumptions that the photophysical properties of the dyes are constant and the relative orientation of the dyes do not change (often accomplished by attaching the dyes via a flexible linker and assuming the dyes will be free to rotate and sample all relative orientations), the rate of energy transfer depends primarily upon the distance between the dyes. Because energy transfer competes with all other pathways for the relaxation of the donor back to the ground state, measuring the FRET efficiency (E_{FRET}), the fraction of excited donors that undergo energy transfer, provides a means of estimating the distance between the donor and acceptor:

$$E_{FRET} = \frac{k_T}{\tau_d^{-1} + k_T} = \frac{1}{1 + \left(\frac{r}{R_o}\right)^6} \tag{1.2.3}$$

$$R_o^6 = Q_D \kappa^2 \left(\frac{9000 \ln(10)}{128 \pi^6 N_A n^4}\right) J(\lambda) \tag{1.2.4}$$

where R_o is called the Förster radius and corresponds empirically to the distance at which $E_{FRET} = 0.5$. Because the typical Förster radii of fluorescent dye pairs lies in the 2-9 nm range [21], FRET is well suited for studying distance changes on the nanometer scale. For the Cy3-Cy5 dye pair that I use, R_o ~6 nm [22].

In order to track the FRET efficiency from individual labeled molecules in real time, I use a prism-type total-internal-reflection-fluorescence (TIRF) microscope to image molecules immobilized onto a microscope slide (Figure 1.2.1). The microscope slide is passivated with a layer of poly(ethylene glycol) (PEG), a hydrophilic polymer that helps prevent the non-specific adhesion of proteins to the slide. A small fraction of these PEG molecules are labeled with a biotin at the end, allowing streptavidin to bind. Because streptavidin is multivalent, the streptavidins can link a biotin-labeled biomolecule to the microscope slide. The molecules are immobilized at a density such that individual molecules are not overlapping and appear

as diffraction limited spots. They are excited by a 532 nm laser, and the emitted fluorescence is collected by an objective lens. The fluorescence is separated by a dichroic mirror, which splits the donor and acceptor fluorescence into two channels that are imaged onto two halves of an electron-multiplying charge-coupled device (EM-CCD) camera. The images are aligned, and the apparent FRET efficiency (E_{app}) is calculated from the fluorescence intensities of the donor (I_d) and acceptor (I_a):

$$E_{app} = \frac{I_a}{I_a + \gamma I_d}$$

where γ is an empirically determined correction factor to account for the different quantum yields and detection efficiencies of donor and acceptor fluorescence. This factor is ~1 for the Cy3-Cy5 pair I use. Although other methods exist for performing smFRET measurements (for example, without immobilization [23]), this method enables the real-time monitoring of FRET from individual molecules over long periods of time.

Measuring FRET at the single molecule level provides a number of advantages in the study of dynamic systems when compared to traditional structural and biochemical methods that involve ensemble averaging. First, single-molecule measurements allow one to measure the full distribution of FRET values in a population whereas ensemble FRET measurements can measure only the average FRET from the billions of molecules in solution. The full distribution of FRET values often provides much more information about the system—for example, Figure 1.2.2a-c shows three different distributions of FRET values that suggest widely different behaviors, but all have the same average FRET and would appear identical in ensemble FRET experiments. The situation is most problematic in Figure 1.2.2c, in which there are two populations of molecules with FRET of 0.2 and 0.8; in this case, the average FRET (0.5) is not representative of any individual molecule in the sample. A second limitation of ensemble measurements is in the characterization of intermediates in multistep processes. Although one can synchronize the start of a reaction, chemical processes occur stochastically, so the individual molecules quickly become desynchronized as the

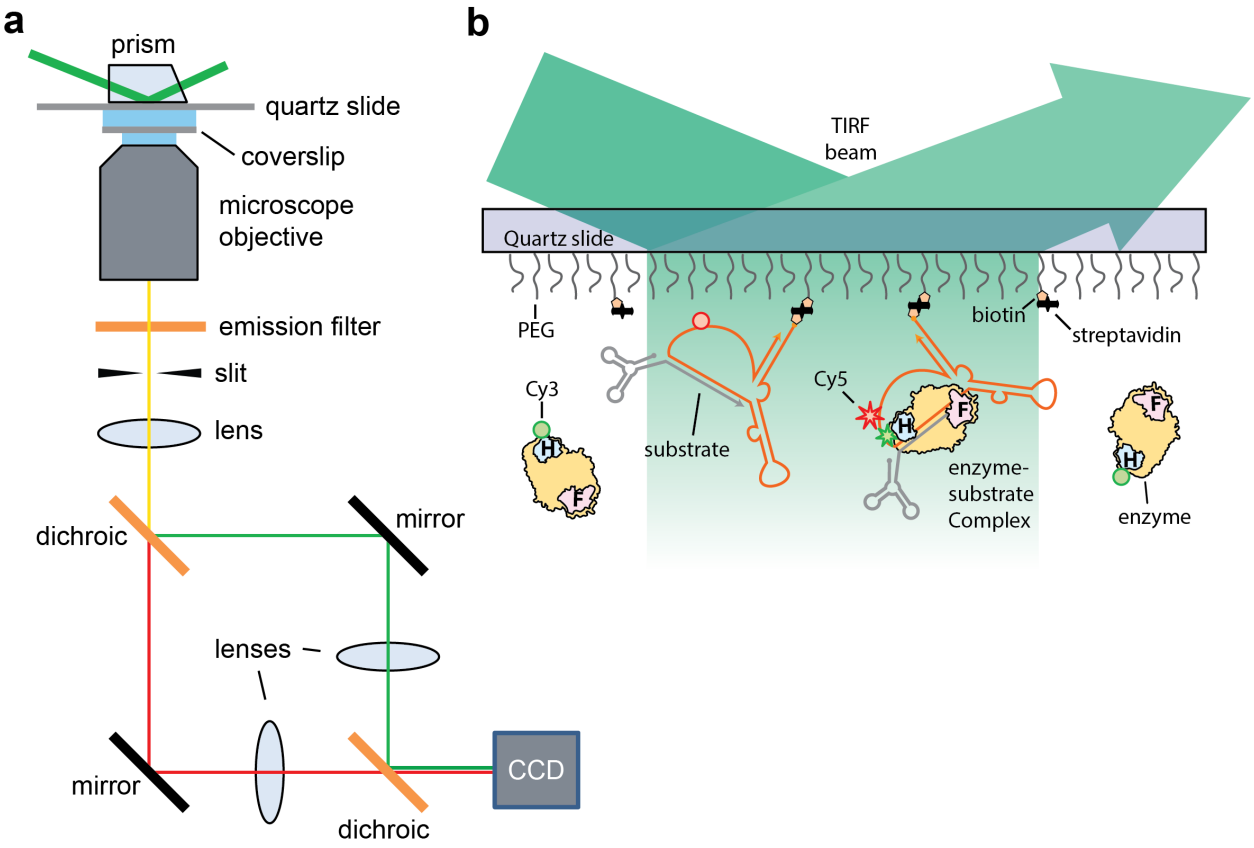


Figure 1.2.1: Monitoring the structural dynamics of biomolecules using smFRET. **a**, Diagram of the instrument used for smFRET experiments (not to scale). Labeled molecules are immobilized onto a quartz microscope slide which is illuminated by TIRF through a prism placed on top of the slide. The emitted fluorescence is imaged through a glass coverslip by a 60x water-immersion objective. The light collected by the objective is sent through an emission filter (550 nm long-pass), a rectangular slit, and a lens. A dichroic mirror (630 nm cutoff) separates the Cy3 and Cy5 emission, which are refocused by a pair of lenses, recombined at a second dichroic, and imaged onto two halves of an EM-CCD camera. **b**, Cartoon of the single molecule assay used for studying reverse transcriptase. The quartz microscope slide is functionalized with a 100:1 mixture of PEG:biotin-PEG polymers. Substrate molecules (orange/gray) labeled with the FRET aceptor dye Cy5 (red) are attached to the functionalized surface via a biotin-streptavidin linkage. Enzyme (yellow) labeled with the FRET donor dye Cy3 (green) are introduced in solution. Illumination by TIRF excites only those molecules within ~100-200 nm of the surface, reducing the background from dye-labeled molecules diffusing through the flow channel that do not bind stably to substrate molecule on the surface. The Cy5 on unbound substrate molecules is not excited by the 532 nm laser used for illumination.

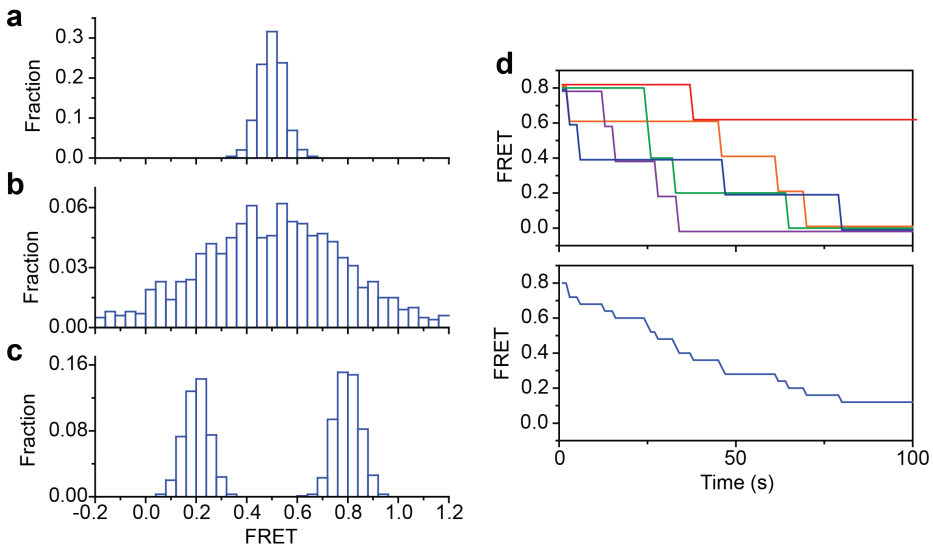


Figure 1.2.2: Advantages of smFRET over ensemble assays. **a-c**, Simulated histograms showing the distribution of FRET values from a molecule showing a single, well-defined distance between the fluorophores (a), a highly variable distance between the two fluorophores (b), and a sample containing two populations of molecule showing two different distances between the two fluorophores (c). **d**, Top: Simulated smFRET traces from five different molecules showing a stepwise decrease in FRET from 0.8 to 0.0 in increments of 0.2. The steps occur stochastically with a time constant of 20 s. Bottom: The FRET time trace obtained from averaging the five single-molecule traces shown above. Whereas the individual pauses are readily identifiable in the single-molecule traces, the major pause locations are difficult to distinguish in the averaged trace.

reaction progresses, limiting the ability of ensemble measurements to detect transient intermediates (Figure 1.2.2d). Thus, traditional studies of intermediates often rely on mutations or small-molecules to trap the system in an intermediate state, a situation that can introduce unwanted perturbations and is not always feasible for some intermediates. In contrast, single-molecule experiments allow one to directly observe and characterize the intermediates and determine their FRET values and lifetimes. Finally, single-molecule methods enable one to deal with heterogeneity in the sample. For example, if the system consists of a heterogeneous population of molecules, each class of molecule can be distinguished and analyzed separately. Similarly, if an enzyme exhibits multiple reaction pathways, the different classes of FRET traces can be sorted and analyzed separately.

In the following sections, I will review previous smFRET studies of two different classes of

enzymes, reverse transcriptases and chromatin remodeling enzymes. In both cases, smFRET enabled insights into the function of these enzymes that would have been difficult to obtain with traditional techniques that rely on ensemble averaging. In the case of reverse transcriptase, the capability of FRET to measure the distribution of FRET values was important for distinguishing the various modes by which it can interact with its substrate [24, 25]. In the study of chromatin remodeling enzymes, the capability of FRET to characterize the structure of intermediates in multistep processes enabled new insights into the mechanism of chromatin remodeling [26, 27].

1.3 smFRET reveals the orientational and translational dynamics of HIV RT

The reverse transcriptase (RT) of the human immunodeficiency virus (HIV) is responsible for converting HIV's single-stranded RNA genome into a double-stranded DNA competent for integration into the host cell's genome [28, 29]. Because of its essential role in viral replication, RT is a major drug target, and understanding its structure and function is of great medical importance [30, 31, 32]. RT is a member of the DNA polymerase family and consists of a heterodimer of p66 and p51 subunits. Like all DNA polymerases, its structure resembles a right hand, with the DNA polymerase active site sitting at a cleft near where the fingers, thumb, and palm domains of the enzyme meet [33]. Unlike other DNA polymerases, however, RT has two additional domains at its C-terminus, the connection domain and the RNase H domain, which contains an active site that catalyzes the hydrolysis of RNA in RNA-DNA duplexes [34]. RT's nucleic acid substrate binds in a cleft spanning the entire enzyme, and the points at which the polymerase and RNase H active sites contact the substrate are separated by about 19-22 bp [35, 36, 37].

Because it has two separate active sites on opposite sides of the enzyme, understanding how RT interacts with its substrates is critically important to understanding how this

multifunctional enzyme regulates its various activities. Indeed, during reverse transcription, RT encounters a number of different nucleic acid substrates on which it performs a number of different activities [30, 31, 32] (Figure 1.3.1). Reverse transcription initiates from a cellular tRNA that is hybridized to a specific sequence in the 5' UTR of the viral RNA (vRNA) [38]. After elongating this RNA-RNA duplex, RT then works on a DNA-RNA hybrid where it performs RNA-directed DNA-synthesis to synthesize the (–)-strand of DNA [28, 29]. Throughout (–)-strand synthesis, RT also uses its RNase H activity to degrade the RNA in order to make room for the synthesis of the (+)-strand of DNA [39, 40, 41, 42]. However, two regions of the vRNA, the polypurine tracts, are resistant to RNase H cleavage and act as primers for (+)-strand synthesis [43, 44, 45, 46]. (+)-strand synthesis involves an intramolecular strand transfer that cyclizes the substrate, allowing the completion of both (+) and (–) strand synthesis [47]. Finally, reverse transcription involves a termination step where strand-displacement synthesis over the central termination sequence (CTS) stalls RT, creating a central flap sequence [48].

Given RT's need to interact with a variety of RNA-RNA, RNA-DNA, and DNA-DNA substrates, Abbondanzieri *et al.* developed a smFRET assay to examine how RT interacts with the various substrates it encounters during reverse transcription [24]. When binding to a substrate containing a short DNA primer hybridized to a longer DNA template, the smFRET measurements are consistent with the binding mode observed in crystal structures of the enzyme bound to DNA [36, 37, 49, 50]: the fingers domain (near the polymerase active site) lies near the 3' end of the DNA primer, and the RNase domain lies near the 5' end of the primer (Figure 1.3.2a). When RT interacts with a RNA primer hybridized to a DNA template, however, RT binds in a mode that is flipped by 180°: the RNase H domain now resides over the 3' end of the RNA primer and the polymerase active site resides near the 5' end of the primer (Figure 1.3.2b). In this flipped orientation, the polymerase active site is far away from the 3' end of the primer, so the enzyme is unable to perform its polymerase activity in this orientation. Examining which features of the substrate determine the enzyme's bind-

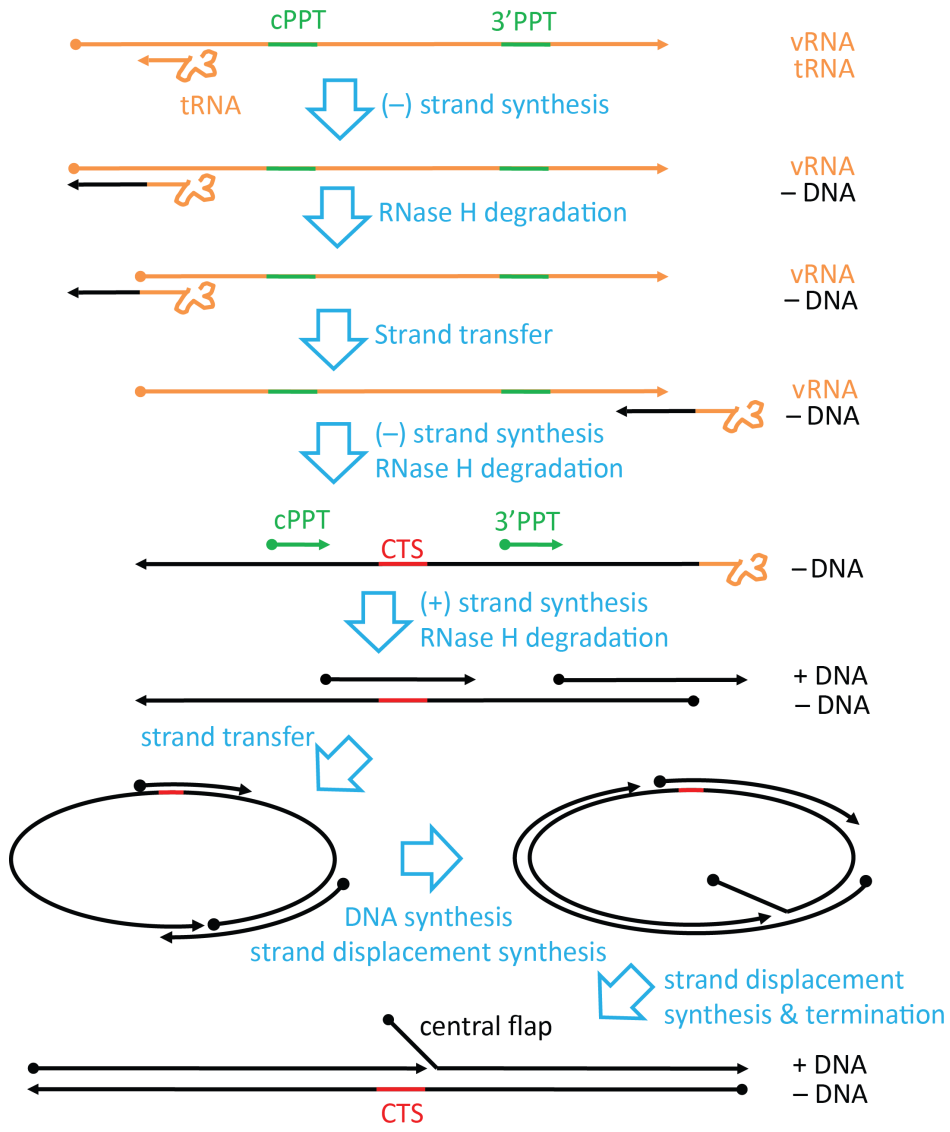


Figure 1.3.1: RT encounters a number of different substrates during reverse transcription. Diagram showing the various intermediates formed during reverse transcription. RNA strands are shown in orange, and DNA strands are shown in black. 5' ends of strands are denoted with a circle and 3' ends of strands are denoted with arrowheads. The PPT RNAs are highlighted in green and the DNA containing the CTS is highlighted in red.

ing orientation revealed that the sugar phosphate backbone at the 5' end of the primer has the strongest effect on determining the binding orientation of RT. When these nucleotides are DNA, RT binds predominantly in the polymerase-competent orientation regardless of whether the 3' end consists of RNA or DNA nucleotides, and similarly, RNA nucleotides at the 5' end of the primer direct RT to bind in the flipped orientation. Furthermore, the ability of RT to extend these primers correlates with the binding orientation of the enzyme: On substrates that RT binds predominantly in the polymerase-competent orientation, primer extension occurs quickly, and on substrates that RT binds predominantly in the flipped orientation, primer extension occurs slowly. These results indicate that binding orientation is an important factor regulating RT's polymerase activity [24].

The ability of RT to bind in the polymerase-competent and flipped orientations is important for its function on the polypurine tract (PPT) RNAs that prime (+)-strand synthesis [24]. Reverse transcriptase must be able to perform RNase H cleavage at the precise 3' end of the PPT sequence because the location of the 3' PPT sets the 5' end of the viral DNA. If the PPT sequence is not cleaved at the correct location, the resulting viral DNA will not be a competent substrate for integration [51]. After RNase H cleavage, RT must then use its DNA polymerase activity to extend the PPT primer to initiate (+)-strand synthesis. Consistent with the notion that RT must perform both of its enzymatic activities on the PPT substrate [52], smFRET experiments revealed that RT binds the PPT substrate in both the polymerase-competent and flipped orientations (Figure 1.3.2c) [24]. Surprisingly, the data also showed RT transitioning between the two orientations without dissociating, indicating that it can sample between the two orientations during a single binding event (Figure 1.3.2c).

While these studies were performed on a short (~19 bp) duplex, the size of RT's binding footprint, examining RT's interaction with longer RNA-DNA hybrids revealed additional dynamics [25]. On longer substrates, RT shuttles back and forth between the two ends of the duplex (Figure 1.3.2d). Because the enzyme could be observed at intermediate locations when transiting between the two ends of the duplex, the movement represents sliding of the

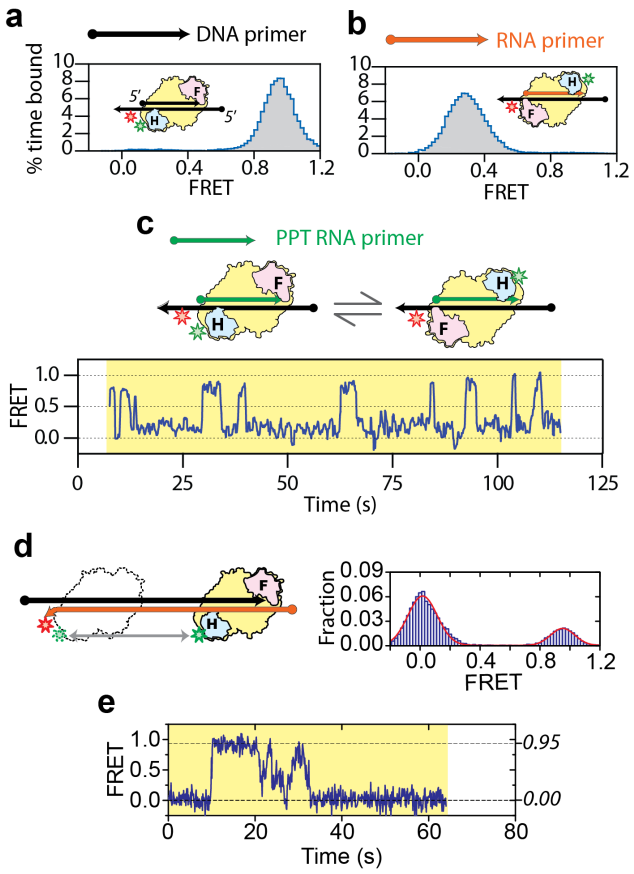


Figure 1.3.2: Orientational and translational dynamics of RT. **a-b,** the FRET histograms obtained from the interaction of RT (yellow) with a DNA primer (black) (a) and a RNA primer (orange) (b) hybridized to a DNA template (black). RT is labeled near its RNase H domain with the FRET donor, Cy3 (green star), while the substrates are labeled near the 5' end of the primer with the FRET acceptor dye, Cy5 (red star). The fingers domain of RT, which resides near the polymerase active site, is indicated with an F and highlighted in pink. The RNase H domain of RT is indicated with an H and highlighted in blue. **c**, Top: Cartoon showing the two binding orientations of RT on a substrate containing the PPT RNA primer (green) annealed to a DNA template. Bottom: FRET time trace from the binding of RT to the PPT RNA showing that RT can bind in both the polymerase (high FRET) orientation and flipped (low FRET) orientation and transition between the two orientations during a single binding event. The enzyme and substrate are labeled with Cy3 and Cy5 as in panels a-b. **d**, Left: Cartoon showing RT's interactions with a 56 bp RNA/DNA duplex. Right: FRET histograms demonstrating that RT binds to the 56 bp duplex at either the front end (low FRET) or back end (high FRET). **e**, FRET time trace showing RT shuttling between the front end and back end binding modes. The transition occurs gradually via states with intermediate FRET, supporting a model in which RT slides along the duplex. Adapted with permission from references [24] and [25].

enzyme along the duplex, rather than discrete hopping between the two sites (Figure 1.3.2e). The ability of RT to slide along the duplex allows the enzyme to find the primer end via facilitated diffusion after dissociation and aids the enzyme to perform strand displacement synthesis [25].

Together, these studies highlight the ability of single-molecule techniques to identify structural heterogeneity in biomolecular interactions. By making use of smFRET to characterize the structural features of the different binding modes and by combining these measurements with traditional biochemistry to elucidate the functional relevance of the binding dynamics, they reveal how the orientational and translational dynamics of RT aid in the process of reverse transcription [24, 25]. Despite these findings, important questions still remain. For example, although these studies addressed the interaction of RT with double-stranded DNA and RNA-DNA hybrids, RT must elongate a double-stranded RNA substrate during the initiation of reverse transcription [38]. Furthermore, both the tRNA primer and viral RNA template contain evolutionarily conserved RNA secondary structures that regulate the elongation of the tRNA primer [53, 54, 55, 56, 57]. How does RT interact with its double-stranded RNA substrate during initiation, and might the enzyme also exhibit binding orientation dynamics that aid in regulating initiation? Does the tRNA-vRNA substrate exhibit any structural dynamics important for the regulation of reverse transcription? Because of its ability to dissect structural heterogeneity in intermolecular complexes, smFRET is an ideal tool to answer these questions.

1.4 Monitoring nucleosome sliding in real time with smFRET

Eukaroytes package their genomes in a protein-DNA complex known as chromatin [58]. The fundamental unit of chromatin is the nucleosome, which consists of 147 bp of DNA wrapped 1.7 times around an octamer of histone proteins (two copies each of H2A, H2B, H3, and

H4) [59, 60]. Nucleosomes restrict access to the DNA sequences they wrap, acting as an important regulatory feature in eukaryotic cells. In order to make nucleosome-occluded sequences available for replication, recombination, repair, transcription and other processes, the cell has evolved an elaborate regulatory system to alter chromatin structure in response to environmental and developmental cues as well as maintain it during normal cell function [61, 62]. Crucial components of this regulatory machinery are the ATP-dependent chromatin remodeling complexes, a family of enzymes that use the energy of ATP hydrolysis to alter chromatin structure [63, 64, 65, 66, 67]. At least four families of chromatin remodeling enzymes—SWI/SNF, ISWI, CHD, and INO80/SWR1—exist in eukaryotes, all exhibiting distinct functions within the cell. For example, in yeast, the SWI/SNF-family remodeler RSC acts to promote transcription by establishing nucleosome-free regions upstream of transcription start sites [68, 69, 70] while the ISWI-family remodelers ISW1 and ISW2 antagonize RSC by moving nucleosomes toward nucleosome-free regions and promoting transcriptional silencing [71, 72, 73]. Through their effects on chromatin structure, the remodeler enzymes play important roles in development, such as in the maintenance and establishment of pluripotent and multipotent states in cells [74], and defects in chromatin remodeling underlie a number of developmental disorders and cancers [75, 76].

Understanding the mechanisms of chromatin remodeling, in particular, how remodelers couple the free energy from the hydrolysis of ATP to the disruption of histone-DNA contacts and repositioning/restructuring of nucelosomes, remains a major challenge in the field. Because this dissertation focuses on the ISWI and SWI/SNF family of remodeler enzymes, I will limit my discussion to these two families. Both families of enzyme are capable of repositioning the octamer *in cis*, altering the translational position of the nucleosome on the DNA [77, 78, 79], but the SWI/SNF enzymes also exhibit histone dimer exchange and dimer/octamer ejection activities [80, 81, 82, 83, 84, 85, 86] while the ISWI remodelers also participate in chromatin assembly [87, 88, 89, 90]. The catalytic subunits of these remodelers contain a SF2 ATPase domain that contacts the nucleosomal DNA at a site

20 bp from the nucleosome dyad (referred to as SHL±2) and translocates along the DNA [91, 92, 93, 94, 95, 96, 97, 98]. Unlike other DNA translocases, however, remodelers face a major challenge: their DNA substrate is highly constrained in a protein-DNA complex, so translocation must somehow be coupled to the disruption of histone-DNA contacts.

Major efforts have been made to understand how the translocation of DNA by the ATPase is coordinated with the movement of DNA into the nucleosome at one edge (termed the entry side) and the movement of DNA out of the nucleosome at the other edge (termed the exit side), and have led to numerous proposals for the mechanism by which remodelers can reposition nucleosomes (reviewed in refs [63, 65, 66, 67, 99]) (Figure 1.4.1). In the twist-diffusion model, DNA translocation slides the DNA along its canonical path around the nucleosome, maintaining the majority of the histone-DNA contacts intact (Figure 1.4.1a). Such models are disfavored because they require the DNA to rotate as it slides around the nucleosome, and obstacles to DNA rotation do not hinder nucleosome repositioning [100, 101, 102, 103]. Another set of models, the loop/bulge propagation models, propose that remodelers generate a transient loop or bulge of DNA on the surface of the nucleosome that propagates around the nucleosome and resolves at the exit side (Figure 1.4.1b). Consistent with this idea, SWI/SNF remodelers have been reported to transiently disrupt the nucleosome structure during remodeling, exposing significant amounts of DNA to nucleases [104, 105, 106, 107]. Furthermore, various approaches have reported evidence that SWI/SNF remodelers can lift DNA from the surface of the histone octamer to form intranuleosomal loops [104, 107, 108, 109, 110, 111]. In contrast, ISWI remodelers expose relatively little DNA to digestion during remodeling [112, 113], though there is evidence that the remodeler may also transiently lift DNA from the surface of the nucleosome [103]. Finally, concerted sliding models have been proposed in which the remodeler globally disrupts histone-DNA contacts to move the DNA around the nucleosome in a concerted manner (Figure 1.4.1a).

Distinguishing between these models requires the ability to structurally characterize intermediates of the nucleosome sliding reaction. While ensemble FRET studies allow character-

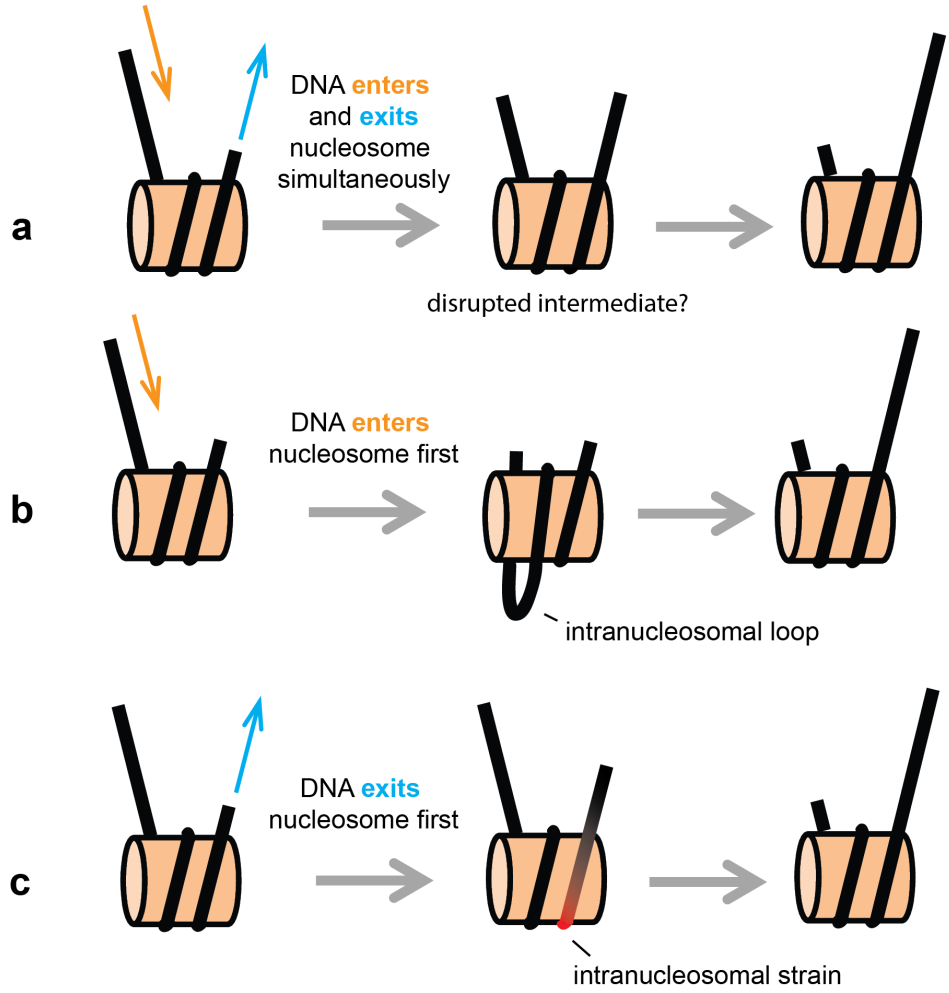


Figure 1.4.1: Nucleosome remodeling mechanisms. The histone octamer is depicted as a tan cylinder and the DNA as a black line. **a**, In twist-diffusion and concerted motion models, the DNA enters and exits the nucleosome simultaneously. Twist-diffusion models propose that intermediates of the sliding process maintain most of the histone-DNA contacts whereas concerted motion models propose that many histone-DNA contacts are disrupted in intermediates. **b,** In loop-propagation models, DNA enters the nucleosome before DNA exits, producing an intranucleosomal loop that propagates around the nucleosome and resolves at the other side. **c,** In tension-based models, DNA exits the nucleosome before any DNA enters, producing intranucleosomal strain (either stretching/unwinding of the DNA or conformational changes to the histone octamer).

ization of the overall kinetics of nucleosome sliding [114, 115], the application of smFRET to these systems enables detailed structural characterization of the intermediates (Figure 1.4.2). Blosser *et al.* developed a smFRET assay to monitor nucleosome sliding by the human ISWI enzyme ACF and found that remodeling occurs in a series of alternating translocation phases where DNA exits the nucleosome, and pause phases where no DNA motion occurs on the exit side [26]. Interestingly, the first pause occurs after 7 bp of DNA exits the nucleosome, but subsequent pauses occur after the translocation of 3-4 bp. Further work demonstrated that the DNA moves in 1 bp increments during the 7- and 3-bp translocation phases [27]. In contrast, the same study found that DNA enters the nucleosome in ~3 bp increments. These experiments also demonstrated that movement of DNA into the nucleosome first requires ~7 bp of DNA to exit the nucleosome.

Together, these results revealed an entirely new mechanism of chromatin remodeling [27]. Unlike the loop propagation model, which predicts that DNA enters the nucleosome before DNA exits, and concerted sliding models, which predict that DNA enters and exits simultaneously, smFRET studies of ISWI remodelers found that DNA exits the nucleosome prior to entering. Furthermore, whereas twist diffusion models predict an elementary step size of 1 bp at both the entry and exit sides of the nucleosome, ISWI remodelers show 1 bp steps at the exit side, but motion in 3 bp increments at the entry side. The results instead, support a "tension-based" mechanism for chromatin remodeling (Figures 1.4.1c and 1.4.3). In this model, DNA translocation by the ATPase results in DNA being pushed out of the exit side one bp at a time. This ATPase action does not result in any movement at the entry side, presumably because the enzyme's HAND-SANT-SLIDE (HSS) domain binds the DNA at the entry side and prevents its movement. Therefore, strain builds up as the enzyme pumps DNA out of the nucleosome without allowing any DNA to enter, creating a deficit of DNA within the nucleosome. When this deficit reaches 7 bp, the strain triggers a conformational change in the enzyme that causes the enzyme to move 3 bp of DNA into the nucleosome at the entry side, reducing the DNA deficit to 4 bp. This reduction in the

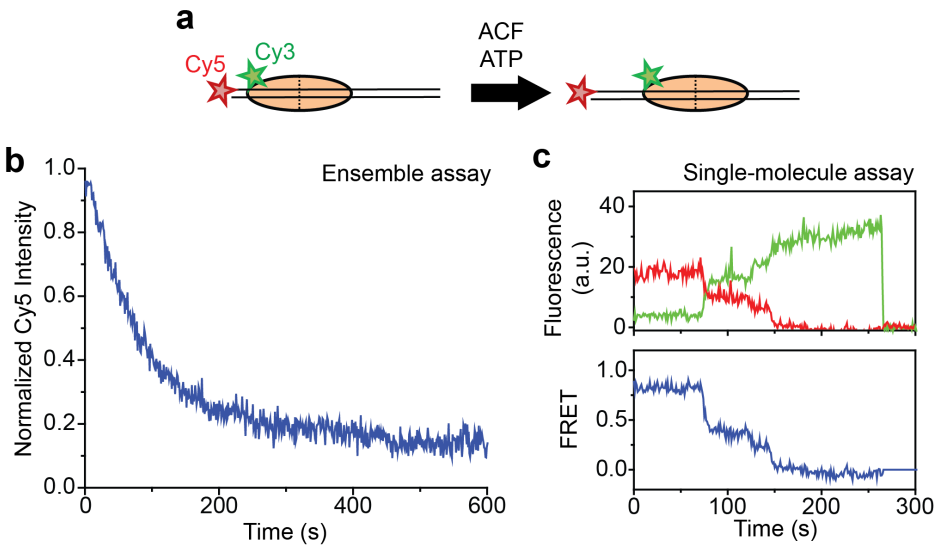


Figure 1.4.2: Ensemble versus single-molecule FRET assays for monitoring nucleosome remodeling. **a,** Cartoon of the remodeling reaction. The footprint of the histone octamer (orange oval) is shown on double-stranded DNA (black lines). The Cy3 label on the histone octamer initially resides close to the Cy5 on the end of the DNA, producing a high FRET signal. Addition of ACF and ATP repositions the octamer to the center of the DNA, causing a decrease in FRET to zero. **b,** Ensemble FRET assay monitoring the nucleosome remodeling reaction in panel a. The assay monitors the Cy5 emission from the nucleosomes under 532 nm illumination. **c,** Single-molecule FRET assay monitoring the nucleosome remodeling reaction in panel a. The Cy3 intensity (green), Cy5 intensity (red), and FRET (blue) time traces are shown. The single-molecule assay reveals that the FRET decrease occurs in a stepwise manner, enabling pauses at intermediate FRET values to be observed.

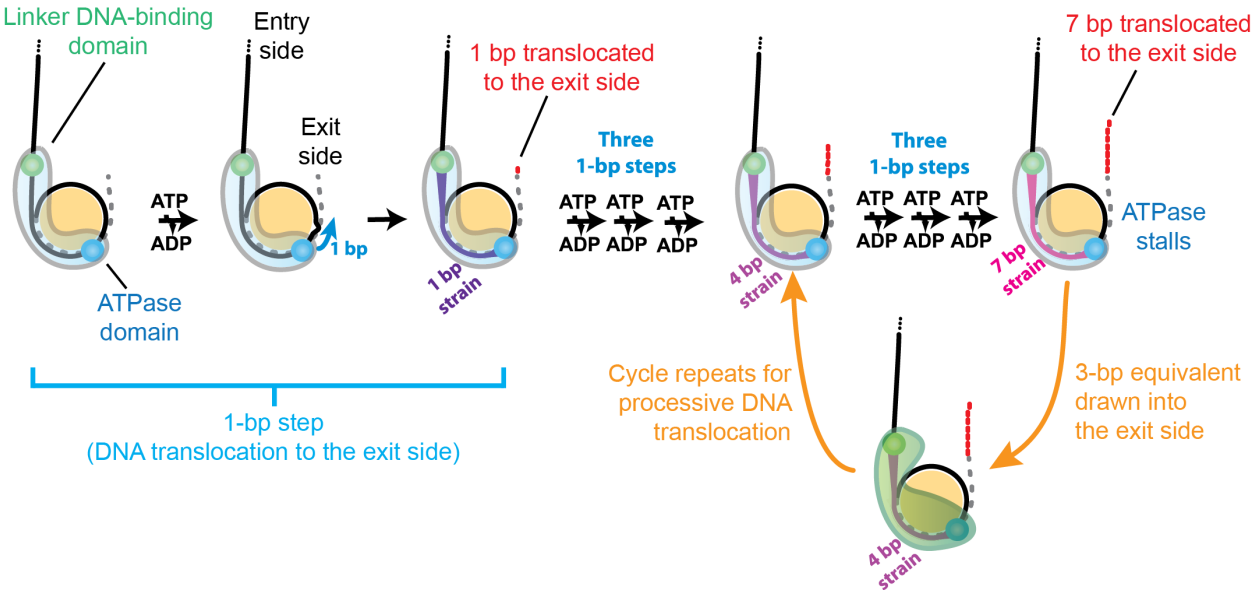


Figure 1.4.3: Tension-based mechanism for ISWI-family chromatin remodeling enzymes. DNA, histone octamer, and remodeler are shown in black/grey/red, yellow, and blue/green, respectively. The upper and lower DNA gyres are depicted as solid black and dashed grey lines, respectively. Each base pair of DNA translocated to the exit side is shown by a red dot. The remodeler is depicted as a semi-transparent shape with the linker DNA-binding domain and ATPase domain domain depicted as green and blue spheres, respectively. Remodeling by ISWI remodelers begins when the ATPase translocates DNA one base pair at a time from SHL2 to the exit side. Because no DNA motion occurs at the entry side during translocation, strain builds up in the DNA between the linker DNA-binding domain and the ATPase domain. Once 7 bp of strain have built up, the ATPase domain stalls, producing a pause, and the linker-DNA binding domain draws a 3-bp equivalent into the entry side, reducing the strain to 4 bp. The ATPase can then translocate an additional three 1-bp steps to repeat the cycle and enable processive DNA translocation. Adapted with permission from reference [27].

DNA deficit allows an additional 3 bp of exit-side translocation, increasing the deficit back to 7 bp, which triggers 3 bp of entry-side movement, and this cycle continues, allowing the enzyme to processively slide DNA around the nucleosome.

Although these studies have made crucial insights into the mechanism of chromatin remodeling, many questions remain. First, ISWI remodelers are regulated by a number of factors, including H4 tail modifications [116, 117, 118, 119, 120] and linker DNA length [114, 118, 121, 122, 123]. Interestingly, these factors alter the catalytic activity of the enzyme, not just their recruitment to the nucleosomes, and it is not known how these factors

impinge upon the remodeling mechanism to alter the kinetics of nucleosome sliding. Second, although SWI/SNF remodelers have an ATPase domain that is homologous to the ISWI remodelers, they are more disruptive to the nucleosome structure than ISWI remodelers [113, 124], so it is unclear if they share a common remodeling mechanism. Indeed, SWI/SNF remodelers lack the HSS domain that seems crucial to the tension-based mechanism of the ISWI remodelers [27, 125]. Therefore, determining how SWI/SNF remodelers differ from ISWI remodelers in their remodeling mechanism, and how these differences contribute to their divergent functions in the cell, is a question of outstanding interest.

1.5 Outline of the dissertation

smFRET assays provide a powerful means of studying the mechanism and regulation of biochemical and biophysical systems. I apply this technique to study three different enzymes: HIV reverse transcriptase (Chapter 2), and the chromatin remodeling enzymes human ACF (Chapter 3) and yeast RSC (Chapter 4). I summarize the main findings from these studies below:

HIV initiates reverse transcription of its viral RNA genome from a cellular $tRNA^{Lys,3}$ primer that binds to a specific region of the vRNA. This process is characterized by a slow initiation phase with specific pauses, followed by a fast elongation phase. In Chapter 2, we use single-molecule FRET to monitor the dynamics of individual initiation complexes, composed of vRNA, tRNA and HIV reverse transcriptase. RT transitions between two opposite binding orientations on the tRNA-vRNA substrate, and the pausing events during initiation are related to RT binding in an inactive orientation, where the enzyme is flipped 180° from the polymerization-competent configuration. We examine the contributions of the tRNA and vRNA secondary structure in regulating RT's behavior, and find that a stem-loop structure within the vRNA is responsible for maintaining the enzyme predominantly in the flipped orientation at the major pause locations. Disruption of the stem-loop structure

triggers the initiation-to-elongation transition. These results demonstrate how the structural dynamics of the initiation complex directs transitions between the early phases of reverse transcription. This work was published in reference [126].

The ISWI-family chromatin remodeling enzymes promote heterochromatin formation and transcriptional silencing by generating regularly spaced nucleosome arrays. The nucleosome-spacing activity arises from the dependence of nucleosome translocation on the length of extranucleosomal linker DNA, but the underlying mechanism for this regulation remains unclear. In Chapter 3, we study nucleosome remodeling by the human ACF complex, an ISWI enzyme comprising a catalytic subunit, Snf2h, and an accessory subunit, Acf1. We find that ACF senses linker DNA length through a mechanism involving the accessory and catalytic subunits of ACF as well as the histone H4 tail of the nucleosome. Mutation of AutoN, an auto-inhibitory region of Snf2h that resembles the H4 tail, eliminates the linker length sensitivity of the enzyme. Addition of exogenous H4 peptide or mutation of the H4 tail also diminishes the linker length sensitivity of ACF. Moreover, Acf1 binds the H4 tail via its N-terminal region, and this interaction is dependent on linker DNA length. Deletion of the N-terminal region eliminates the linker length sensitivity of ACF *in vitro*, and deletion of the homologous region of yeast ISW2 causes a severe growth defect. Together, these results suggest a mechanism for nucleosome spacing where Acf1 senses linker DNA and communicates this information to Snf2h by regulating the accessibility of the H4 tail. On nucleosomes with short linker DNA, Acf1 binds the H4 tail, allowing AutoN to inhibit the ATPase activity of Snf2h. On substrates with longer linkers, Acf1 instead binds to the linker DNA, leaving the H4 tail free to displace AutoN from the ATPase and activate ACF. This work was published in reference [127].

The SWI/SNF chromatin-remodeling enzymes are important regulators of chromatin structure in transcriptional activation and in the creation and maintenance of nucleosome free regions, but the mechanisms by which these enzymes couple the free energy of nucleotide hydrolysis to the repositioning and rectructuring of nucleosomes remains incompletely under-

stood. In Chapter 4, we use single molecule FRET to monitor the remodeling of individual nucleosomes by the yeast SWI/SNF remodeler, RSC. The observed remodeling activity primarily involves translocation of DNA across the nucleosome without large-amplitude lifting of the DNA away from the nucleosome or substantial displacement of the H2A-H2B dimer. By monitoring nucleosome remodeling under a variety of different labeling schemes, we find that RSC translocates DNA across the nucleosome in a stepwise manner with a step size of approximately 1-2 bp. These results suggest that the movement of DNA across the nucleosome is coupled directly to DNA translocation by the ATPase.

Chapter 2

Initiation complex dynamics direct the transitions between distinct phases of early HIV reverse transcription

2.1 Introduction

As a key step in the life cycle of the human immunodefficiency virus (HIV), reverse transcription converts the single-stranded viral RNA (vRNA) genome into an integration-competent double-stranded DNA [30]. This multi-step reaction is catalyzed by the viral enzyme reverse transcriptase (RT), which initiates DNA synthesis from the 3' terminus of a cell-derived $tRNA^{Lys,3}$ annealed to an 18-nucleotide (nt) primer binding site (PBS) in the vRNA [128]. As a prerequisite of reverse transcription, the tRNA primer, vRNA template and RT must assemble into a productive ternary ribonucleoprotein complex, called the initiation complex [129, 130, 131]. Various chemical and enzymatic probing assays have revealed extensive intermolecular interactions within this complex, which are important for the efficiency and specificity of the initiation of minus-strand DNA synthesis [53, 54, 55, 56, 57].

The early stages of reverse transcription comprise a slow, distributive initiation phase followed by a fast, processive elongation phase [132, 133]. The initiation phase is characterized by slow DNA polymerization, rapid dissociation of RT, and frequent kinetic pauses during primer extension. These pausing events result in the accumulation of short extension products, such as the tRNA+3 and tRNA+5 intermediates. Interestingly, after addition of

the 6th nucleotide to the tRNA primer, reverse transcription transitions to the elongation phase with dramatically increased polymerization rate and processivity [134]. The unique and complex nature of the initiation process makes it an attractive target for antiviral drugs. Indeed, it has been shown that some RT inhibitors and drug-resistant RT mutants have different effects on initiation versus elongation [56]. Elucidating the mechanism underlying the initiation process could therefore set the stage for developing novel inhibitors for combating HIV infection. However, a consensus on the structure of the initiation complex has yet to be reached [54, 135, 136]. Furthermore, the dynamics of the complex and the mechanism underlying the transitions between various phases of early reverse transcription remain incompletely understood [134, 137, 138, 139, 140].

We recently demonstrated that RT displays remarkable orientational and translational dynamics on its nucleic acid substrates [24, 25]. Such large-scale motions facilitate various stages of reverse transcription, including DNA synthesis, RNase H cleavage, and strand-displacement synthesis. In particular, RT can bind its substrates in two opposite orientations [24, 141]. In one orientation, the DNA polymerase active site is positioned over the 3' end of the primer, ready to catalyze primer extension. In the other, RT is flipped ~180 degrees such that the DNA polymerase active site is physically separated from the primer 3' end, and thereby unable to support synthesis. Here, we investigated whether such orientational dynamics regulate the initiation of reverse transcription.

To test this hypothesis, we applied single-molecule fluorescence resonance energy transfer (FRET) [15, 16] and ensemble primer-extension assays to probe the conformation and enzymatic activity of the initiation complex. Interestingly, we found that this complex supports both the polymerase-competent and the flipped RT binding orientations. Furthermore, the equilibrium between these two binding orientations evolves as RT elongates the tRNA primer, and changes in the primer extension activity correlate with changes in the enzyme's binding orientation. In particular, a stem-loop structure within the vRNA adjacent to the PBS forces RT to bind predominantly in the flipped, polymerization-incompetent orientation, causing

reverse transcription pauses. Disrupting this structure by strand-displacement synthesis allows RT to bind in the polymerase-competent mode and triggers the initiation-to-elongation transition.

2.2 Results

2.2.1 Single-molecule FRET assay for studying initiation complex dynamics

To facilitate single-molecule FRET measurements, we labeled specific sites of the initiation complex with FRET donor and acceptor dyes. First, we constructed a vRNA template using the sequence corresponding to nucleotides 106-245 of the HIV-1 genome (NL4.3 isolate), with the PBS located at nucleotides 183-200. It has been shown that the sequence within the initiation site of the NL4.3 isolate is representative of 86% known HIV-1 isolates [135] and that sequences outside this region do not affect the efficiency of reverse transcription [142]. A FRET acceptor dye (Cy5) was site-specifically attached to U208 near the 3' end of the PBS (Figure 2.2.1a). The natural $tRNA^{Lys,3}$ primer was then annealed to the PBS. Next, tRNA:vRNA complexes were anchored to a slide surface and immersed in a reaction buffer containing HIV-1 RT labeled with a FRET donor dye (Cy3) at its RNase H domain (Figure 2.2.1a). Dye labeling of RT and vRNA did not alter reverse transcription kinetics appreciably (Figure 2.2.2). Fluorescence emission from individual tRNA:vRNA:RT ternary complexes was monitored using a total-internal-reflection fluorescence microscope. Freely diffusing RT was observed to bind and dissociate from the tRNA:vRNA substrates in real time. Each binding event caused an increase in the total fluorescence signal due to excitation of the FRET donor dye (Figure 2.2.1b). FRET values recorded during the binding event allowed us to determine the binding configuration of the enzyme. In this scheme, dissociation of the enzyme would result in a loss of the total fluorescence signal while a FRET value change without loss of total fluorescence could be interpreted as a change of binding configuration.

In addition to the tRNA:vRNA substrates, we also probed the binding configuration of RT on substrates comprising a variety of primer and template structures, as shown in Figure 2.2.1c,d.

2.2.2 RT adopts two opposite orientations in initiation complexes

It has been shown that HIV-1 RT can bind to simple primer:template substrates in two opposite orientations [24, 141]: In the case of a DNA primer, RT binds almost exclusively in the polymerase-competent orientation; in the case of a random-sequence RNA primer annealed to a simple DNA template, RT binds in the flipped orientation that inhibits primer extension. It is thus interesting to ask which orientation RT would adopt on the substrate formed between two complex RNA structures (tRNA and vRNA), which supports primer extension despite its RNA primer composition. Should RT bind in the polymerase-competent orientation, we expect to observe a high FRET value due to the proximity of the FRET donor and acceptor dyes according to previous crystallographic and footprinting data [37, 49, 143]. In contrast, the flipped orientation should give a low FRET value of ~0.2-0.3 with the donor and acceptor separated by the 18-base-pair tRNA:PBS duplex (Figure 2.2.3a).

Notably, the binding of RT to the tRNA:vRNA complex gave two distinct FRET peaks centered at ~0.9 and ~0.2 (Figure 2.2.3b), suggesting that RT binds in both the polymerase-competent and flipped orientations on this substrate. To confirm the assignments of the two FRET states, we took advantage of the knowledge that cognate deoxyribonucleotides preferentially stabilize RT binding in the polymerase-competent orientation [24]. Indeed, when the FRET distribution was acquired in the presence of the cognate nucleotide (over a short duration to avoid substantial primer extension), the equilibrium shifted towards the high FRET state (Figure 2.2.3b), supporting the assignment of this state to the polymerase-competent binding mode. The two binding modes were also observed when the $tRNA^{Lys,3}$ primer was annealed to a simple RNA template containing the PBS sequence without any secondary structures (Figures 2.2.1c and 2.2.5), further supporting the notion that the two

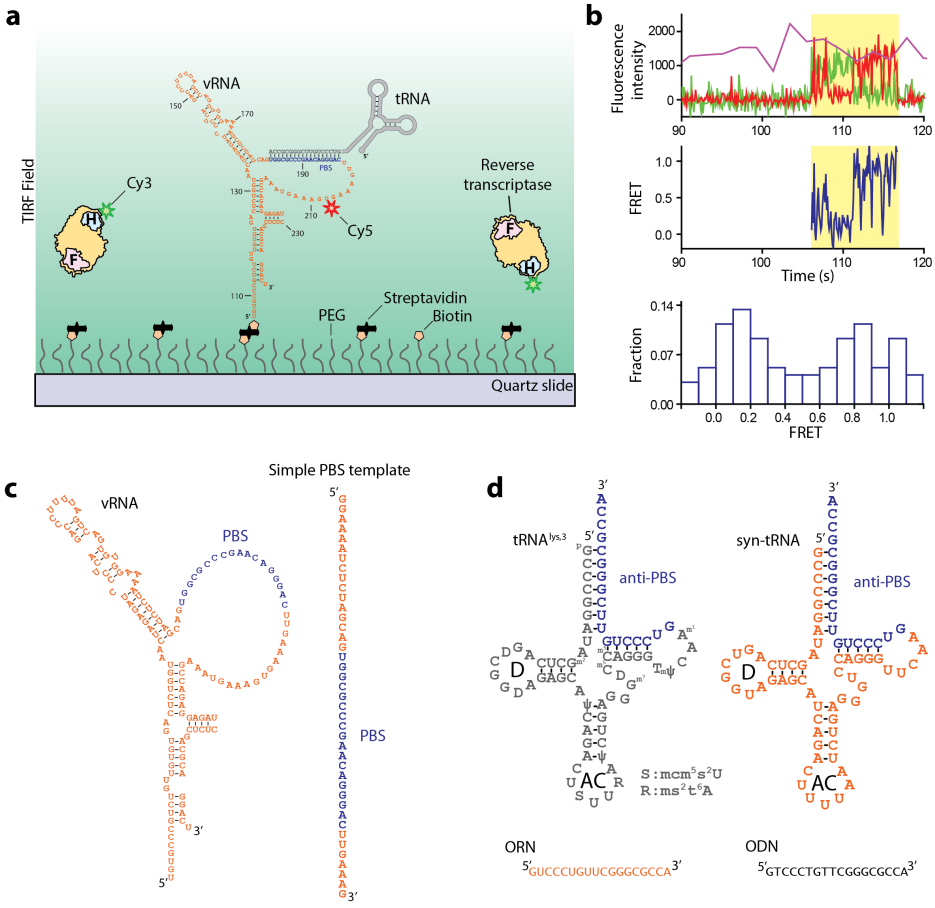


Figure 2.2.1: Single-molecule FRET assay for probing the structural dynamics of the initiation complex. **a**, The vRNA template (orange) is labeled with a FRET acceptor (Cy5, red star) near the PBS (blue), annealed to a tRNA primer (grey), and immobilized to the PEG-coated surface via a streptavidin-biotin linkage. The surface-anchored tRNA:vRNA substrates are immersed in a solution containing RT (yellow) labeled with the FRET donor (Cy3, green star). The fingers and RNase H domains of RT are indicated by F and H, respectively. Fluorescence signal from single tRNA:vRNA:RT complexes are detected using a TIRF microscope. **b**, FRET analysis of RT binding events. The upper panel shows the fluorescence signals from Cy3 (green) and Cy5 (red) under 532 nm illumination and the signal from Cy5 directly excited by 635 nm illumination (purple). Binding of RT to the substrate (highlighted by yellow) results in an increase in the total fluorescence signals from Cy3 and Cy5 under 532 nm illumination due to excitation of the FRET donor, but does not affect the Cy5 signal from direct excitation by 635 nm light. (Middle panel) FRET values during the binding event. (Lower panel) FRET histogram of the binding event. **c**, Sequences of the vRNA and simple RNA PBS templates studied in this work. **d**, Sequences of the natural tRNALys,3, synthetic tRNA (syn-tRNA), oligoribonucleotide (ORN) and oligodeoxyribonucleotide (ODN) primers. Adapted with permission from [126].

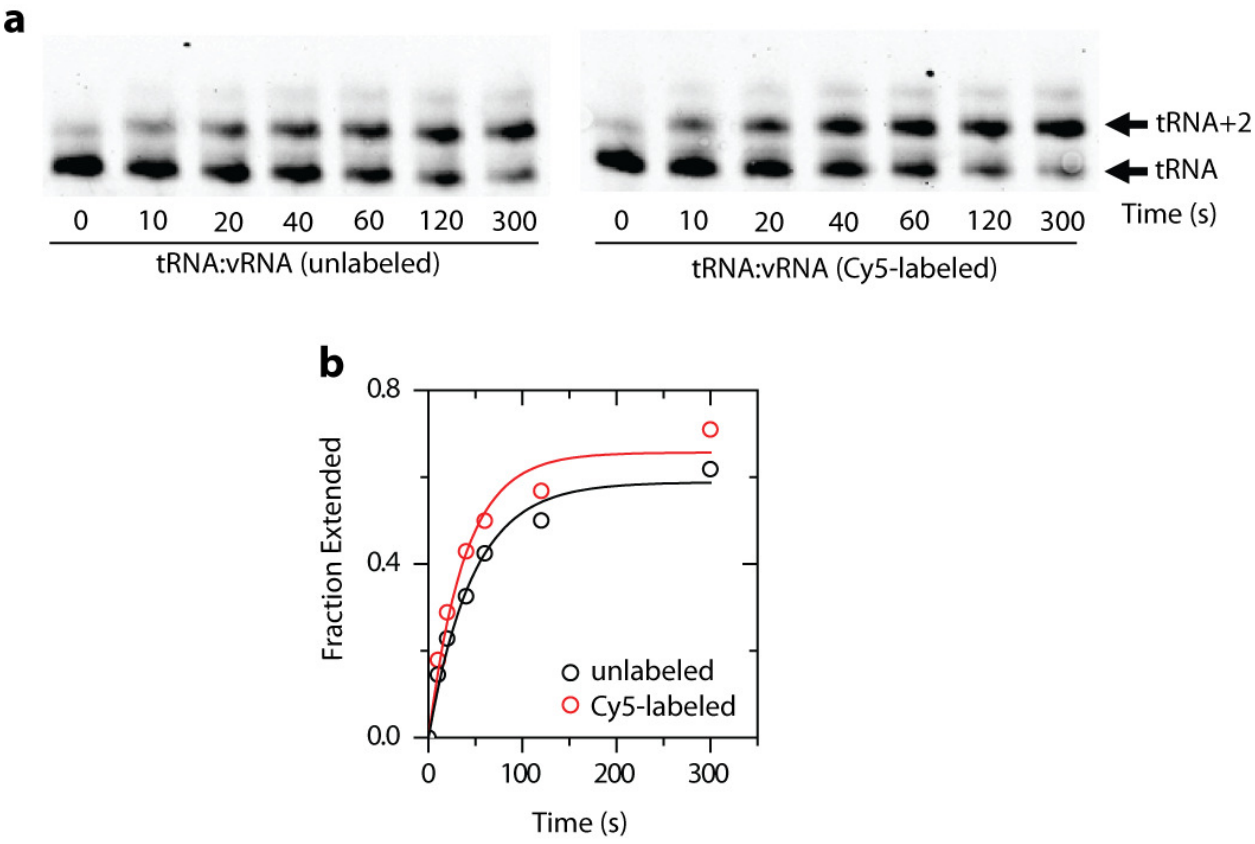


Figure 2.2.2: Dye labeling does not perturb the primer extension activity of the tRNA:vRNA substrate and the RT enzyme. A nucleotide incorporation assay in the presence of dCTP and dTTP is used to measure the kinetics of primer extension. **a**, Polyacrylamide gel showing the time course of extension of a tRNA primer annealed to an unlabeled vRNA template (left) or a Cy5-labeled vRNA (right). **b**, Fraction of the primer extended as a function of time. The extended primer fraction is determined as the ratio of the signal from the product band (tRNA+2) over the total signal from both the unextended (tRNA) and product (tRNA+2) bands. The lines are fitting of the data to a single exponential function, $f = A(1 - e^{-kt})$, which allows the determination of the rate constant for primer extension (k). The similar primer extension rate constants observed for the unlabeled vRNA template (0.021 s^{-1}) and the Cy5-labeled vRNA template (0.026 s^{-1}) indicate that dye-labeling of the substrate does not substantially alter the activity. We have previously shown that labeling RT at the C-terminus of the p66 subunit also does not affect its polymerase activity appreciably [24, 25]. Adapted with permission from [126].

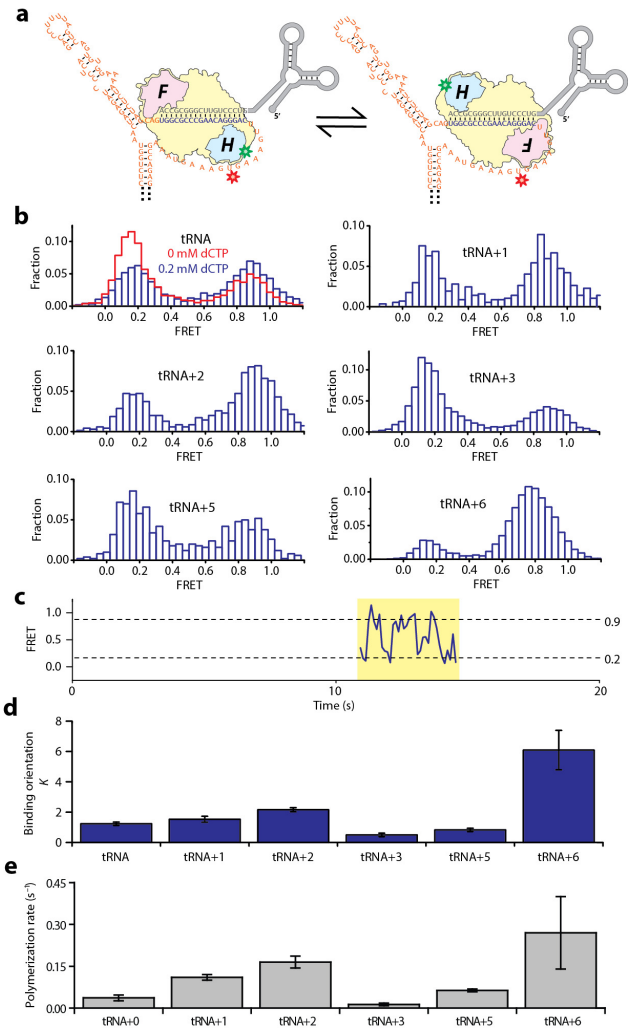
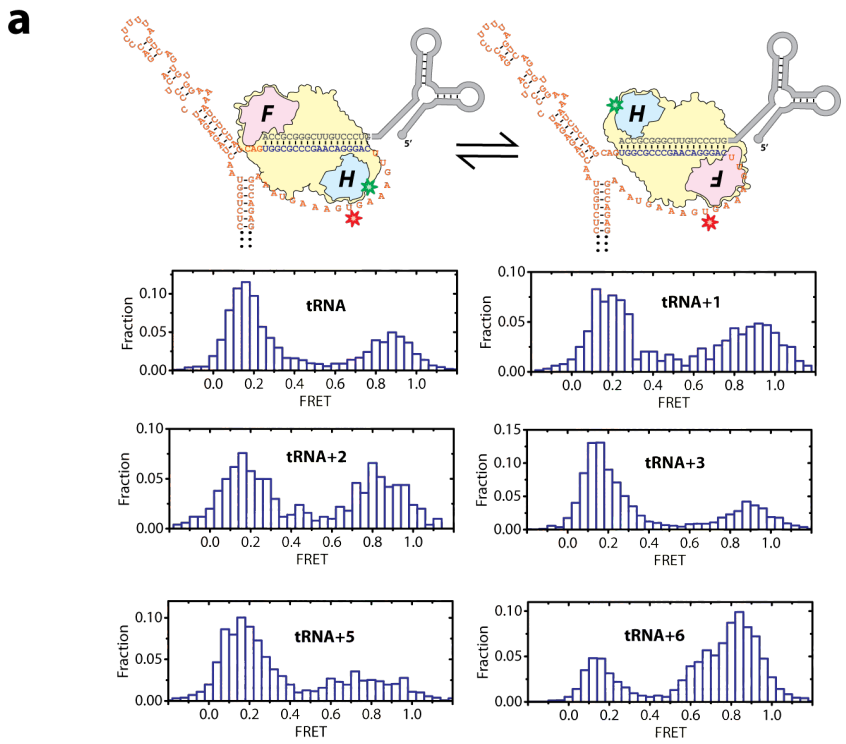


Figure 2.2.3: The RNA-dependent DNA polymerase activity of RT correlates with its binding orientation in the initiation complex. **a**, A cartoon illustrating RT bound to the tRNA:vRNA complex in the polymerase-competent (left) and flipped (right) orientations. **b**, The FRET distribution obtained when Cy3-labeled RT binds to Cy5-labeled tRNA+n:vRNA complexes in the presence of 200 µM cognate dNTP. In the case of tRNA ($n = 0$), the FRET distribution in the absence of dNTP (red) is also shown. The FRET distributions for other n values in the absence of dNTP are shown in Figure 2.2.4. **c**, A representative time trace of an RT binding event (highlighted in yellow) shows spontaneous transitions between the high and low FRET states. **d**, The equilibrium constants K between the polymerase-competent and the flipped binding orientations of RT in the presence of 200 µM cognate dNTP. Error bars are s.e.m. from at least 3 independent experiments. **e**, Primer extension rates on the vRNA template. The primer extension rate is highly correlated with the binding orientation equilibrium K, exhibiting a correlation coefficient of 0.94. Error bars are s.d. from at least 3 independent experiments. Adapted with permission from [126].



b

	$k_{low \rightarrow high}$	$k_{high \rightarrow low}$	K_{eq}
tRNA	0.50±0.08 s^{-1}	0.82±0.11 s^{-1}	0.54±0.06
tRNA+1	0.58±0.12 s^{-1}	0.57±0.14 s^{-1}	0.95±0.12
tRNA+2	0.57±0.17 s^{-1}	0.46±0.14 s^{-1}	0.98±0.18
tRNA+3	0.47±0.11 s^{-1}	1.0±0.3 s^{-1}	0.42±0.05
tRNA+5	0.34±0.16 s^{-1}	0.51±0.26 s^{-1}	0.60±0.15
tRNA+6	0.60±0.14 s^{-1}	0.18±0.05 s^{-1}	3.3±0.7

Figure 2.2.4: The flipping equilibrium and kinetics of RT on the tRNA+n:vRNA substrate in the absence of dNTP. **a**, FRET distributions obtained from Cy3-labeled RT bound to Cy5-labeled tRNA+n:vRNA complexes in the absence of dNTP. **b**, Rate constants for the transitions between the high FRET polymerase-competent and low FRET flipped binding modes as well as the equilibrium constants for the transitions in the absence of dNTP. The equilibrium constants, derived from the areas of the high and low FRET peaks in the histograms, match well with the ratios of the rate constants, which are derived from analyzing the dwell times in the single-molecule traces. The uncertainties are s.e.m. from at least 3 independent experiments. Adapted with permission from [126].

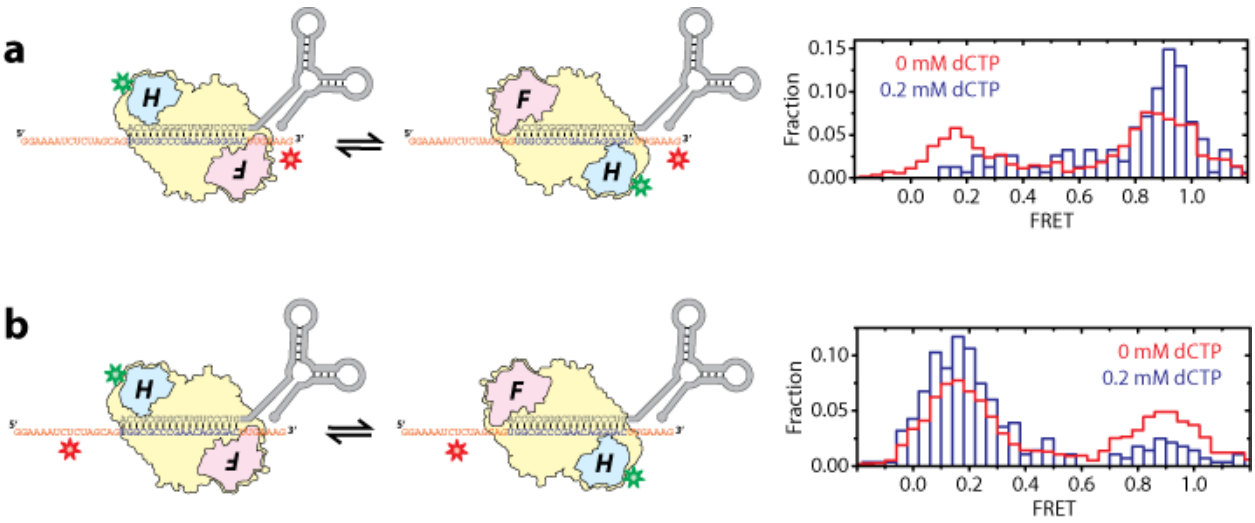


Figure 2.2.5: RT binds to tRNA:simple-PBS complexes in two orientations. The FRET distributions are obtained from the binding of Cy3-labeled RT to tRNA primers annealed to Cy5-labeled simple-PBS templates in the absence (red) and presence (blue) of 200 µM dCTP. Two different labeling schemes of the simple-PBS template are used, one with Cy5 attached to the 3' terminus (a) and one with Cy5 attached near the 5' end, at the nucleotide corresponding to U175 in the vRNA (b). Adapted with permission from [126].

FRET states originate from different binding configurations of RT on the tRNA:PBS duplex rather than binding to regions of vRNA outside the PBS. Moreover, the FRET time trace showed spontaneous transitions between the ~0.9 and ~0.2 FRET states within individual binding events (Figures 2.2.3c), indicating that RT could dynamically transition between the two orientations without dissociation.

We next studied the influence of tRNA structure on RT binding orientation by testing substrates formed by various other primers (Figure 2.2.1d) and the vRNA template. First, we replaced the natural $tRNA^{Lys,3}$ with a synthetic, unmodified tRNA (syn-tRNA) to test the effect of the modified tRNA bases. The FRET distribution observed for the syn-tRNA:vRNA:RT complexes was similar to that of $tRNA^{Lys,3}$:vRNA:RT (Figure 2.2.6a,b). The equilibrium constant (K) between the high-FRET and low-FRET orientations was 0.54 ± 0.06 for natural $tRNA^{Lys,3}$ and 0.64 ± 0.10 for syn-tRNA, respectively. Next, an 18-nt oligoribonucleotide (ORN) primer complementary to the PBS was used in place of $tRNA^{Lys,3}$. In this case, RT can still transition between the two orientations, the equilibrium shifting to-

wards the high-FRET orientation with $K = 1.8 \pm 0.2$ (Figure 2.2.6c). It has been suggested that the anti-codon loop of the tRNA base-pairs with the A-rich loop upstream of the PBS, although results on whether such interactions exist in the NL4.3 isolate vary [135, 136, 144]. Since such interactions would not be present in the case of the ORN primer, our results indicate that the ability of RT to adopt and flip between the two orientations does not require these interactions outside the PBS. Finally, when an 18-nt oligodeoxyribonucleotide (ODN) primer was annealed to the vRNA template, RT nearly exclusively bound in the high-FRET, polymerase-competent orientation (Figure 2.2.6d). In addition, the average binding time of RT was substantially longer on the ODN primer (40 sec) than on the ORN (2.0 sec) and tRNA (1.6 sec) primers. These results indicate that both the nucleotide nature (deoxyribonucleotide vs. ribonucleotide) and the secondary structure of the primer play a role in determining RT's binding orientation.

2.2.3 Polymerase activity correlates with RT binding orientation

To investigate how RT's binding orientation changes during the initiation of minus-strand DNA synthesis, we constructed a series of tRNA-DNA chimeras, tRNA+n, representing intermediates encountered at varying primer extension steps (n denotes the number of deoxyribonucleotides added to the tRNA 3' terminus) (Figure 2.2.7). Cy3-labeled RT was then added to the Cy5-labeled tRNA+n:vRNA substrates. All tRNA+n:vRNA:RT initiation complexes ($n = 0 - 6$) supported both high and low FRET binding orientations (Figures 2.2.3b and 2.2.4a). The equilibrium constant K between the two FRET states increased when n increased from 0 to 2 (Figures 2.2.3b,d and 2.2.4). Remarkably, when tRNA was further extended by one extra nucleotide (i.e. tRNA+3), RT binding shifted substantially towards the flipped orientation, which remained dominant until position tRNA+5 (Fig. 2.2.3b,d and 2.2.4). Finally, when the 6th nucleotide was added to the tRNA primer, another major shift occurred, resulting in a predominant polymerase-competent binding orientation for the enzyme (Figures 2.2.3b,d and 2.2.4). Analysis of the flipping rate constants indicates that the

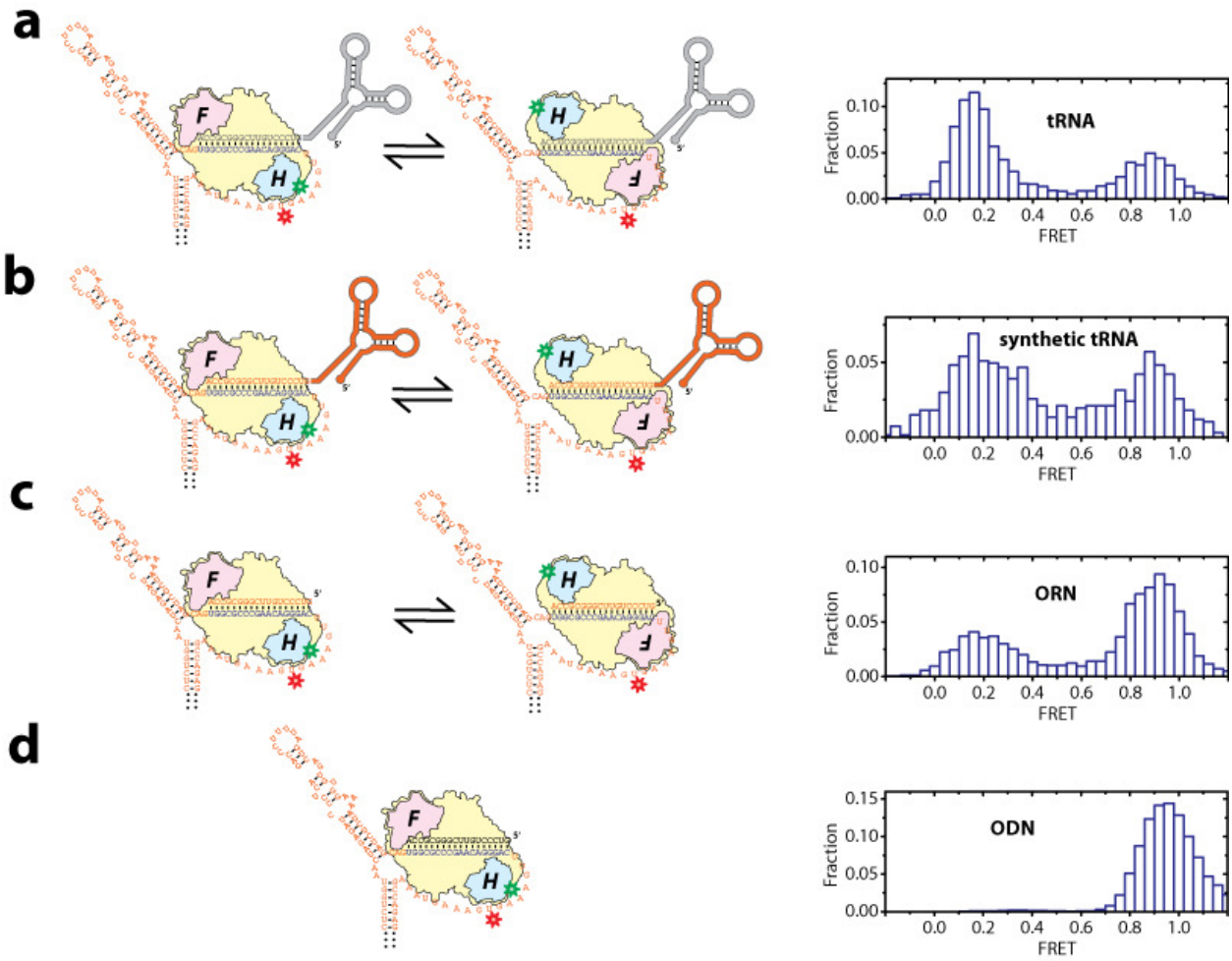


Figure 2.2.6: Effect of primer structure on RT binding orientation. FRET histograms from the binding of Cy3-labeled RT to (a) natural $tRNA^{Lys,3}$, (b) synthetic, unmodified tRNA (syn-tRNA), (c) oligoribonucleotide (ORN), and (d) oligodeoxyribonucleotide (ODN) primers annealed to a Cy5-labeled vRNA template. Adapted with permission from [126].

change in the flipping equilibrium is modulated primarily by the rate at which RT transitions from the polymerase-competent to the flipped orientation whereas the reverse rate does not vary substantially with the length of the DNA extension (Figure 2.2.4b).

Intriguingly, the positions at which the binding orientation equilibrium undergoes major changes (tRNA+3 and tRNA+6) correspond to major transitions in primer extension kinetics: polymerization exhibits a strong pause at tRNA+3 and a transition from the slow initiation phase to the rapid elongation phase at tRNA+6 [132, 134]. These correlations raise the possibility that the DNA synthesis activity of RT during initiation may be regulated by its binding orientation. To test this notion, we performed single-turnover nucleotide incorporation assays to determine the primer extension activity of RT on all reaction intermediates used in the single-molecule FRET experiments. Indeed, the primer-extension rate was closely correlated with the equilibrium between the two binding orientations (Figure 2.2.3d,e). In particular, RT exhibited the lowest primer extension activity on the tRNA+3 primer, where the enzyme preferentially bound in the flipped orientation, and the fastest extension kinetics on tRNA+6, where the enzyme predominantly bound in the polymerase-competent orientation. These observations suggest that the polymerase activity of the enzyme is at least in part regulated by its binding orientation.

2.2.4 The stem-loop near the PBS causes the early pausing

Next we investigated the origin of the pauses during early reverse transcription, namely the structural features of the initiation complex responsible for the flipped RT binding orientation and slow primer extension rate. The most prominent pause during initiation is at position +3, for which two possible scenarios may be hypothesized: (i) the 3-nt DNA addition to the tRNA primer may confer on it a unique structure (or primer:PBS duplex structure) that disfavors RT binding in the polymerase-competent orientation; (ii) alternatively, the stem-loop structure in the vRNA template 3-nt upstream of the PBS [135, 137] may force RT to bind in the flipped orientation and decrease DNA synthesis activity. To distinguish between

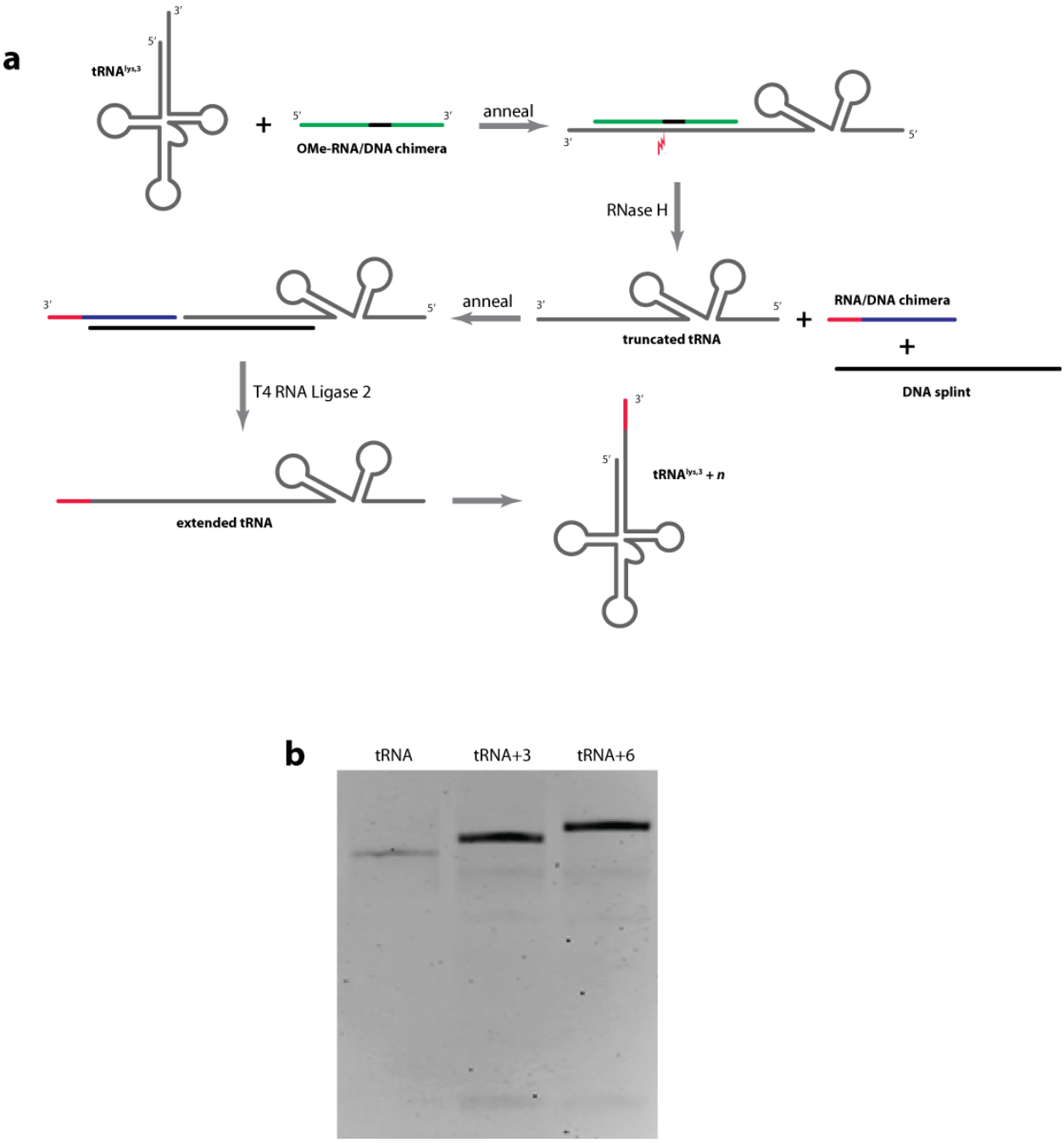


Figure 2.2.7: Scheme for the construction of tRNA+n primers. **a**, Natural $tRNA^{Lys,3}$ is annealed to a 2'O-methyl RNA/DNA chimera which directs E. coli RNase H to cleave between C62 and U63 of the tRNA. A splinted ligation reaction catalyzed by T4 RNA ligase 2 is used to ligate a synthetic RNA/DNA chimeric oligonucleotide onto the end of the truncated tRNA. **b**, An ethidium bromide-stained polyacrylamide gel showing the migration of the tRNA+3 and tRNA+6 constructs relative to the unextended tRNA. Adapted with permission from [126].

these possibilities, we compared the binding configuration and primer extension kinetics of RT on the simple PBS template (Figure 2.2.1c) to those on the vRNA template. While RT preferentially bound in the low FRET state on the tRNA+3:vRNA substrate, it strongly favored the high FRET, polymerase-competent orientation on the tRNA+3:PBS substrate (Figure 2.2.8a). Moreover, the substantial drop of primer extension rate from the tRNA+2 to tRNA+3 position observed on the vRNA template was not reproduced on the simple PBS template (Figure 2.2.8a). As a result, unlike in the case of the vRNA template, no substantial pause was detected at position +3 on the simple PBS template (Figure 2.2.9a, b). These results indicate that the strong +3 pause on the vRNA template is not due to an intrinsic structural property of the tRNA+3 primer, but rather originates from the template structure.

To test the role of the vRNA stem-loop structure in regulating RT's binding dynamics, we created a mutant, $vRNA_{h-}$, by altering the sequence at positions 135-142 such that the 8 base pairs at the base of the stem-loop could not form (Figure 2.2.8b). When RT was added to the tRNA+3:$vRNA_{h-}$ substrate, the FRET distribution showed a single peak at ~0.9 (Figure 2.2.8b), indicating binding predominantly in the polymerase-competent orientation. This result differs drastically from the behavior of RT on the wild type vRNA template but mimics its interaction with the simple PBS template. Moreover, the single-nucleotide incorporation assay showed that eliminating the stem-loop structure in the $vRNA_{h-}$ template increased the primer extension rate at the +3 position by ~8 fold compared to the wild type vRNA (Figure 2.2.8b), consistent with a previous report that altering vRNA sequences upstream of the PBS can eliminate the +3 pause site [137]. Overall, the above observations demonstrate that the secondary structure of the vRNA template strongly influences the binding configuration of RT: the stem-loop structure upstream of the PBS reorients RT into a flipped, polymerase-incompetent orientation and induces pausing at position +3.

Conflicting results have been reported on whether the same +3 pause occurs when a DNA primer is used [132, 137]. Interestingly, we found that the ability of the stem-loop to reorient

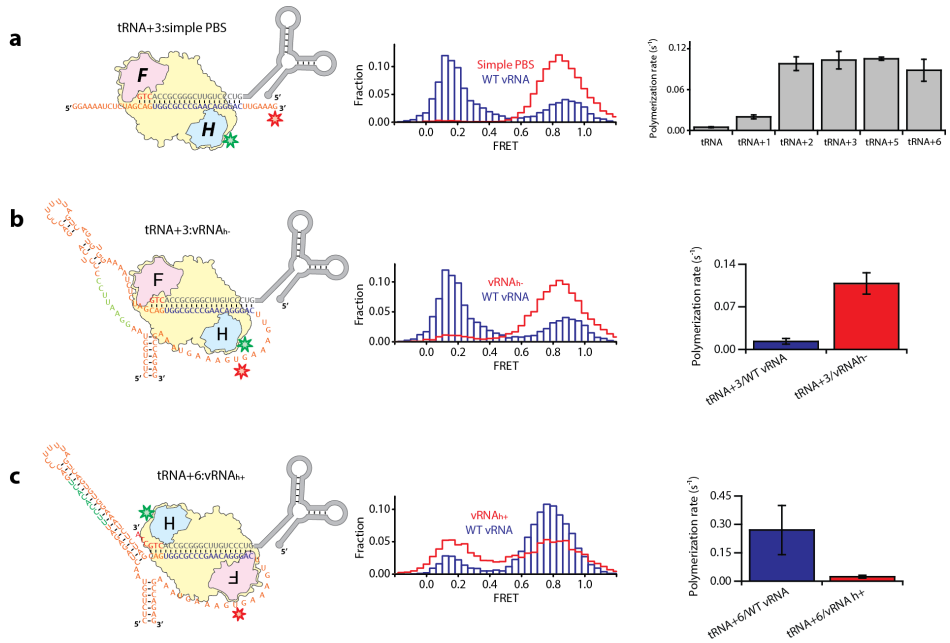


Figure 2.2.8: The stem-loop structure upstream of the PBS causes the major pauses during initiation and governs the initiation-to-elongation transition. **a**, Left panel: Cartoon of RT bound to tRNA+3:simple PBS substrate. Middle Panel: FRET histograms for RT bound to tRNA+3 primers annealed to the vRNA template (blue) and simple PBS template (red). Right panel: Rates of single nucleotide addition to various tRNA primers on the simple PBS template. **b**, Left panel: Cartoon of RT bound to tRNA+3:vRNA$_{h-}$ substrate. Middle Panel: FRET histograms for RT bound to tRNA+3 primers annealed to wild type (wt) vRNA (blue) and vRNA$_{h-}$ (red) templates. Right panel: Rates of single nucleotide addition to tRNA+3 primers annealed to wt vRNA (blue) and vRNA$_{h-}$ (red) templates. **c**, Left panel: Cartoon of RT bound to tRNA+6:vRNA$_{h+}$ substrate. Middle Panel: FRET histograms for RT bound to tRNA+6 primers annealed to wt vRNA (blue) and vRNA$_{h+}$ (red) templates. Right panel: Rates of single nucleotide addition to tRNA+6 primers annealed to wt vRNA (blue) and vRNA$_{h+}$ (red) templates. The nucleotides mutated to create the vRNA$_{h-}$ and vRNA$_{h+}$ constructs are shown in green. Error bars are s.d. from at least 3 independent experiments. Adapted with permission from [126].

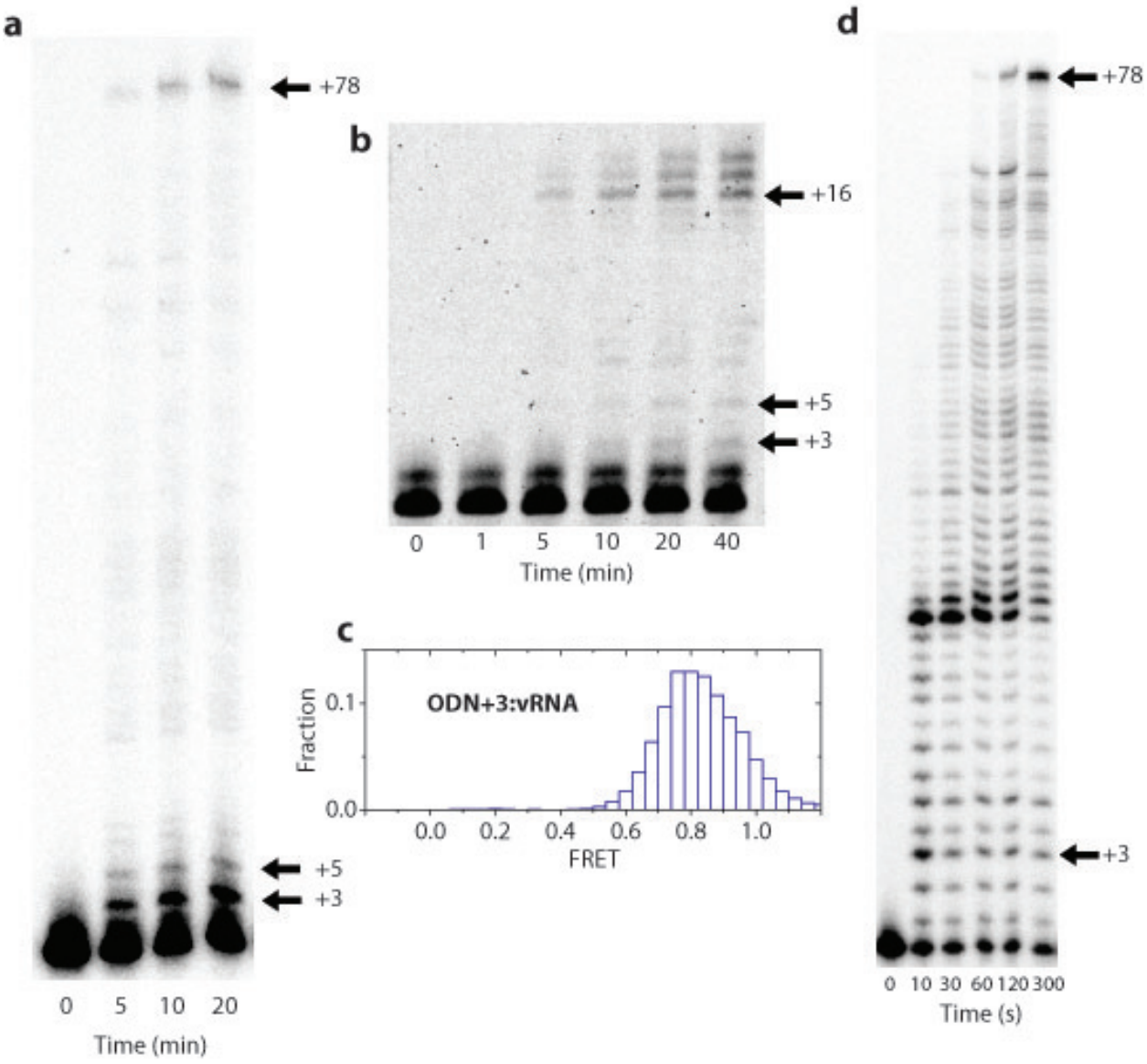


Figure 2.2.9: Polymerization pauses (or lack thereof) of the tRNA:vRNA, tRNA:simple-PBS and ODN:vRNA substrates. **a**, Polyacrylamide gel showing the time course for the extension of a tRNA primer annealed to the vRNA template. All four types of dNTP are present to allow full-length extension. **b**, Time course for the extension of a tRNA primer annealed to the simple-PBS template. Extension beyond position +16 (the 5' end of the simple PBS template) is due to RT's ability to perform blunt-end nucleotide addition [145]. **c**, The FRET histogram from Cy3-labeled RT bound to an ODN+3 primer annealed to the Cy5-labeled vRNA template. **d**, Time course for the extension of an ODN primer annealed to the vRNA template. Adapted with permission from [126].

RT vanished when the tRNA was replaced by an ODN primer. RT bound predominantly in the polymerization orientation on an ODN+3:vRNA substrate despite the presence of the stem-loop structure (Figure 2.2.9c). Furthermore, RT exhibited much weaker pausing at the +3 position when reverse transcription was initiated from the ODN primer (Figure 2.2.9d). This observation rules out the possibility that the strong +3 pause observed with the tRNA primer is due to the intrinsically slower rate of strand-displacement synthesis [146] as the enzyme passes through the vRNA stem-loop, which should occur for both tRNA and ODN primers, and further supports the notion that the flipped RT orientation on the tRNA:vRNA substrate imposed by the stem-loop was a major determinant of pausing.

2.2.5 Disruption of stem-loop triggers transition to elongation

Because the stem-loop structure induces strong pausing and consequently the slow, distributive initiation phase of DNA synthesis, we hypothesize that disassembly of the stem-loop structure may cause the transition to the elongation phase that occurs between tRNA+5 and tRNA+6. Supporting this hypothesis, such a transition was not observed on the simple PBS template lacking the stem-loop structure, on which the tRNA+6 and tRNA+5 primers displayed similar extension kinetics (Figure 2.2.8a).

If unfolding of the stem-loop structure mediates the initiation-to-elongation transition, we anticipate that preventing hairpin disassembly would inhibit entry into the elongation phase. We therefore created another mutant vRNA template containing a hairpin structure lacking bulges ($vRNA_{h+}$), which should be more stable and difficult to disrupt. Indeed, this modification shifted the binding configuration of RT toward the low FRET orientation (Figure 2.2.8c). In addition, the tRNA+6:$vRNA_{h+}$ substrate displayed a primer extension rate that was substantially slower as compared to the wild type tRNA+6:vRNA case but was more similar to those of the initiation phase (Figure 2.2.8c). These results suggest that entry into the elongation phase requires disassembly of the stem-loop structure.

To further test this notion, we directly probed the stem-loop structure by placing Cy3 and

Cy5 on U132 and U177 of the vRNA, respectively (Figure 2.2.10a). A fully formed hairpin would bring the two dyes into proximity and generate a high FRET value, whereas an opened hairpin would force them apart and yield a low FRET value. The doubly-labeled vRNA template was then annealed to various tRNA+n primers and incubated with unlabeled RT. In the case of an unextended tRNA primer, a major FRET peak at 0.95 was observed (Figure 2.2.10b), reflecting a folded stem-loop structure. Similar FRET distributions were observed for the tRNA+3:vRNA and tRNA+5:vRNA substrates (Figure 2.2.10c,d), indicating that the stem-loop was still largely intact after incorporation of up to five deoxyribonucleotides. In contrast, in the presence of tRNA+6, the majority of complexes exhibited a low FRET value ~0.15 (Figure 2.2.10e), indicating opening of the hairpin structure. Presumably, in the tRNA+6:vRNA:RT complex, competition from the terminal three nucleotides of tRNA+6 destabilizes the stem-loop and, together with the bulges in the upper half of the stem-loop, makes this structure relatively unstable. Unfolding of the stem-loop structure explains why RT can engage this substrate predominantly in the polymerase-competent mode (Figure 2.2.3b,d) and synthesize DNA efficiently (Figure 2.2.3e).

2.2.6 Role of NC in the initiation of DNA synthesis

The viral nucleocapsid (NC) protein has been reported to facilitate various stages of HIV replication, including reverse transcription [147, 148]. NC is a small basic protein that non-specifically binds to nucleic acids and possesses nucleic-acid chaperon activity [147, 149]. In particular, it has been suggested that NC could disrupt template secondary structures at which RT pauses during the elongation phase [150, 151, 152]. It is thus interesting to ask whether NC could disrupt the stem-loop structure upstream of the PBS and increase the primer extension rate at the +3 position. Addition of NC to the doubly-labeled tRNA:vRNA complex broadened the FRET distribution considerably (Figure 2.2.11a). FRET traces of individual initiation complexes revealed fluctuations among a wide range of FRET values (Figure 2.2.11a), suggesting NC-mediated destabilization of the stem-loop. A high-

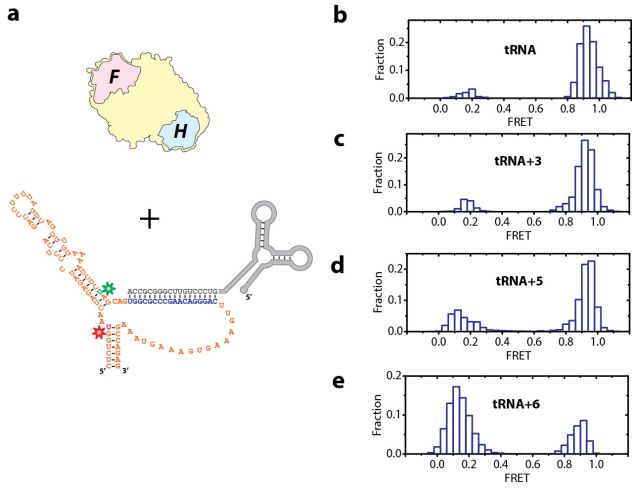


Figure 2.2.10: Disruption of the stem-loop structure upstream of the PBS occurs upon addition of the 6th nucleotide to the tRNA primer. **a**, Diagram of the doubly-labeled construct used to monitor the folding of the stem-loop structure. Cy3 (green star) and Cy5 (red star) are attached to vRNA positions U132 and U177. **b-e**, The FRET distributions obtained when 100 nM RT was added to the doubly-labeled template annealed to the tRNA (b), tRNA+3 (c), tRNA+5 (d), and tRNA+6 (e) primers. Adapted with permission from [126].

throughput footprinting study also shows that the stem-loop upstream of the PBS is at least partially open in the virus particle while disruption of the NC-vRNA interactions restores base-pairing [136]. With the stem-loop structure disrupted, we anticipate that RT would bind primarily in the polymerase-competent orientation. Indeed, when added to the Cy5-labeled tRNA+3:vRNA substrate in the presence of NC, Cy3-labeled RT bound predominantly in the high FRET, polymerase-competent orientation, in contrast to the case without NC (compare Figure 2.2.11b with the tRNA+3 panel in Figure 2.2.3b). Intriguingly, despite the disruption of the stem-loop, the addition of NC did not result in appreciable increase in the extension rate of the tRNA+3 primer (Figure 2.2.11c), consistent with previous observations [137, 140]. Considering that NC binding may reduce binding of RT to the primer:template complex [151] and inhibit DNA synthesis due to steric hindrance [153], these negative effects could counter-balance its effect on disrupting the stem-loop structure. Indeed, we observed a substantial reduction in the binding frequency of RT to the tRNA+3:vRNA substrate in the presence of NC in single-molecule FRET traces (not shown), and NC has been previously

found to have both positive and negative effects on reverse transcription depending on the experimental conditions [150, 151, 153, 154, 155]. It is worth noting that NC is present at an extremely high concentration in the virions and such a condition is difficult to simulate *in vitro* due to protein aggregation. Future experiments that probe the virus directly are required to further elucidate whether the initiation pauses are alleviated by NC.

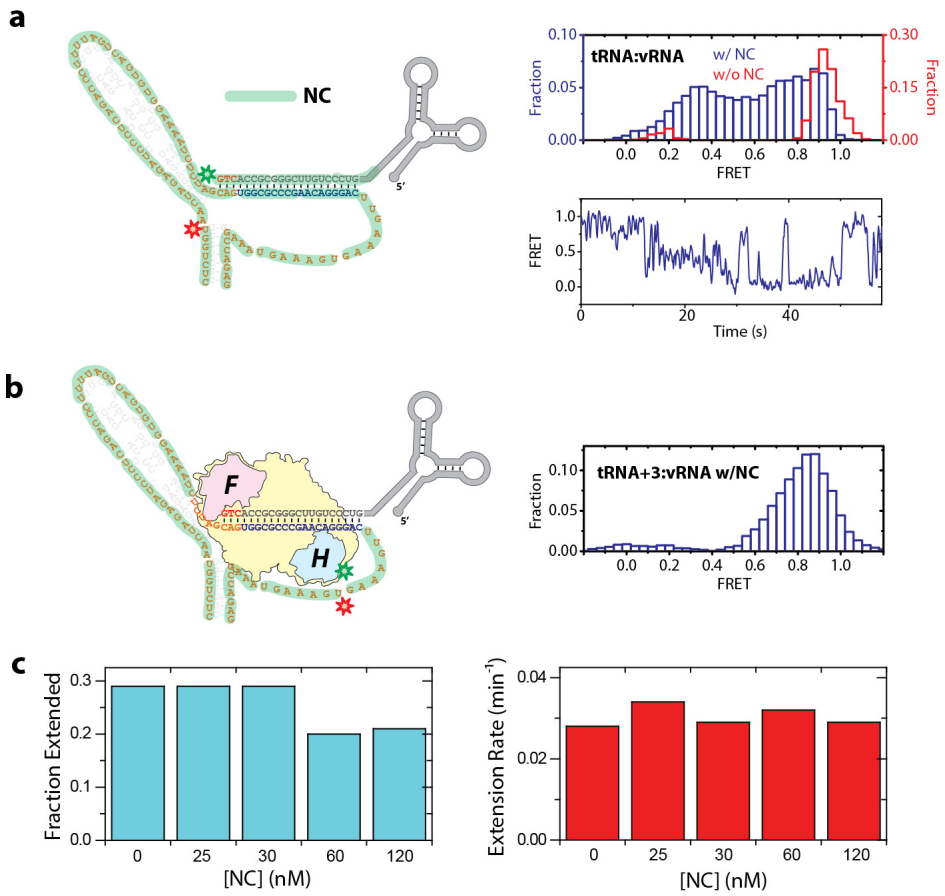


Figure 2.2.11: HIV-1 NC destabilizes the stem-loop structure. **a**, Left: Cartoon of the doubly-labeled tRNA:vRNA construct complexed with NC (green). Upper right: FRET distributions obtained from the doubly labeled tRNA:vRNA substrate in the absence (red) or presence (blue) of 1 µM NC. Lower right: a representative FRET time trace in the presence of 1 µM NC. **b**, FRET distribution obtained for Cy3-labeled RT bound to Cy5-labeled vRNA:tRNA+3 complexes in the presence of 10 nM NC. **c**, Single-nucleotide extension kinetics of tRNA+3:vRNA complex in the presence of 20 nM RT and various concentrations of NC. Left: The fraction of tRNA primers extended by a single nucleotide. Right: Primer extension rate constants. Adapted with permission from [126].

2.3 Discussion

Reverse transcription of the HIV genome is initiated from a cellular tRNA primer bound to the viral RNA. The specific structure and dynamics of the tRNA:vRNA complex result in an initiation phase with properties markedly different from those of later stages of reverse transcription, providing unique opportunities for anti-HIV drug design. It is thus important to understand the correlation between the structural dynamics of the initiation complex and their functional consequences.

By monitoring the dynamics of individual initiation complexes in real time, we investigated in detail how HIV-1 RT interacts with tRNA:vRNA substrates. Our results reveal that RT adopts two orientations in the initiation complex—a polymerase-competent orientation and a flipped orientation in which the polymerase active site is placed many nucleotides from the 3' end of the tRNA, precluding primer extension. The equilibrium between these two orientations regulates RT activity throughout the initiation process (Figure 2.3.1). The relatively slow polymerization rates during initiation result in part from the fact that RT spends a large portion of time bound to the tRNA:vRNA substrate in the flipped, polymerase-incompetent orientation (Figure 2.3.1), though the nucleotide nature of the primer (RNA vs. DNA) and the secondary structure of the template, which determines whether RT has to undergo strand-displacement synthesis, could also affect polymerization kinetics directly [134, 137, 138]. Interestingly, the ability to support RT binding in the flipped, polymerase-inactive orientation is a unique property of the RNA primer. When the tRNA primer is replaced with a DNA primer complementary to the PBS sequence, RT binds to the primer:template complex nearly exclusively in the polymerase-competent orientation and a slow initiation phase is not observed (Figure 2.2.9).

Addition of the first few nucleotides to the tRNA primer leads to a shift of the RT binding equilibrium towards the polymerase-competent orientation and an increase in the DNA synthesis rate (Figure 2.3.1). This rate then drops dramatically at position +3, where RT encounters a stem-loop structure on the vRNA template that forces the enzyme to bind

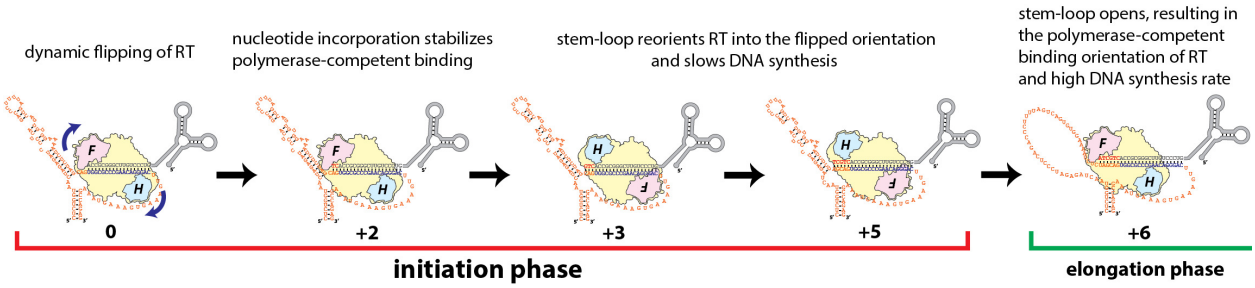


Figure 2.3.1: Structural dynamics of the HIV-1 initiation complex regulate the early phases of reverse transcription. RT binds to the initiation complex in two orientations — a polymerase-competent orientation and a flipped, polymerase-inactive orientation. RT spends a large portion of time bound to the tRNA:vRNA substrate in the flipped orientation. Addition of the first couple of deoxyribonucleotides to the tRNA primer shifts the RT binding equilibrium towards the polymerase-competent orientation and increases the DNA synthesis rate. The synthesis rate drops dramatically at position +3, where RT encounters a stem-loop structure in the vRNA template that forces the enzyme to bind predominantly in the flipped orientation, thereby increasing the probability of pausing. Strand-displacement synthesis until the +6 position eventually leads to unfolding of the stem-loop, which allows RT to reorient into the polymerase-competent binding mode and enter the fast, processive elongation phase of DNA synthesis. Adapted with permission from [126].

predominantly in the flipped, polymerase-inactive orientation, increasing the probability of pausing (Figure 2.3.1). RT cannot escape this slow mode of synthesis until it has synthesized enough DNA to disrupt the stem-loop at position +6. Disassembly of the stem-loop allows RT to reorient into the polymerase-competent binding mode and triggers transition to the fast, processive elongation phase of DNA synthesis (Figure 2.3.1). The single-stranded bulges within the vRNA stem-loop appear to promote the initiation-to-elongation transition (Figure 2.2.8c). Taken together, the above results provide a structural basis for understanding the early stages of reverse transcription and demonstrate how the structural dynamics of the initiation ribonucleoprotein complex regulate RT activity.

Considering the critical role played by the stem-loop in regulating RT binding orientation and activity, it is interesting to speculate that HIV may have evolved this structure to act as a temporal brake to slow initiation of reverse transcription. Such a mechanism may help prevent reverse transcription of the RNA genome prior to virus budding [156, 157], which is detrimental to viral infectivity [148, 158]. This speculation is supported by the observations

that both mutations of the viral NC protein that cause premature reverse transcription and mutations near the vRNA PBS that increase the initiation efficiency result in virus replication defects [159, 160, 161]. Conveniently, the NC protein has an ability to unwind the stem-loop (Figure 2.2.11a) and evidence suggests that this stem-loop is partially disrupted within the virus particle [136]. Therefore, this mechanism to inhibit reverse transcription in the initiation phase may be partially alleviated in a matured virion. On the other hand, the precursor form of NC within the Gag polyprotein (the form found in infected cells prior to virus budding) appears to have a weaker chaperone activity than the cleaved form of NC present in mature virions [152], potentially making the inhibition mechanism stronger prior to virus budding to prevent premature reverse transcription. Although it is difficult to mimic the exact effects of NC *in vitro*, future *in virio* and *in vivo* experiments should help test these hypotheses.

These unique properties of the initiation complex could make it an attractive target for the development of anti-HIV agents [56]. Indeed, initiation can be suppressed by both nucleoside-analog and non-nucleoside RT inhibitors, some of which display higher efficacy during initiation than elongation [162, 163]. The structural and mechanistic understanding of the initiation process demonstrated here could potentially help the development of novel antiviral agents, in particular by designing drugs that regulate the structural dynamics of the viral RNA and the binding orientation of RT, two factors of critical importance during early reverse transcription.

2.4 Materials and methods

2.4.1 Protein preparation

Recombinant HIV-1 RT p66 and p51 subunits were expressed and purified as described previously [164]. An E478Q mutation was introduced into the p66 RNase H domain to eliminate RNase H activity and prevent degradation of the substrates [165]. To produce Cy3-

labeled RT, the native cysteines on RT were changed to serine and a unique cysteine residue was introduced at the C-terminus of the p66 subunit. Purified RT molecules were incubated with Cy3 maleimide (GE Healthcare) for 60 minutes in 100 mM pH 7.0 NaH_2PO_4/Na_2HPO_4 buffer. Labeled RT was then dialyzed for ~ 48 hours to remove excess dye molecules. NC protein was a gift from Dr. Robert Gorelick (National Cancer Institute, Frederick, MD).

2.4.2 Nucleic acid preparation

Natural $tRNA^{Lys,3}$ purified from human placenta was purchased from Bio S&T. Natural tRNA-DNA chimeras were generated by employing RNase H cleavage [166] followed by a splinted ligation as shown in Figure 2.2.7. Synthetic DNA (Integrated DNA Technologies) and RNA (Dharmacon) oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis. vRNA templates longer than 80 nt were constructed by ligating two synthetic RNA pieces with T4 RNA ligase 2 (New England Biolabs). When necessary, the strands were specifically derivatized with a biotin or an amino modifier during synthesis. Cy3 or Cy5 mono-reactive NHS ester (GE Healthcare) was post-synthetically conjugated to the primary amine group on DNA or RNA. Labeled oligonucleotides were HPLC purified by reverse phase chromatography on a C8 column (GE Healthcare). To assemble a primer:template complex, primer and template strands were annealed with a 10:1 ratio in the annealing buffer (50 mM pH 8.0 Tris-HCl, 100 mM KCl, and 1 mM EDTA) using a Bio-Rad thermocycler with the following program: 2 minute at 95 °C followed by 10 minutes at 70 °C, then slowly cool down to 4 °C over 1 hour. Annealed products were analyzed on 5% native polyacrylamide gels to confirm successful assembly. Excess, unannealed primer molecules do not affect analysis as they are neither labeled by dye for detection nor labeled by biotin to allow surface anchoring.

2.4.3 Single-molecule FRET measurement

Nucleic acid substrates were immobilized on PEG-coated quartz microscope slides through biotin-streptavidin linkage as previously described [24]. The biotin group was placed at the

5' end of the template strand in all constructs. Surface-immobilized Cy5-labeled substrates were immersed in a solution containing Cy3-labeled RT. Measurements were performed in standard RT reaction buffer (50 mM pH 8.0 Tris-HCl, 50 mM KCl, 6 mM $MgCl_2$, and 0.1 mg ml^{-1} BSA) at room temperature (23 °C). An oxygen scavenger system (10% w/v glucose, 300 μg ml^{-1} glucose oxidase, and 40 μg ml^{-1} catalase) and a reducing reagent (2 mM Trolox) [167] were included in the buffer to minimize photobleaching and blinking. The concentration of Cy3-labeled p66 subunits used in the single-molecule experiments was 10 – 20 nM. A large excess of unlabeled wildtype p51 subunits, at a concentration (100 – 200 nM) comparable to or larger than the dissociation constant of the p66/p51 heterodimer [168] were added such that a substantial fraction of the Cy3-labeled p66 form heterodimers with p51. As the p66 and p51 subunits alone exhibited much lower affinities to the nucleic acid substrates and the p66 subunits were present at a much lower concentration than the dissociation constant of the p66/p66 homodimer [168], the majority of the binding events that we observed should involve a p66/p51 heterodimer. The observation of single-step RT dissociation or Cy3 photobleaching indicates that the observed binding events were from a single RT molecule.

The fluorescence signals for the FRET donor Cy3 and acceptor Cy5 were detected using a prism-type total-internal-reflection fluorescence microscope [24]. Alternating 532 nm (Crystal Laser) and 635 nm (Coherent) laser light were used to excite the sample [169]. Emission from the donor and acceptor were separated using a dichroic mirror (Chroma Technology) and imaged onto the two halves of a back-illuminated electron multiplying CCD camera (Andor Ixon 887). The FRET value during RT binding events was defined as $\frac{I_A}{I_A+I_D}$, where I_A and I_D were the fluorescence signals detected from the acceptor and donor channel, respectively.

To accurately identify RT binding events to nucleic-acid substrates, RT molecules non-specifically adhered to the slide surface need to be excluded as these molecules give a donor-only signal that could be mistaken as substrate-specific binding events with FRET = 0. Here,

in addition to using PEG-coated slides to minimize nonspecific sticking of RT to the surface, we further developed an algorithm to identify the remaining nonspecific surface-binding events using high-precision single-molecule localization, which allowed us to exclude these events during data analysis. In this assay, images of individual RT and substrate molecules, taken during 532 nm and 635 nm illumination, respectively, were fit to a Gaussian function to locate their centroid positions. A RT molecule within 0.5 pixels (~70 nm) of a substrate molecule was considered bound specifically to the substrate (Figure 2.4.1). While the image of a single molecule is diffraction limited (image width ~ a few hundred nanometers), the position of the molecule can be determined to a much higher precision depending on the number of photons detected [170, 171]. Thus, this single-molecule localization assay allows us to determine specific binding events with much higher fidelity than simply identifying overlapping Cy3 and Cy5 images.

2.4.4 Single-turnover nucleotide incorporation kinetics assay

The primer:template substrates were annealed as described above. Cy5 (GE Healthcare) was conjugated to the primers at the 5' end (ODN and ORN primers) or at position U67 (tRNA primers) to monitor DNA synthesis. Substrates (5 nM) were incubated with a saturating concentration of RT (~500 nM) in RT reaction buffer at room temperature for 5 minutes in a final reaction volume of 100 μl. Single nucleotide incorporation was initiated by the addition of selected dNTP (200 μM final concentration) and terminated by mixing a 5 μl aliquot of the reaction mixture with 15 μl of stop buffer (96% v/v formamide, 20 mM EDTA) at various time points. Reaction products were heated to 95 °C for 5 minutes, fractionated over a 6% (tRNA primers) or 10% (ODN and ORN primers) denaturing polyacrylamide gel, and quantified using a Typhoon Trio Imager (GE Healthcare).

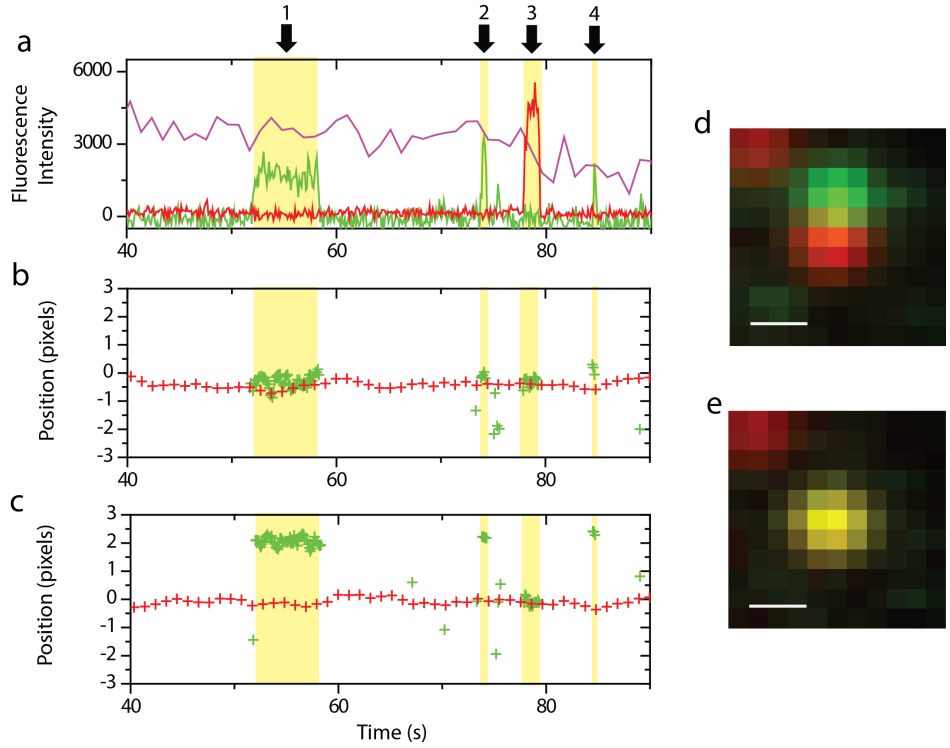


Figure 2.4.1: Rejecting non-specific protein binding events using high-precision single-molecule localization. Sample data from the interaction between Cy3-labeled RT and Cy5-labeled tRNA:vRNA substrate. **a,** Fluorescence time traces showing the Cy3 signal (green line) and Cy5 signal (red line) under 532 nm illumination and the Cy5 signal under direct 635 nm illumination (purple line). Four binding events detected by the increase of total fluorescence signal under 532 nm illumination are highlighted in yellow. **b-c,** The calculated x-coordinates (b) and y-coordinates (c) of the centroid positions for RT (green crosses) and the tRNA:vRNA substrate (red crosses). One pixel corresponds to 140 nm. **d-e,** Fluorescence images of the binding events taken under 532 nm illumination (green) overlaid over images taken under 635 nm illumination (red) for binding event 1 (d) and event 3 (e). The scale bar represents 0.4 µm. In this example, four apparent RT binding events are detected for the surface-anchored tRNA:vRNA substrate. For events 1, 2, and 4, the centroid positions show that RT binds at a location about two pixels away from the tRNA:vRNA substrate in the y-axis, indicating that the binding events represent the non-specific binding of RT to the surface. In contrast, for event 3, the centroid position of RT colocalizes with the centroid of the tRNA:vRNA substrate, indicating that this event is a specific binding event. Adapted with permission from [126].

2.4.5 RT pausing assays

The reactions were prepared as above with the following changes. To better detect the faint intermediate bands, primers were radiolabeled with ^{32}P using the KinaseMax kit (Ambion). Subsaturating concentrations of RT were used (20 nM for ODN, 50 nM for tRNA). Reactions were initiated by the addition of all four dNTPs (200 µM final concentrations) and terminated by mixing a 5 µl aliquot with 15 µl of stop buffer. Reaction products were fractionated over a denaturing polyacrylamide gel as above, exposed onto a PhosphoImager cassette (GE Healthcare), and quantified using the Typhoon Trio Imager.

Chapter 3

Histone H4 tail mediates allosteric regulation of nucleosome remodleing by linker DNA

3.1 Introduction

The ISWI family of ATP-dependent chromatin remodeling enzymes promotes transcriptional silencing and heterochromatin formation by generating and maintaining evenly spaced nucleosome arrays [63, 78, 87, 172, 173]. Biochemically, this activity manifests itself as the ability of these remodelers to center a mononucleosome on a short stretch of DNA [78, 87, 172, 173]. This nucleosome centering activity arises from the regulation of the enzymes' catalytic activity by the amount of linker DNA: The remodelers exhibit the fastest remodeling rates on substrates with long linkers, and the remodeling rate decreases as the linker DNA becomes shorter [114, 118, 121, 122, 123]. Thus, the enzyme repositions the histone octamer away from shorter linkers and toward longer linkers, centering the octamer on the DNA. The ISWI enzymes are sensitive to linker lengths of up to 60-70 bp [114, 118, 121, 122, 123], but the binding footprint of the catalytic subunits of these remodelers extends only ~20 bp away from the edge of the nucleosome [93, 174]. While it is known that the accessory subunits of these remodelers can bind linker DNA ~60 bp away from the edge of the nucleosome [93], how these accessory subunits communicate information about linker length to regulate the activity of the catalytic subunit remains poorly understood.

Previous studies of ISWI remodelers found that nucleosome sliding occurs in a series of translocation phases, where the ATPase pumps DNA around the nucleosome, interrupted by pause phases, where no DNA motion occurs [26, 27]. It is not known which of these phases is regulated by linker DNA length. In addition, the Cairns group identified two regions of the ISWI catalytic subunit, AutoN and NegC, that regulate enzyme activity [175]. The AutoN region resides in the N-terminus of the catalytic subunit and inhibits ATPase activity by binding to the ATPase domain. AutoN inhibition is relieved when the histone H4 tail displaces AutoN from the ATPase domain, allowing ATP hydrolysis to occur. The NegC region also binds to the ATPase domain, and this binding inhibits the coupling of ATPase activity to nucleosome sliding. Autoinhbition by NegC is relieved by binding of the HAND-SANT-SLIDE module of the catalytic subunit to linker DNA. Although NegC seems to regulate linker length sensing in the context of the isolated catalytic subunit, the role of these two regions in mediating linker length sensing in the context of the full remodeler complex is not known.

In order to investigate the linker length sensing mechanism, we studied the human ACF enzyme, a prototypical ISWI remodeler consisting of a catalytic subunit Snf2h and an accessory subunit Acf1 [87, 88, 176, 177]. Using FRET-based single-molecule and ensemble assays to monitor ACF activity, we found that linker length sensing occurs during the pause phases and not the translocation phases of nucleosome remodeling. When we examined various mutants of Snf2h, we found that the AutoN region but not the NegC region regulates linker length sensing, and that the H4 tail is also involved in the linker length sensing mechanism. Further biochemical experiments found that the N-terminal region of Acf1 aids in linker length sensing by sequestering the H4 tail in a linker length dependent manner. Together these results reveal how Snf2h and Acf1 coordinate with the histone H4 tail to regulate linker DNA length sensing by ISWI remodelers: on substrates with short linker DNA, the N-terminus of Acf1 binds to the H4 tail, preventing it from displacing AutoN. This leaves the ATPase in an inhibited state, resulting in slow remodeling on short-linker substrates.

On long linker substrates, however, the N-terminus of Acf1 instead binds to the linker DNA, leaving the H4 tail free to displace AutoN from the ATPase domain, and activate the ATPase to cause DNA translocation.

3.2 Results

3.2.1 Linker DNA length and the histone H4 tail regulate the remodeling pause phases

In order to monitor ACF activity, we used a FRET-based assay [114] to track the position of the nucleosome in real time during nucleosome sliding (Figure 3.2.1a). We reconstituted end-positioned nucleosomes labeled with the FRET donor dye Cy3 on the C-terminal tail of H2A and the FRET acceptor dye Cy5 on the DNA end [26]. This construct began with a high FRET value because the dyes were in close proximity, but after the addition of ACF and ATP, the enzyme repositioned the octamer toward the center of the DNA, increasing the distance between the dyes and lowering the FRET (Figure 3.2.1a,b). This assay recapitulated the linker length sensitivity of ACF as shortening the amount of linker DNA from 78 to 40 to 20 bp reduced the rate of sliding. ISWI enzymes also require an unmodified H4 tail for activity [116, 117, 118, 119, 120], and remodeling a substrate lacking an H4 tail, resulted in almost no sliding (Figure 3.2.1b).

We also monitored the FRET change at the single molecule level [15, 26], allowing us to resolve intermediates of the sliding process. As observed previously [26, 27], remodeling occurred in a series of alternating translocation and pause phases (Figure 3.2.1c). To determine how linker length and the H4 tail regulate remodeling, we first investigated which of these intermediate states linker length and the H4 tail regulate: the translocation phase or the pause phase. To address this question we examined remodeling on substrates with differing lengths of linker DNA or various H4 tail mutants and measured the average durations of each state to calculate the translocation rate and the pause phase exit rate. As we varied the

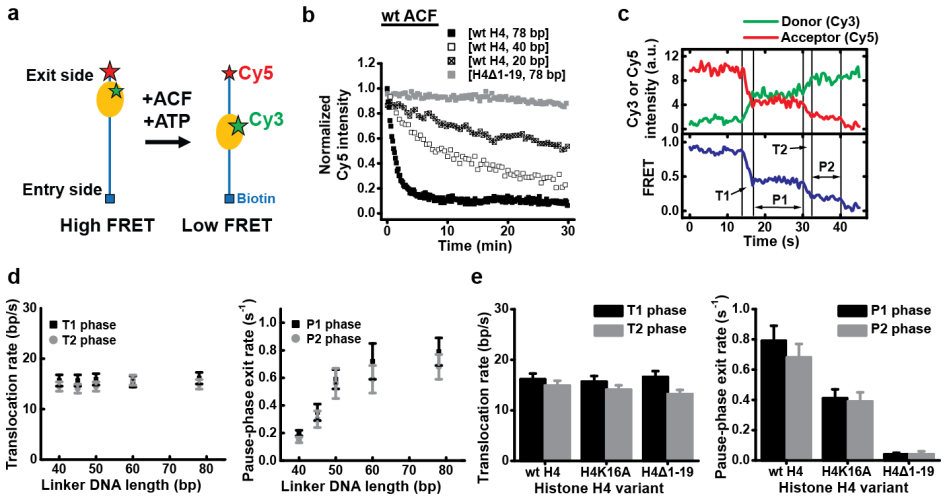


Figure 3.2.1: The linker DNA length and histone H4 tail regulate the remodelling pause phases but not the translocation phases. **a**, Schematic of a FRET-labelled mononucleosome undergoing remodelling by ACF. Nucleosome constructs are refered to as [WT H4/H4Δ1-19/H4K16A, n bp] to indicate the length of the longer DNA linker and the type of histone H4 (wild-type histone H4, a H4 tail deletion mutant, or H4 containing the K16A mutation) in the octamer. **b,** Ensemble remodelling time courses of [wt H4, 78 bp], [wt H4, 40 bp], [wt H4, 20 bp], and [H4Δ1-19, 78 bp] nucleosomes by 40 nM ACF at 5 µM ATP. Nucleosome translocation is monitored by the emission intensity of the FRET acceptor Cy5 under excitation of the FRET donor Cy3. **c**, Cy3 and Cy5 fluorescence (top) and FRET (bottom) time traces during the remodelling of a single [wt H4, 78 bp] nucleosome with the translocation (T1, T2) and pause (P1, P2) phases indicated. **d**, Linker DNA length dependence of the translocation rates between pauses (left, defined as the average number of base pairs moved per second) and pause-phase exit rates (right, defined as the inverse of the average pause durations). **e**, Dependence of the translocation rates between pauses (left) and pause-phase exit rates (right) on the H4 variants. In (d) and (e), [ACF] = 10 nM and [ATP] = 2 mM. Data are mean ± s.e.m. derived from at least 100 (d) or at least 50 (e) individual nucleosome remodelling traces from three independent experiments. Adapted with permission from [127].

length of the linker DNA, we saw that while the translocation rate does not vary, the pause phase exit rate decreased dramatically as linker length decreased (Figure 3.2.1d), indicating that exit from the pause phase is the key regulatory step. Similarly, when we examined nucleosomes with wild-type, mutant or deleted H4 tails, we saw that the translocation rate was not affected by perturbations to the H4 tail, but again, exit from the pause phase was the regulated step (Figure 3.2.1e).

3.2.2 AutoN and the histone H4 tail are involved in linker length sensing

Because the pause phase appeared to be the key regulatory step in remodeling, we next asked what mechanism inhibits the ATPase during the pause phase and regulates exit from this phase. The Cairns group identified two regions of the catalytic subunit of ISWI remodelers that inhibit its ATPase activity: AutoN and NegC [175]. Because NegC mediates the linker length sensitivity of the isolated catalytic subunit [175], we hypothesized that it may mediate the linker length sensitivity of the full ACF complex. To determine whether NegC is involved in regulating exit from the pause phase, we expressed a mutant ACF complex with NegC deleted from the catalytic subunit (Figures 3.2.2a). The resulting complex retained its linker length sensitivity and showed a reduced remodeling rate on short linker substrates compared to long linker substrates (Figure 3.2.2b). This result indicated that NegC was not involved in linker length sensing in the context of the full ACF complex. Surprisingly, when we mutated AutoN in the full ACF complex, we saw that this mutant lost its linker length sensitivity and exhibited full activity on both short and long linker substrates (Figure 3.2.3a,b). We observed quantitatively similar results when we measured the pause phase exit rates in our single molecule assay (Figure 3.2.2c and 3.2.3c). These results indicate that the slow remodeling rate and slow pause phase exit observed on short linker substrates were due to AutoN inhibition.

Because the H4 tail competes with AutoN for binding to the ATPase [175], we hypothe-

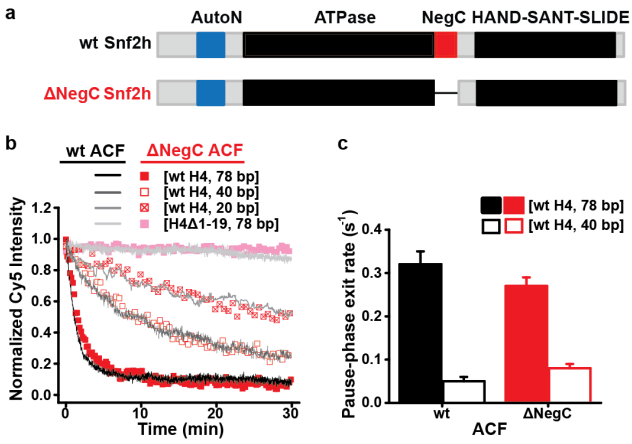


Figure 3.2.2: Deletion of the NegC domain of the Snf2h catalytic subunit does not substantially affect linker DNA length sensing by the ACF complex. **a**, Domain architecture of wt and ΔNegC Snf2h (residues 669-700 replaced with a SGSGS linker). **b**, Ensemble remodelling time courses of [wt H4, 78 bp], [wt H4, 40 bp], [wt H4, 20 bp], and [H4Δ1-19, 78 bp] nucleosomes by 40 nM wt ACF (black/gray lines, duplicated from Fig. 1b) and ΔNegC ACF (red/pink symbols) at 5 μM ATP. **c**, Linker DNA length dependence of the pause-phase exit rate (P1 phase) measured for wt ACF (black) and ΔNegC ACF (red). [ACF] = 10 nM and [ATP] = 20 μM. Data are mean ± s.e.m. derived from at least 100 individual nucleosome remodelling traces from three independent experiments. Adapted with permission from [127].

sized that ACF regulates AutoN inhibition by altering H4 tail accessibility. If this were true, the H4 tail would only be available to displace AutoN from the catalytic subunit on long linker substrates. Thus, AutoN should remain bound to the catalytic subunit on short linker substrates. We tested this hypothesis in two ways. First, we examined the effect of exogenously added H4 tail peptide. This peptide should increase the remodeling rate (and thus decrease the pause phase exit rate) on short linker substrates because it can help compete AutoN off of the ATPase, but it should have no effect on long linker substrates where AutoN is already displaced. Consistent with these expectations, adding H4 tail peptide dramatically decreased the pause phase exit rate for short linker nucleosomes but it had no effect on long linker nucleosomes (Figure 3.2.3d). Second, we looked for linker length dependence on substrates lacking an H4 tail. In these experiments, both long linker and short linker constructs should show similar remodeling rates because, in the absence of H4 tail, AutoN will be engaged in both cases. Consistent with these expectations, both long linker and short

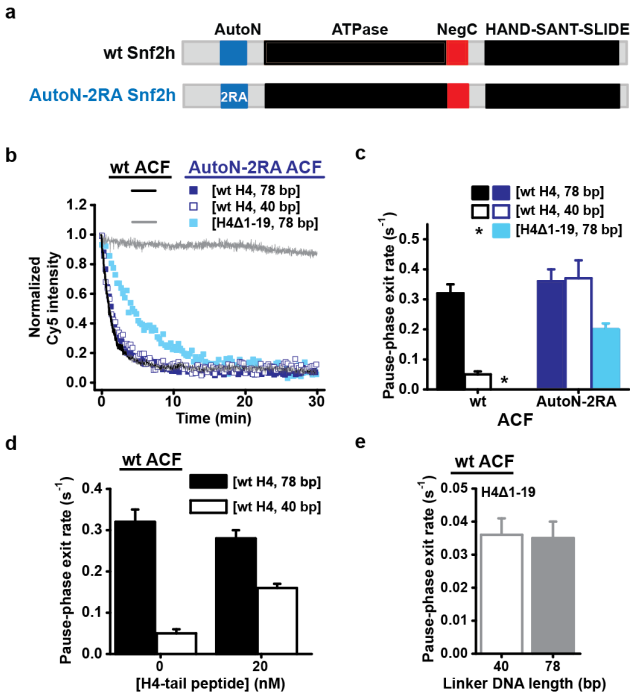


Figure 3.2.3: The AutoN domain of Snf2h and the nucleosomal H4 tail are important for linker DNA length sensing by the ACF complex. **a**, Domain architecture of wt and AutoN-2RA (R142A and R144A) Snf2h. **b**, Ensemble remodelling time courses of [wt H4, 78 bp], [wt H4, 40 bp] and [H4Δ1-19, 78 bp] nucleosomes by 40 nM wt ACF (black/gray lines, duplicated from Fig. 1b) and AutoN-2RA ACF (blue/cyan symbols) at 5 µM ATP. **c**, Dependence of the pause-phase exit rate on the linker DNA length and H4 tail for wt (black) or AutoN-2RA ACF (blue/cyan). * too slow to be measured. **d**, Effect of the exogenously added H4-tail peptide on the pause-phase exit rates during remodelling by wt ACF. **e,** Pause-phase exit rates of nucleosomes lacking the H4 tail during remodelling by wt ACF. In (c-e), [ACF] = 10 nM and [ATP] = 20 µM, except that 2 mM of ATP was used in (e) to make the pause exit rates measurable for nucleosomes lacking the H4 tail. Data are mean ± s.e.m from at least 100 (c, d) or at least 50 (e) individual nucleosome remodelling traces from three independent experiments. Adapted with permission from [127].

linker substrates lacking an H4 tail gave similar pause phase exit rates in our single molecule assay (Figure 3.2.3e).

3.2.3 The N-terminal region of Acf1 regulates H4 tail accessibility

Because Acf1 is responsible for sensing linker DNA $>$ 30 bp from the nucleosome [93], we hypothesized that Acf1 might bind the H4 tail on short linker substrates, preventing the H4 tail from relieving H4 inhibition. Interaction with linker DNA might then prevent Acf1 from binding the H4 tail, allowing the tail to relieve AutoN inhibition. This model makes a few predictions we can test. First, we should be able to see a direct interaction between Acf1 and the H4 tail. Second, if we identify a mutant of Acf1 unable to bind H4 tail, ACF complex containing this mutant should be insensitive to linker length. Finally, the Acf1-H4 tail interaction should decrease with increasing linker length.

To probe for a direct interaction between Acf1 and the H4 tail we performed a fluorescence anisotropy assay to measure the binding isotherm for the interaction between the isolated Acf1 subunit and a fluorescently-labeled H4 tail peptide. These measurements showed that they interacted with a K_d of 3 nM (Figure 3.2.4b). Next, we looked for the region of Acf1 responsible for binding to the H4 tail and found that deletion of the N-terminus of Acf1 greatly decreased binding to the H4 tail (Figure 3.2.4a,b). When we expressed ACF complex containing Acf1 with its N-terminal region deleted, we saw that this complex exhibited no linker length sensitivity, while retaining a requirement for the H4 tail (Figure 3.2.4c,d). This result is consistent with the idea that the ΔN-term complex could no longer bind the H4 tail to inhibit remodeling at short linker lengths. These results further suggest a mechanism for why Acf1 does not bind the H4 tail on long linker substrates; the DNA binding domain of Acf1 also resides in the N-terminus [178], so interaction of Acf1 with linker DNA may physically remove Acf1 from the proximity of the H4 tail.

To test whether the Acf1-H4 tail interaction is regulated by the length of linker DNA, we used a crosslinking assay to probe for interactions with the H4 tail when ACF is bound to

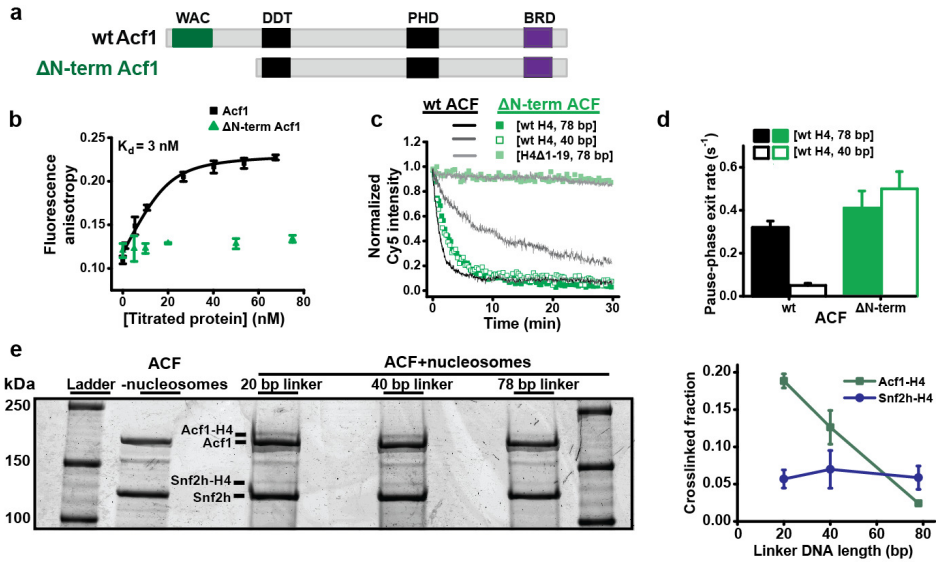


Figure 3.2.4: The N-terminal region of the Acf1 accessory subunit is important for linker DNA length sensing by the ACF complex. **a**, Domain architecture of wt and ΔN-term (residues 1-371 deleted) Acf1. **b**, Fluorescence anisotropy of dye-labelled H4-tail peptide in the presence of varying amounts of wt (black symbols) or ΔN-term (green symbols) Acf1. Data are mean ± s.e.m. (n = 3 independent experiments). The K_d for wt Acf1 is 3 ± 9 nM (error represents the 95% confidence interval). **c**, Ensemble remodelling time courses of [wt H4, 78 bp], [wt H4, 40 bp], and [H4Δ1-19, 78 bp] nucleosomes by 40 nM wt ACF (black/gray lines, duplicated from Fig. 1b) and ΔN-term ACF (green/light green symbols) at 5 µM ATP. **d**, Dependence of the pause-phase exit rate on the linker DNA length for wt ACF (black) or ΔN-term ACF (green). [ACF] = 10 nM and [ATP] = 20 µM. Data are mean ± s.e.m derived from at least 100 individual nucleosome remodelling traces from three independent experiments. **e**, Crosslinking of the H4 tail to Acf1 depends on the linker DNA length. The crosslinking products were analysed by SDS-PAGE (left). The Acf1-H4 crosslinking band was absent for ACF without nucleosomes (lane "- nucleosomes"). (right) Quantification of the H4-crosslinked fractions of Acf1 and Snf2h as a function of linker DNA length. Data are mean ± s.e.m. (n = 3 independent crosslinking experiments). Adapted with permission from [127].

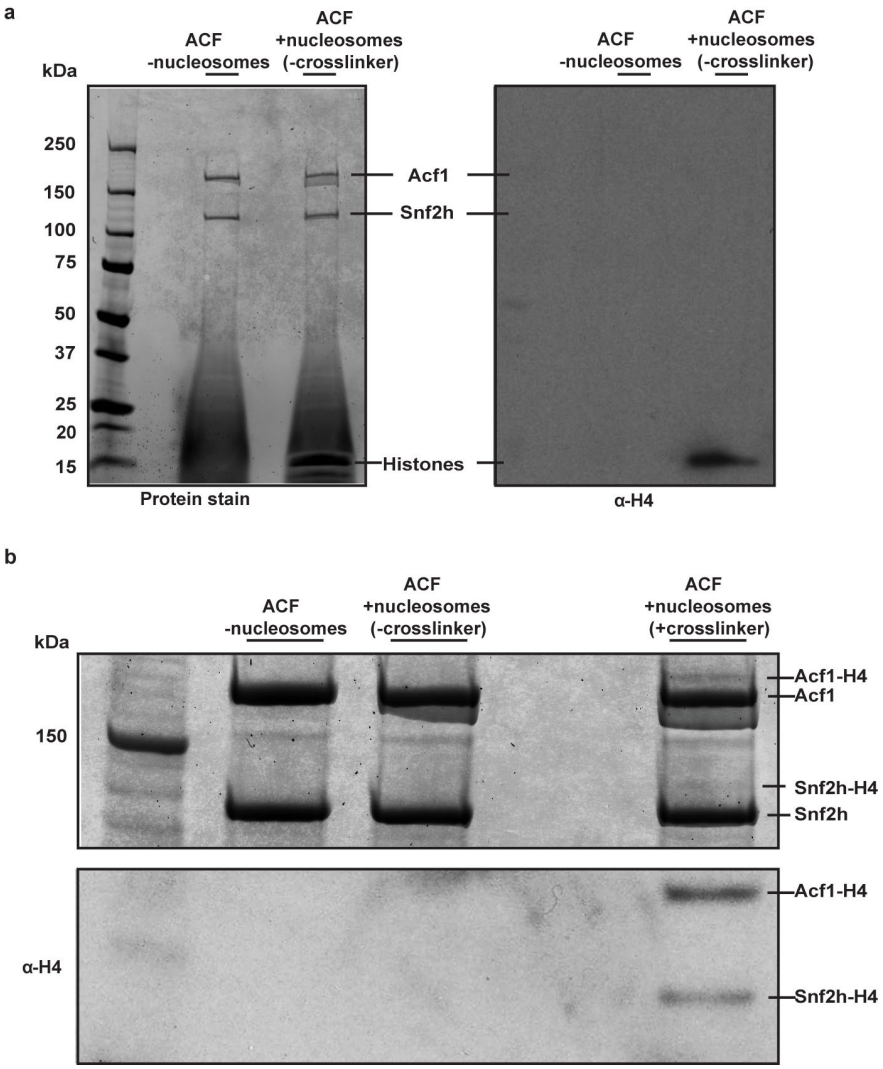


Figure 3.2.5: α-histone H4 immunoblot analysis validates the formation of the Acf1-H4 crosslinked product. **a**, Left: SDS PAGE analysis of samples containing ACF alone or ACF with nucleosomes (20 bp linker DNA) that do not possess the cysteine-reactive crosslinker on the H4 tail. Both samples yield two distinct bands corresponding to the Acf1 and Snf2h subunits (180 kDa and 122 kDa, respectively). Additionally, histone bands at low molecular weights are present in the lane for the sample containing nucleosomes. Right: Corresponding immunoblot using α-H4 antibody. In the presence of nucleosomes without crosslinker, a single H4 band is visible at ~11 kDa corresponding to the histone itself. **b**, Top: Incubation of ACF and nucleosomes that contain a crosslinker at the H4 tail yield Acf1-H4 and Snf2h-H4 crosslinking bands. These bands are absent for ACF without addition of nucleosomes ("- nucleosomes") or upon addition of nucleosomes without a crosslinker. Proteolytic degradation of Acf1 gave rise to a fainter band immediately below Acf1. Bottom: α-histone H4 immunoblotting reveals specific Acf1-H4 and Snf2h-H4 bands that are absent for ACF without addition of nucleosomes ("- nucleosomes") or upon addition of nucleosomes without a crosslinker. Adapted with permission from [127].

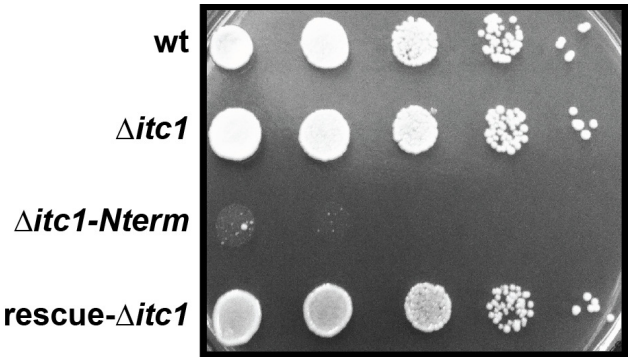


Figure 3.2.6: Effects of deletion of Itc1 (Acf1 homolog) and its N-terminal region on the growth of yeast cells. Top row: wt. Second row: the *itc1* gene is deleted (Δ*itc1*). Third row: the coding sequence of the N-terminal region of Itc1 is deleted (Δ*itc1-Nterm*). Bottom row: The remaining portion of *itc1* is deleted from Δ*itc1-Nterm* (rescue-Δ*itc1*). One representative of three independent growth experiments is shown. Adapted with permission from [127].

nucleosomes. We installed a maleimide crosslinker onto the H4 tail of the nucleosomes with varying linker lengths, incubated these nucleosomes with ACF complex, then analyzed the reaction on a SDS-PAGE gel to assess the amount of Snf2h and Acf1 that were crosslinked to the H4 tail. The crosslinked products were detected as bands that exhibited reduced mobility compared to the uncrosslinked Snf2h and Acf1 proteins (Figure 3.2.4e). Notably, these bands were not present in the absence of nucleosomes or crosslinker, and a western blot for histone H4 showed that these bands contained histone H4 (Figure 3.2.5). Consistent with our hypothesis, we observed a reproducible decrease in the amount of Acf1-H4 crosslinking as the linker DNA length increased (Figure 3.2.4e). Snf2h-H4 crosslinking remained fairly constant, likely because Snf2h stayed in close proximity to the H4 tail regardless of whether it bound the H4 tail or not (Figure 3.2.4e). These results provide a direct confirmation of our model that Acf1 regulates Snf2h catalytic activity by binding the H4 tail in a linker length dependent manner.

Finally, we tested the physiological importance of our model by turning to the yeast homolog of ACF, ISW2. The Tsukiyama group has previously shown that deletion of the accessory subunit of the yeast ACF homolog produces no growth defect [179], but we hypothesized that deletion of only the N-terminus should create a much stronger defect. The

resulting mutant would still be active on ISWI target genes, but it will lack spacing activity and may therefore disrupt nucleosome spacing in silenced regions, resulting in a global disruption of chromatin structure. Indeed, when we made this deletion mutant in yeast, we saw a very strong growth defect (Figure 3.2.6), indicating that this linker length sensing mechanism is crucial for the proper *in vivo* function of ISWI remodelers.

3.3 Discussion

In this study, we investigated the mechanism of linker length sensing by the ISWI chromatin remodeling complex ACF. Through a combination of single molecule and ensemble FRET measurements, crosslinking experiments, and yeast genetics, we identified the key structural features of the nucleosome and ACF complex that cooperate to transmit information about the presence of linker DNA from the accessory subunit, Acf1, to the catalytic subunit, Snf2h, in order to modulate the rate of remodeling.

Our single molecule FRET experiments enabled us to determine that linker DNA length modulates the length of the pause phases that occur during remodeling but does not modulate the speed of DNA translocation during the translocation phases. This linker length regulation requires the AutoN region of Snf2h and the H4 tail on the octamer because disrupting either of these regions resulted in the loss of linker length sensitivity by ACF. We further found that the N-terminal region of Acf1 binds to the H4 tail in a linker length dependent manner, and that this interaction is required for linker length sensing. Disruption of the homologous region of the yeast ISW2 complex causes a severe growth defect, indicating its importance to the linker sensing mechanism of ISWI remodelers.

Together, our data lead us to propose the following model for the mechanism of linker length sensing by ISWI remodelers (Figure 3.3.1). Remodeling consists of an alternating series of translocation phases, where the H4 tail is engaged with the ATPase allowing sliding to occur, and pause phases, where linker length sensing occurs. During the pause phase,

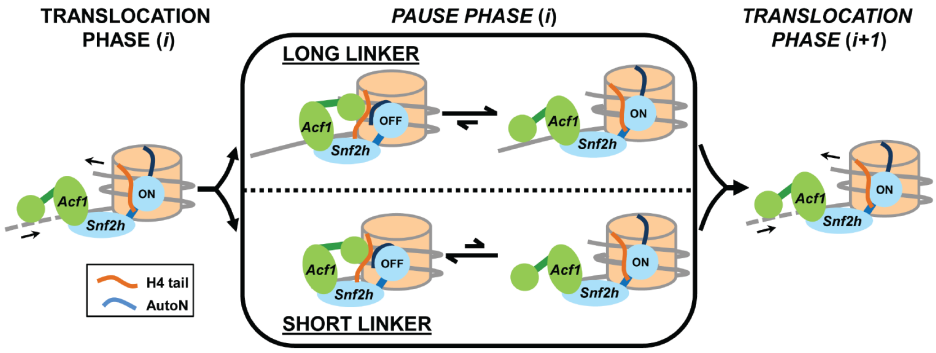


Figure 3.3.1: Model for linker DNA length sensing by the ACF complex. DNA: gray lines, histone octamer: beige cylinders, Snf2h: blue/cyan, Acf1: green. The ATPase domain of Snf2h is depicted as a cyan sphere and labelled "ON" when active and "OFF" when inactive. Adapted with permission from [127].

Acf1 samples for the presence of linker DNA. If the linker is short, Acf1 binds to the H4 tail, preventing it from relieving AutoN inhibition and leaving the ATPase inactive. If the linker is long, the Acf1 N-terminus binds the linker DNA, removing the H4 tail binding site from the proximity of the H4 tail. This leaves the H4 tail free to displace AutoN, activate the ATPase, and promote entry into the next translocation phase. These results provide a mechanistic picture of the molecular interactions underlying the linker length sensing ability of an important family of chromatin-regulatory enzymes.

3.4 Materials and methods

3.4.1 DNA constructs

Double-stranded (ds) DNA constructs containing the 601 nucleosome positioning sequence [180], varying DNA linker lengths on the entry side of the nucleosome, 3 bp linker DNA on the exit side of the nucleosome, a FRET acceptor Cy5 attached to the 5' end of the exit-side linker DNA, and a biotin moiety attached to the 5' end of the entry-side linker DNA were generated by PCR and purified by PAGE. The dsDNA constructs with an additional 38 nucleotide (nt) single-stranded (ss) DNA spacer used for single-molecule mononucleosome remodelling experiments were created by annealing and ligating a set of overlapping, complementary

oligonucleotides. The 38 nt ssDNA spacer was used to prevent surface perturbation in remodelling. These HPLC-purified oligonucleotides (Integrated DNA Technologies) were mixed at equimolar concentrations in 50 mM Tris pH 8.0, 100 mM KCl, 1 mM EDTA and annealed with an temperature ramp (95°C - 3°C), ligated with T4 DNA ligase (New England Biolabs), and purified by PAGE. Successful ligation was confirmed by denaturing PAGE.

3.4.2 Histone octamers

Recombinant *X. laevis* H2A, H2A K119C, H2B, H3, H3 C110A, H4, H4 K16A, H4Δ1-19, and H4 NCys (with an N-terminal cysteine residue inserted) histones were expressed in BL21 (DE3) pLysS competent cells (Stratagene) and purified from inclusion bodies by gel filtration (Sephacryl S-200 column, GE Healthcare) and ion exchange (Resource-S column, GE Healthcare) chromatography under denaturing conditions [59, 181]. The H4 K16A, H4Δ1-19, and H4 NCys mutants were created by site-directed mutagenesis.

For site-specific labelling of histone H2A, a cysteine substitution was introduced at residue 119 (K119C) as described previously [114]. Lyophilized H2A was dissolved in labelling buffer (7 M guanidine-HCl, 20 mM Tris pH 7.0, 5 mM Na-EDTA, 1.25 mM TCEP) and incubated for 2 hr in the dark. Cy3-maleimide was dissolved in DMSO and added to the reaction to a final concentration of 2.5 mM. After incubating for 3 hr in the dark, the reaction was quenched by adding 80 mM β-mercaptoethanol. The labelled H2A was dialyzed three times against dialysis buffer (7 M guanidine HCl, 20 mM Tris pH 7.0, 1 mM DTT) in a 7K MWCO dialysis cassette (Pierce) and quantified. To reconstitute histone octamer, each lyophilized core histone was dissolved in unfolding buffer (7 M guanidine-HCl, 20 mM Tris-HCl pH 7.5, 10 mM DTT), mixed in a molar ratio of 1.2:1.2:1:1 H2A:H2B:H3:H4, dialyzed against refolding buffer (2 M NaCl, 10 mM Tris-HCl pH 7.5, 1 mM Na-EDTA, 5 mM β-mercaptoethanol), and purified by gel filtration chromatography (Superdex 200 10/100 GL column, GE Healthcare) as previously described [59]. For making FRET-labelled nucleosomes, a ~1:1 ratio of Cy3-labelled and unlabelled H2A mixture was used. Because each histone octamer contains

two H2A subunits, this reconstitution yielded three distinct populations of Cy3-labelled nucleosomes: (1) with Cy3 attached only to the H2A subunit proximal to the FRET acceptor Cy5 on the DNA, (2) with Cy3 attached only to the H2A subunit distal to the Cy5 on the DNA, and (3) with both H2A labelled by Cy3. These three labelling configurations yielded distinct FRET levels that could be clearly distinguished at the single-molecule level, as shown previously [26]. To maximize the dynamic range for single-molecule FRET measurements, we selected the population with the highest FRET value, which corresponded to nucleosomes with a single Cy3 on the proximal H2A, for further analysis. For ensemble FRET measurements, because the overall remodelling rate was derived from the decay curve of the acceptor signal, distinguishing the three populations was unnecessary. For crosslinking experiments, only unlabelled H2A was used.

3.4.3 Nucleosomes

Mononucleosomes were reconstituted from DNA and histone octamers at 4°C [59]. The nucleosome reconstitution reaction consisted of 2 M KCl, 20 mM Tris-HCl pH 7.5, 1 mM Na-EDTA, 1.2 µM histone octamer, 1 µM DNA construct, 10 mM DTT, and 0.5 mM benzamidine. The sample was injected into a pre-hydrated 7K MWCO dialysis cassette (Pierce) and subjected to salt gradient dialysis (2 M KCl to 250 mM KCl over 60 hours) followed by incubation in TCS buffer (20 mM Tris-HCl pH 7.5, 1 mM EDTA, 1 mM DTT) for 8 hr at 4°C. Assembled nucleosomes were purified on a 10-30% glycerol gradient in 20 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.1% Igepal CA-630. Fractions were analysed by PAGE (5% 0.5x TBE) and fractions containing pure and correctly assembled nucleosomes were pooled and concentrated with a 100K MWCO Amicon Ultra-0.5 mL centrifugal filter (Millipore) to ~30 µL final volume and stored at 4°C. For the assembly of nucleosomes used in crosslinking experiments, reducing agent was omitted.

3.4.4 ATP-dependent chromatin assembly and remodelling factor (ACF)

To produce isolated Snf2h-FLAG or isolated Acf1-FLAG, these proteins were overexpressed in Sf9 cells using a baculovirus expression system (Kinnakeet Biotechnology). To produce ACF complexes, Snf2h and Acf1-FLAG were co-expressed in Sf9 cells. Nuclear extraction was performed as described previously [114, 121]. Nuclear extract was fortified with 0.5 mM benzamidine, 60 µg/mL TLCK, and 1x Roche Complete Protease Inhibitor Cocktail and then purified by M2-affinity chromatography (Anti-FLAG M2 beads and FLAG peptide, Sigma-Aldrich), as described previously [114, 121]. Specifically, the FLAG-tagged subunit or complex was bound to M2 beads, and the beads were washed several times with wash buffer containing 20 mM HEPES, pH 7.9, 20% glycerol, 0.2 mM EDTA, 1 mM benzamidine, 1 mM DTT, Roche EDTA-free complete protease inhibitor, and various concentrations of KCl (300, 150, and 100 mM sequentially). Protein was subsequently eluted with a 1 mg/mL FLAG peptide solution in wash buffer with 100 mM KCl. Elution fractions were analysed by PAGE and fractions containing the isolated subunits or the intact ACF complex were pooled and concentrated with a 50K MWCO Amicon Ultra-15 mL centrifugal filter (Millipore). Concentrated ACF complexes, Snf2h subunit, and Acf1 subunit were then aliquoted, flash frozen, and stored at -80°C in 20 mM HEPES, pH 7.9, 20% Glycerol, 0.2 mM EDTA, 100 mM KCl, and 1 mM DTT. To purify enzyme for crosslinking experiments, the complete protease inhibitor pills and DTT were omitted from the buffers during the affinity purification steps. Typical yields were ~50-100 µg, 20-40 µg, and 1-2 µg of ACF complexes, Snf2h subunit, and Acf1 subunit, respectively, per litre of insect cell culture. Quantification was performed by SYPRO-red staining and comparison to BSA standards. The Snf2h and Acf1 mutants were created by site-directed mutagenesis. SDS-PAGE analysis of the purified protein indicated that the mutants eluted from the column as a complex with the correct stoichiometry (Figure 3.4.1).

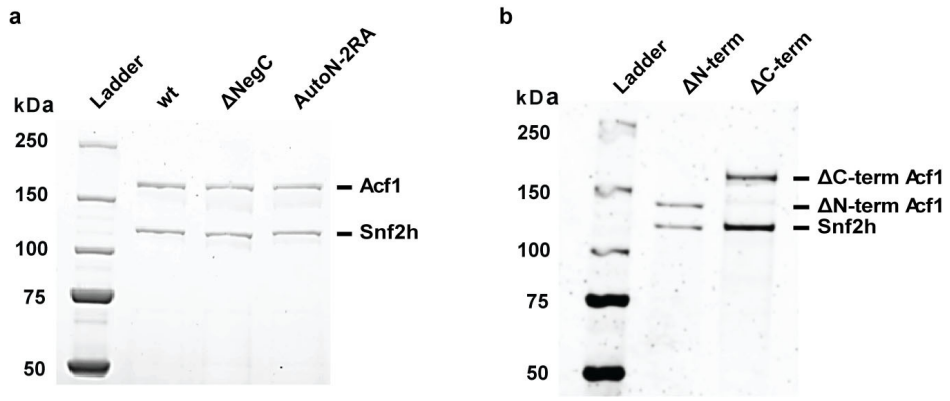


Figure 3.4.1: SDS-PAGE analysis of wt, ΔNegC, AutoN-2RA, ΔN-term and ΔC-term ACF complexes. **a**, wt Snf2h, ΔNegC Snf2h, or AutoN-2RA Snf2h were co-expressed with Acf1-FLAG in Sf9 insect cells and purified by affinity chromatography. **b**, ΔN-term Acf1 or ΔC-term Acf1 was co-expressed with Snf2h-FLAG and purified by affinity chromatography. The presence of both Acf1 and Snf2h in each case indicated that wt SNF2h, ΔNegC SNF2h and AutoN-2RA SNF2h can all form complexes with Acf1 and that ΔN-term Acf1 and ΔC-term Acf1 can both form complexes with Snf2h. Adapted with permission from [127].

3.4.5 Ensemble FRET assay for nucleosome remodelling

The ensemble remodelling kinetics of nucleosomes were measured by monitoring the total acceptor (Cy5) signal (under a constant 532 nm illumination that excites the donor Cy3) of a solution of FRET-labelled nucleosomes after addition of ACF and ATP to initiate remodelling. As nucleosomes were translocated towards the centre of the DNA, the FRET efficiency decreased, which resulted in a decrease in the Cy5 signal. No remodeling was seen in the absence of ATP (Figure 3.4.2a).

The ensemble nucleosome remodelling assay was performed by diluting nucleosomes in remodelling buffer (40 mM Tris pH 7.5, 12 mM HEPES pH 7.9, 60 mM KCl, 0.32 mM EDTA, 3 mM $MgCl_2$, 100 µg/mL acetylated BSA, 10% (v/v) glycerol, 0.02% (v/v) Igepal CA-630, 10% glucose, 2 mM Trolox, and 2.5 mM protocatechuic acid) in a cuvette fitted for a Cary Eclipse Fluorescence Spectrophotometer (Varian). Immediately before starting a kinetic scan (ex = 532 ± 10 nm, em = 670 ± 20 nm, 1 Hz acquisition frequency, 20°C), ACF, ATP, and additional $MgCl_2$ equimolar to ATP were mixed in the remodelling buffer and added to the nucleosomes, achieving a final concentration of 2.5 nM nucleosomes and

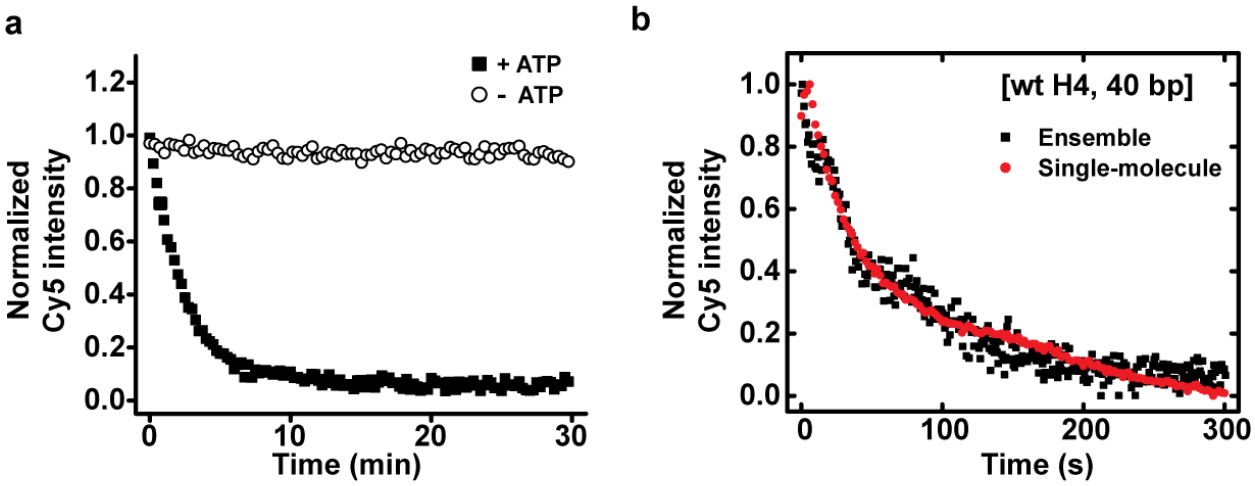


Figure 3.4.2: Control experiments to test the effect of ATP and surface-anchoring on nucleosome remodelling. **a**, Ensemble remodelling time courses of [wt H4, 78 bp] nucleosomes by 10 nM ACF with 5 µM ATP (filled symbols) or without (open symbols) ATP. Nucleosome translocation is monitored by the emission intensity of the FRET acceptor Cy5 under constant 532 nm illumination that excites the FRET donor Cy3. **b**, Comparison of the average remodelling kinetics for surface-anchored [wt H4, 40 bp] nucleosomes (measured by the single-molecule assay, $>$ 250 nucleosomes) and freely diffusing [wt H4, 40 bp] nucleosomes in solution (measured by the ensemble assay). [ACF] = 5 nM and [ATP] = 20 µM. Adapted with permission from [127].

varying concentrations of enzyme and ATP, as reported. For remodelling experiments in the presence of the H4-tail peptide, synthetic peptide containing the H4-tail sequence (residues 1-21 of histone H4, SignalChem) was mixed together with ACF, ATP, and $MgCl_2$ in the remodelling buffer before being added to the nucleosomes. The final concentration of the H4-tail peptide was 20 nM.

3.4.6 Single-molecule FRET assay for nucleosome remodelling

Quartz slides were cleaned, functionalized with poly(ethylene glycol) (PEG, Laysan Bio), and assembled into flow chambers as previously described [26]. Reconstituted nucleosomes were anchored to coated quartz slides via a streptavidin-biotin linkage. To eliminate potential perturbation of remodelling by surface anchoring, especially for mononucleosomes with relatively short entry-side linker DNA, we inserted a 38 nt ssDNA spacer between the biotin and dsDNA. With the ssDNA spacer, the remodelling kinetics of surface-anchored nucleosomes with 40 bp or longer entry-side linker DNA were indistinguishable from those of freely diffusing nucleosomes measured using the ensemble FRET assay (Figure 3.4.2b). We performed our single-molecule FRET studies on nucleosomes with at least 40 bp of entry-side linker DNA.

The FRET donor Cy3 was directly excited with a 532 nm Nd:YAG laser (CrystaLaser) in a total-internal-reflection-fluorescence (TIRF) configuration and fluorescence emission from Cy3 and Cy5 was captured with a 60x water immersion objective (Olympus), filtered with a 550 nm long-pass filter (Chroma Technology), spectrally split by a 630 nm dichroic mirror (Chroma Technology), and imaged onto two halves of a CCD camera (Andor iXonEM+888). The imaging conditions were 30°C in the imaging buffer (40 mM Tris pH 7.5, 12 mM HEPES pH 7.9, 60 mM KCl, 0.32 mM EDTA, 3 mM $MgCl_2$, 100 µg/mL acetylated BSA, 10% (v/v) glycerol, 0.02% (v/v) Igepal CA-630), 10% (w/v) glucose, 800 µg/mL glucose oxidase, 50 µg/mL catalase, 2 mM Trolox, and 2.5 mM protocatechuic acid. In certain experiments, 20 nM H4-tail peptide was included. The enzymatic oxygen scavenging system and antioxidants

(10% glucose, 800 µg/mL glucose oxidase, 50 µg/mL catalase, 2 mM Trolox, and 2.5 mM protocatechuic acid) were used to reduce fluorophore photobleaching and blinking [167, 182]. Remodelling was initiated by infusing the flow chamber with a solution containing ACF and ATP in imaging buffer using a syringe pump (KD Scientific).

3.4.7 Fluorescence anisotropy binding assay

For binding studies, we used a synthetic peptide (SignalChem) that contained residues 1-21 of human histone H4 as well as an additional C-terminal cysteine to enable thiol-based labelling. Tetramethylrhodamine-5-iodoacetamide (TMRIA) was coupled to the C-terminus of the synthetic H4-tail peptide to yield TMR-labelled H4 peptide (H4-tail-TMR). The coupling reaction was carried out by shaking 0.2 µmol peptide, ~2 fold molar excess of TMRIA, and 10 fold molar excess of TCEP in 1 mL of 20 mM Tris (pH 7.5) in the dark for 2 h at room temperature. The labelled peptide was purified by RP-HPLC and its identity confirmed by liquid chromatography–mass spectrometry analysis.

Fluorescence anisotropy measurements were carried out at 25°C using a Molecular Devices SpectraMax microplate reader. Binding curves were recorded in low-volume 384-well assay plates. Individual wells contained 20 nM H4-tail-TMR and the indicated concentration of Acf1 in 20 mM Tris (pH 8.0), 100 mM KCl, and 20% (vol/vol) glycerol. The fluorescence anisotropy was measured as a function of the Acf1 concentration. Dissociation constants were determined by fitting the fluorescence anisotropy data to the solution of a quadratic equation derived from the binding isotherm, which takes depletion of Acf1 protein into account:

$$r = \frac{\alpha}{2[H4]_t}\left([P]_t + [H4]_t + K_d - \sqrt{([P]_t + [H4]_t + K_d)^2 - 4[P]_t[H4]_t}\right) + r_o$$

where r is the fluorescence anisotropy, r_o represents an offset value, α is a scaling factor, K_d is the dissociation constant, $[H4]_t$ represents the total concentration of the H4-tail-

TMR peptide, and $[P]_t$ is the total concentration of titrated Acf1 protein. Errors in the determination of K_d were estimated by calculating the 95% confidence interval.

3.4.8 Protein crosslinking

Homobifunctional maleimide crosslinker BM(PEG)$_3$ (Pierce) was added to concentrated H4NCys nucleosomes (defined below) to achieve a final concentration of ~5 µM nucleosomes and 500 µM crosslinker. The H4 NCys nucleosome contains a single labelling site in the form of a cysteine added to the N-terminus of histone H4, immediately following the initial methionine. The other histone proteins used to prepare H4 NCys nucleosomes were wt H2A, wt H2B, and H3 C110A, and do not contain additional cysteine residues. The crosslinker-labelling reaction was performed at room temperature for 1.5 hr. Buffer exchange with wash buffer (20 mM HEPES pH 7.0, 1 mM EDTA, 10% glycerol, 10 mM KCl, 0.1% Igepal CA-630) was performed using a 100K MWCO Amicon Ultra-0.5 mL centrifugal filter (Millipore) five times to remove unreacted crosslinker. ACF was then added to the BM(PEG)$_3$-derivatised nucleosomes to achieve a final enzyme concentration of 200–400 nM and nucleosome concentration of 5 µM, and incubated for 2 hr. Crosslinking products were analysed by SDS-PAGE (4-15%, 1x Tris/Glycine/SDS) and SYPRO-red staining. Under each condition, the Acf1-H4 crosslinked fraction was calculated as:

$$\frac{I_{Acf1-H4}}{(I_{Acf1-H4} + I_{Acf1}) \times I_{rel,histones}}$$

where $I_{Acf1-H4}$ and I_{Acf1} are the integrated intensities of the Acf1-H4 and Acf1 bands, respectively, and $I_{rel,histones}$ is the relative histone amount obtained by dividing the histone band intensity in a given lane by the maximum histone band intensity in the gel. The Snf2h-H4 crosslinked fraction was obtained analogously. Normalization by the relative histone intensity in each lane compensates for any small variation in the nucleosome concentration used, even though efforts were made to maintain identical nucleosome concentrations for all

conditions. For validation of the H4-crosslinking product, immunoblotting was performed using a primary rabbit polyclonal α-histone H4 antibody (Abcam ab10158, tested for Western Blotting applications by the "Abpromise" guarantee) and a HRP–conjugated secondary antibody, and H4 was detected by chemiluminescence (GE Healthcare ECL).

3.4.9 Yeast genetics experiments

Mutant yeast strains were generated in the BY4741 background (ATCC 4006733, MATa; *his3*Δ1; *leu2*Δ0; *met15*Δ0; *ura3*Δ0). This strain lacks the *ura3* gene that encodes an enzyme involved in uracil synthesis and can only grow in media supplemented with uracil (or uridine). The Δ*itc1* strain was derived from BY4741 by replacing (through yeast transformation and integration by homologous recombination) the *itc1* coding sequence with a *URA3* cassette that includes the *ura3* gene. Positive transformants were selected on media lacking uracil (uracil-dropout). The *URA3* cassette was generated by PCR from the pRS306 plasmid. The Δ*itc1-Nterm* strain was derived from the Δ*itc1* strain by replacing the *URA3* cassette with the *itc1* coding sequence lacking the portion that encodes residues 2-374 at the N-terminus of Itc1. DNA for transformation was prepared by PCR from a plasmid encoding Itc1 lacking the N-terminal region (ΔN-term Itc1). Positive transformants were selected on uracil-containing media additionally supplemented with 5-Fluoroorotic acid (5-FOA) for counterselection against cells that still contained the *URA3* cassette. The rescue-Δ*itc1* strain was derived from the Δ*itc1-Nterm* strain by replacing the inserted coding sequence for the ΔN-term Itc1 construct with a *URA3* cassette and selecting on uracil-dropout media. All strains were confirmed by PCR screening and sequencing at the *itc1* locus. For growth assays, liquid overnight cultures of the different strains were diluted to the same cell density, and 10-fold serial dilutions of each strain were spotted onto YPD (complete medium including uracil) plates and grown at 30°C.

3.4.10 Data analysis

Single-molecule time traces were analysed using a custom-written Matlab code. Nucleosome translocation steps were identified manually or by fitting a staircase function to the FRET time traces using a step-finding algorithm [183]. Non-linear curve fitting of the fluorescence anisotropy data was carried out in Matlab, and 95% confidence intervals for fit parameters were obtained using the "confint" routine. Experimental sample sizes as indicated in the figure captions gave the reported s.e.m. (standard error of the mean) values that were sufficiently low to allow meaningful interpretation of the data.

Chapter 4

Stepwise nucleosome translocation by RSC remodeling complexes

4.1 Introduction

The SWI/SNF-family chromatin remodelers are important regulators of chromatin structure in transcriptional activation and in the creation and maintenance of nucleosome free regions [63, 64, 65, 66, 67]. Consistent with these roles, SWI/SNF remodelers are capable of disrupting nucleosomes *in vitro* and *in vivo* in a number of ways, including repositioning the histone octamer along the DNA, ejecting H2A-H2B dimers from the octamer, and ejecting the histone octamer from the DNA [63, 64, 65, 66, 67]. All SWI/SNF complexes comprise a catalytic subunit, which harbors an SF2-family ATPase domain, and a variety of accessory subunits [63, 64]. During remodeling, the ATPase domain contacts the nucleosome at an internal site 20 bp from the nucleosome dyad, referred to as the super-helical 2 (SHL2) site, and translocates DNA [91, 95, 96, 98]. However, how DNA translocation by the ATPase domain is coupled to the various activities of the remodeling complexes remains unclear.

SWI/SNF remodelers expose substantial amounts of intra-nucleosomal DNA to nucleases during remodeling, suggesting that the remodelers disrupt the wrapping of DNA around the nucleosome [104, 105, 106, 107]. The nature of the disrupted intermediates remains unclear, and previous studies have suggested unpeeling of DNA from the edges of the nucleosome [106, 184, 185] as well as the formation of loops or bulges of DNA inside the nucleosome

[104, 107, 108, 109, 110, 111], though propagation of large loops around the nucleosome does not appear to be required for nucleosome remodeling [186].

Given that the ATPase domain of the enzymes acts at the SHL2 site, approximately 50 and 100 bp away from the sites where the DNA enters and exits the nucleosome, respectively, how ATPase activity drives the overall movement of DNA around the nucleosome remains poorly understood. The isolated catalytic subunit of a SWI/SNF remodeler has been shown to translocate naked DNA (in the absence of a histone octamer) in ~2 bp steps [187]. However, DNA translocation around the nucleosome by SWI/SNF holoenzyme complexes has been reported to occur in large steps of ~50 bp in size [96]. It is unknown whether this 50-bp step represents the fundamental step size of nucleosome translocation by SWI/SNF complexes. The related ISWI-family of chromatin remodelers move DNA in multi-bp compound steps that are composed of 1-bp fundamental steps [26, 27]. Although the ISWI-family remodelers share a homologous ATPase domain with the SWI/SNF-family remodelers, they differ in both the domains flanking the ATPase domain on their catalytic subunits and the accessory subunits that associate with the catalytic subunits [63, 64], and they bind the nucleosome differently [124, 188]. Thus, it is unclear whether SWI/SNF remodelers share a similar nucleosome translocation mechanism.

In this study, we used single-molecule FRET [15] with a variety of labeling schemes to monitor nucleosome remodeling by RSC, a prototypical SWI/SNF-family enzyme [189], in real time. These experiments allowed us to observe transient intermediates of the remodeling process and characterize the motion of the DNA into and out of the nucleosome. We found that the remodeling activity primarily involved DNA translocation around the nucleosome without large-amplitude lifting of the DNA away from the nucleosomal surface or substantial displacement of the H2A-H2B dimer, and that DNA was translocated across the nucleosome in a stepwise manner with a step size of ~1-2 bp.

4.2 Results

4.2.1 Single-molecule FRET assay for monitoring nucleosome remodeling

We reconstituted dye-labeled mononucleosomes on a double-stranded DNA containing the 601 positioning sequence [180] to ensure reproducible positioning of the histone octamer. A variety of labeling schemes with the FRET donor dye on various sites of the histone octamer and the FRET acceptor dye on various locations of the DNA were used. In the first labeling scheme, the mononucleosome was flanked by a shorter 6 bp linker on one side and a longer 78 bp linker on the other side. The FRET acceptor dye, Cy5, was attached to the end of the shorter linker and the longer linker contained a biotin at its end for attachment to the microscope slide (Figure 4.2.1a). The histone octamer was labeled with the FRET donor dye, Cy3, on the C-terminal tail of H2A (position 119). We refer to this construct as the H2A/[end, +6] construct to indicate the position of the Cy3 label (histone H2A) and Cy5 label (6 bp from the edge of the nucleosome at the end of the linker DNA). This labeling scheme allowed us to monitor the movement of DNA relative to the histone octamer in real time during nucleosome remodeling.

We anticipate that remodeling of the dye-labeled mononucleosomes by RSC will generate two distinct products, depending on which side of the nucleosome the ATPase engages (Figure 4.2.1a). If RSC engages the SHL+2 site, which we define to be the SHL2 site near the longer linker DNA, RSC will translocate DNA toward the shorter linker. Previous studies of SWI/SNF-family remodelers have shown that the enzyme can translocate DNA around the nucleosome until the end of the DNA reaches the SHL2 site, which is ~50 bp past the edge of the nucleosome [108, 190]. Therefore, this type of action should generate a ~130 bp movement of the DNA toward the shorter linker, moving the Cy5 dye away from the Cy3 dye on the octamer and causing a monotonic decrease in FRET. This action will eventually position the Cy5-labeled DNA end >40 nm from the Cy3 label on the H2A, resulting in a

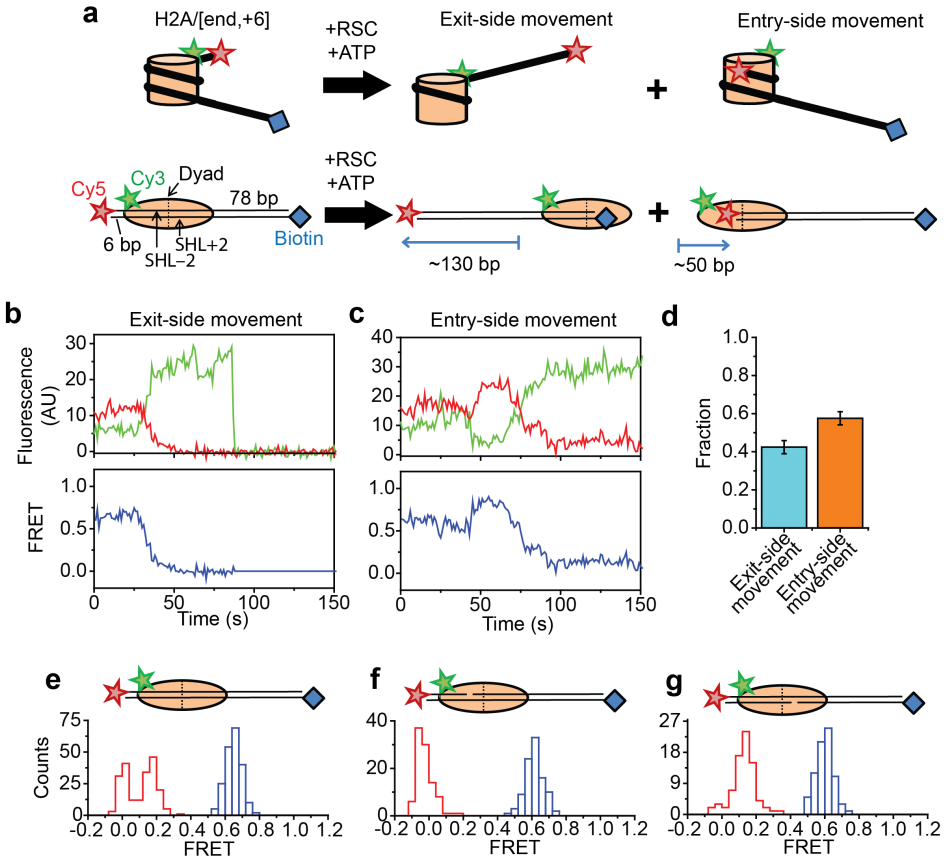


Figure 4.2.1: Single-molecule FRET assay for monitoring nucleosome translocation by RSC. **a,** Diagram depicting the nucleosome substrates before (left) and after (right) remodeling by RSC. The top row depicts the nucleosomes in cartoon form while the bottom row shows the footprint of the histone octamer (tan oval) on the DNA (black line). **b,c,** Representative traces showing the Cy3 intensity (green), Cy5 intensity (red), and FRET value (blue) for exit-side movement (b) and entry-side movement (c). [RSC] = 5 nM, [ATP] = 5 µM. **d,** The fraction of exit-side movement (cyan) and entry-side movement (orange) traces observed with 1 nM RSC and 20 µM ATP. Error bars represent the standard error from > 200 nucleosomes. **e-g,** Histograms showing the distribution of FRET values before (blue) and after (red) addition of 1 nM RSC and 20 µM ATP for nucleosomes lacking any ssDNA gap (e), nucleosomes with a 2-nt ssDNA gap at the SHL–2 site (f), and nucleosomes with a 2-nt ssDNA gap at the SHL+2 site (f).

product that gives zero FRET. On the other hand, if RSC engages the SHL–2 site, the SHL2 site near the shorter linker DNA, RSC will translocate DNA toward the longer linker, first moving the Cy5 dye closer to the Cy3 dye on the octamer and then further away from the Cy3, causing an initial increase in FRET followed by a decrease. This action will generate a final product where the labeled DNA end resides at the SHL–2 site. Based on the crystal structure of the nucleosome [59], this should place the Cy5 ~6.8 nm from the Cy3 labeling site, giving a low but non-zero FRET value.

Consistent with these expectations, we observed two major classes of traces upon addition of RSC and ATP to the surface-tethered nucleosomes. One class of traces showed a monotonic decrease in FRET to zero FRET (Figure 4.2.1b). We assigned these traces to the case where the ATPase domain of RSC bound to SHL+2 and translocated the DNA toward the shorter linker. Because the dye labels monitored the dynamics of the DNA end moving away from the octamer, we refer to these traces as monitoring exit-side movement. The second class of traces showed an initial increase in FRET followed by a decrease to a final FRET of ~0.17 (Figure 4.2.1c). We assigned these traces to the case where the ATPase domain bound to SHL–2 and translocated DNA toward the longer linker, which is expected to first bring the FRET donor and acceptor dyes closer and then farther apart. Because the dye labels in these cases monitored the dynamics of the DNA moving into the nucleosome, we refer to these traces as monitoring entry-side movement. We classified the traces as reflecting entry-side or exit-side movement based on the presence or absence of an initial FRET increase during remodeling and identified a roughly equal number traces showing entry-side and exit-side movement when RSC and ATP were added to the H2A/[end, +6] nucleosome construct (Figure 4.2.1d). This result is consistent with previous results indicating that translocation is bidirectional and does not depend on linker DNA length [108, 190, 191]. The overall kinetics of the FRET changes observed in the single-molecule assay were similar to the kinetics observed in solution-based ensemble FRET measurements, indicating that surface attachment did not substantially affect remodeling by RSC (Figure 4.2.2). As expected, no

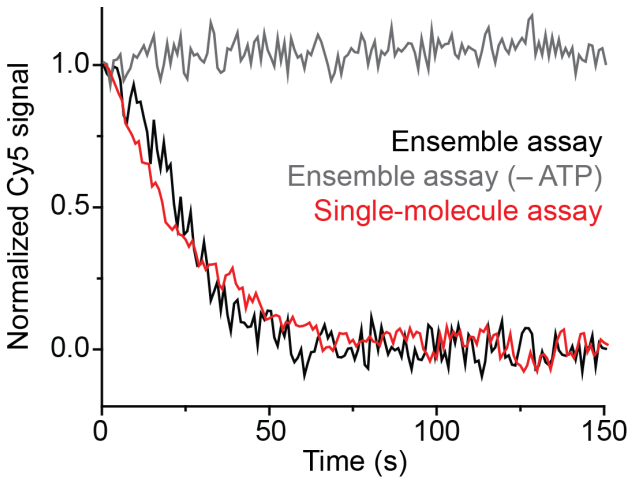


Figure 4.2.2: Surface anchoring of the nucleosomes does not affect the kinetics of remodeling. Comparison of the average remodeling kinetics for freely diffusing nucleosomes in the absence (grey, ensemble assay) or presence of ATP (black, ensemble assay), and surface-tethered nucleosomes in the presence of ATP (red, single-molecule assay). The remodeling kinetics are monitored by reading the Cy5 intensity over time and normalizing the initial and steady-state values to 1 and 0, respectively. The Cy5 intensities in the single-molecule assay are determined by summing the Cy5 signals from ~2000 single-nucleosome traces. [RSC] = 6 nM, [ATP] = 20 μM.

major FRET changes were observed when RSC was added to the nucleosomes in the absence of ATP (Figure 4.2.2).

In order to confirm our assignment of entry-side and exit-side movement traces, we made use of the fact that the ATPase of RSC is incapable of translocating past a 2-nt single-stranded (ss) gap in the DNA [95, 96]. Therefore, placing such a gap at the SHL+2 or SHL–2 site (Figure 4.2.3) should allow us to control the direction of DNA translocation around the nucleosome. After remodeling by RSC, nucleosomes without any ssDNA gap showed two distinct populations of nucleosomes with FRET values of ~0 and ~0.17, corresponding to the products of exit-side and entry-side movement, respectively (Figure 4.2.1e). As expected, remodeling of a construct containing a gap at the SHL–2 site showed a peak only at zero FRET after remodeling (Figure 4.2.1f), consistent with the gap preventing RSC from engaging the SHL–2 site to generate entry-side movement traces. Furthermore, >95% of the remodeling traces from this construct were classified as exit-side movement traces and

Top Strand (5'->3')

GCCGCCCTGGAGAATCCCGGTCTGCAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCG
CTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTC
TCCAGGCACGTGTCAGATATATACATCCTGTGCATGTATTGAACAGCGACCTTGCCGGTGC
CAGTCGGATAGTGTTCCGAGCTCCCACTCTAGAGGATCCCCGGGTACC

Bottom Strand (5'->3')

GGTACCCGGGGATCCTCTAGAGTGGGAGCTCGGAACACTATCCGACTGGCACCGGCAAGGT
CGCTGTTCAATACATGCACAGGATGTATATATCTGACACGTGCCTGGAGACTAGGGAGTAA
TCCCCTTGGCGGTTAAAACGCGGGGGACAGCGCGTACGTGCGTTTAAGCGGTGCTAGAGCT
GTCTACGACCAATTGAGCGGCCTGCAGACCGGGATTCTCCAGGGCGGC

Figure 4.2.3: Locations of the 2-nt ssDNA gaps on the nucleosomes. The sequence of the DNA used for the top and bottom strands of the H2A/[end,+6] and H3/[end,+6] nucleosomes is shown with the 601 positioning sequence colored orange. To create a gap at SHL–2, the underlined nucleotides in the top strand were replaced with a 2-nt gap. To create a gap at SHL+2, the underlined nucleotides in the bottom strand were replaced with a 2-nt gap.

showed only a monotonic FRET decrease (Figure 4.2.4a,b). Similarly, a gap at SHL+2, eliminated the zero FRET population (Figure 4.2.1g) and resulted in a single population with a FRET of ~0.14 after remodeling. Moreover, >95% of the traces observed were classified as entry-side movement traces and showed an initial FRET increase (Figure 4.2.4c,d). These results confirm that the traces classified as reflecting entry-side movement resulted from ATPase action at the SHL–2 site and the traces classified as reflecting exit-side movement resulted from ATPase action at the SHL+2 site.

4.2.2 Probing for potential H2A-H2B dimer movement during remodeling

In addition to nucleosome sliding, RSC has been reported to facilitate a number of other changes to the nucleosome, such as ejecting H2A-H2B dimers or the entire histone octamer, but we do not expect to observe these activities in our assay because they require free nucleosomes or free DNA as acceptors, the activity of histone chaperones, or a dinucleosome construct [63, 65, 66, 67]. However, it is possible that RSC could destabilize the H2A-H2B dimer [106] and lead to transient H2A-H2B dimer displacement during remodeling. To investigate whether movement of the H2A-H2B dimer (and hence the Cy3-label on H2A) relative

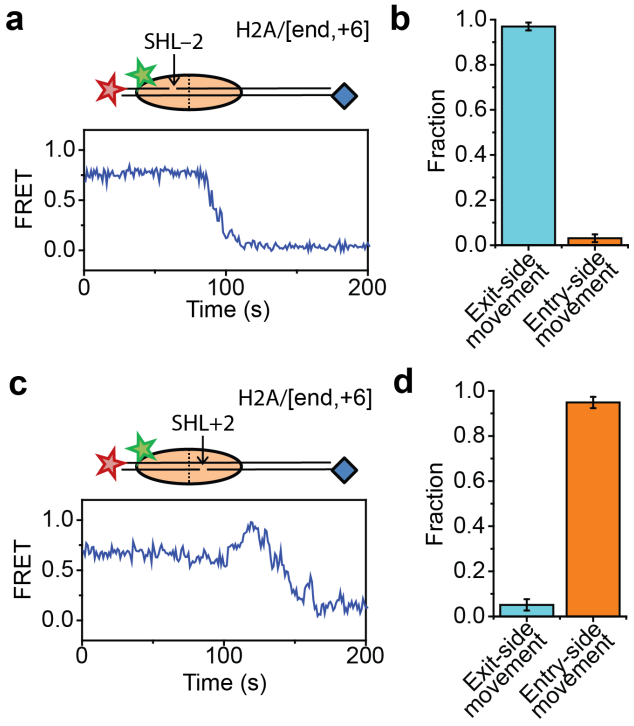


Figure 4.2.4: Effects of 2-nt ssDNA gaps at the SHL±2 sites on the direction of remodeling. **a,c,** Top: cartoon of the H2A/[end+6] nucleosome with a 2-nt ssDNA gap at the SHL–2 site (a) or SHL+2 site (c). Bottom: Representative FRET traces from the remodeling of constructs with a gap at the SHL–2 site (a) and SHL+2 site (c). [RSC] = 1 nM, [ATP] = 5 µM. **b,d,** The fraction of traces showing exit-side movement (cyan) and entry-side movement (orange) observed during the remodeling of constructs with a gap at the SHL–2 (b) or SHL+2 (d) sites. Error bars represent the standard error from > 100 nucleosomes per construct. [RSC] = 1nM, [ATP] = 20µM.

to the rest of the octamer contributed to the FRET changes we observed during RSC remodeling, we took two approaches. First, we created a H3/[end,+6] construct (Figure 4.2.5a) by moving the Cy3 label from the C-terminal tail of histone H2A to the N-terminal tail of histone H3 (position 33). Ensemble FRET remodeling assays showed similar remodeling kinetics for the H2A-labeled and H3-labeled nucleosomes substrates (Figure 4.2.6). If movement of the H2A-H2B dimer relative to the H3-H4 tetramer was responsible for the FRET changes that we observed on the H2A/[end,+6] construct, we expected that moving the Cy3 dye to histone H3 would eliminate these FRET changes. However, if H2A-H2B dimer dynamics are not involved in generating the observed FRET changes, the H3-labeled nucleosomes should produce FRET dynamics similar to those observed with the H2A-labeled nucleosomes because the N-terminal tail of histone H3 lies near the C-terminal tail of histone H2A. In our experiments, remodeling of the H3-labeled nucleosomes produced two classes of FRET traces: exit-side movement traces showing a monotonic decrease in FRET and entry-side movement traces showing an initial FRET increase followed by a FRET decrease (Figure 4.2.5b,c). As in the case of the H2A/[end,+6] constructs, placing a gap at the SHL–2 site of the H3/[end,+6] nucleosomes eliminated nearly all of the traces showing entry-side movement, and placing a gap at the SHL+2 site of the H3/[end,+6] nucleosomes eliminated nearly all of the traces showing exit-side movement (Figure 4.2.5d), confirming our assignment of entry-side and exit-side movement traces. Because the exit-side and entry-side movement traces resemble those observed with the H2A-labeled nucleosome (Figure 4.2.1b,c), this result indicates that the observed FRET changes were not specific to the dynamics of the H2A-H2B dimer.

Second, to directly probe for potential movement of the H2A-H2B dimer during remodeling, we reconstituted mononucleosomes with a 1:1 mixture of Cy3-labeled H2A and Cy5-labeled H2A on unlabeled DNA to create an H2A/H2A construct where both donor and accept dyes are on the H2A subunits of the octamer (Figure 4.2.5e). Mononucleosomes with a single Cy3 on one of the H2A subunits and a single Cy5 on the other H2A subunit

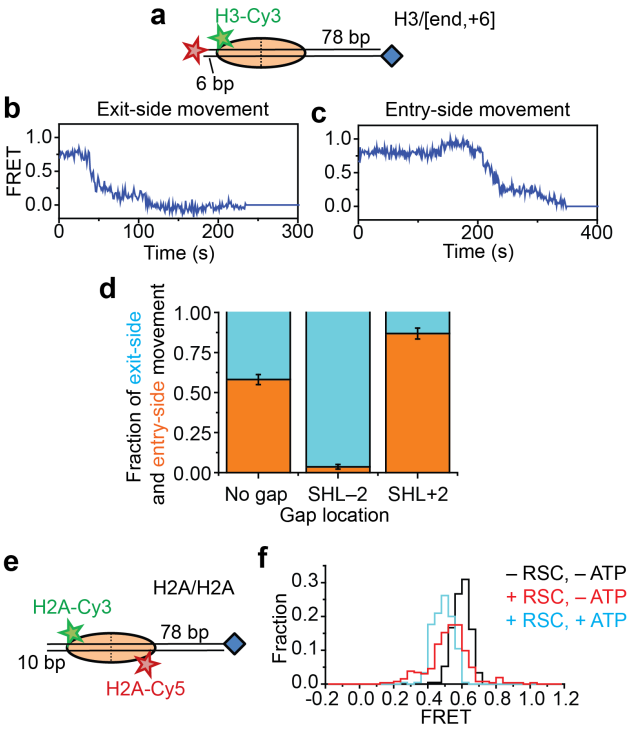


Figure 4.2.5: RSC-mediated nucleosome remodeling does not involve substantial displacement of the H2A-H2B dimer. **a,** Diagram of the H3/[end,+6] nucleosome. **b,c,** Representative FRET traces showing H3/[end,+6] nucleosomes undergoing exit-side movement (b) and entry-side movement (c). [RSC] = 1 nM, [ATP] = 2 µM. **d,** Fraction of traces showing entry-side movement (orange) and exit-side movement (cyan) observed with 1 nM RSC and 20 µM ATP for nucleosomes lacking any ssDNA gap, nucleosomes with a 2-nt ssDNA gap at the SHL–2 site, and nucleosomes with a 2-nt ssDNA gap at the SHL+2 site. Error bars represent the standard error from > 100 nucleosomes per construct. **e,** Diagram of the H2A/H2A nucleosomes. **f,** Histogram of FRET values from H2A/H2A nucleosomes in the absence of RSC (black), after the addition of 6 nM RSC (red), or after the addition of 6 nM RSC and 5 µM ATP (cyan). Histograms were constructed from > 200 nucleosomes per condition.

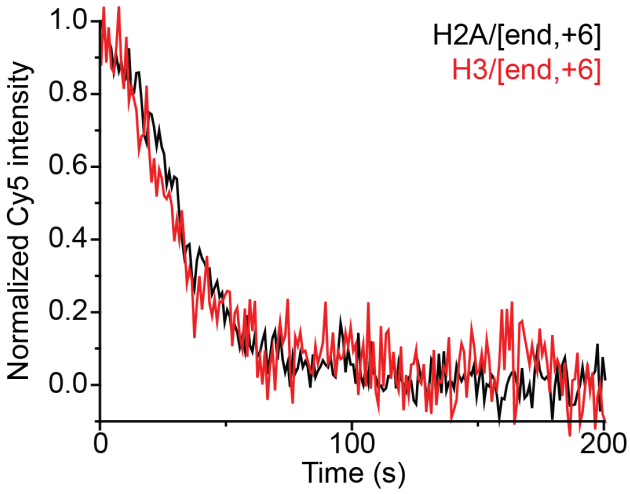


Figure 4.2.6: RSC remodeling kinetics with alternative histone labeling schemes. Comparison of the remodeling kinetics of H2A\[end,+6] (black trace) and H3\[end,+6] (red trace) nucleosomes measured in an ensemble FRET assay. The intensity traces are the average of three independent experiments. [RSC] = 6 nM, [ATP] = 20 µM.

could be readily identified at the single-molecule level and gave FRET ~ 0.6 (Figure 4.2.5f), whereas nucleosomes lacking a Cy5 showed zero FRET and nucleosomes lacking a Cy3 were not visible under green illumination. Addition of RSC and ATP to this H2A/H2A construct resulted in only a small decrease in FRET, ΔFRET ~0.1 (Figure 4.2.5f and Figure 4.2.7a). This FRET decrease was partially recapitulated by the addition of enzyme in the absence of ATP (Figure 4.2.5f), suggesting that this FRET change was associated, at least in part, with RSC binding. Furthermore, this apparent FRET change was due almost entirely to an increase in Cy3 fluorescence without a corresponding decrease in Cy5 fluorescence (Figure 4.2.7a,b), suggesting that this apparent FRET change was probably due to a change in the photophysical properties of the Cy3 dye upon RSC binding rather than a *bona fide* change in distance between the dyes. In any case, the small magnitude of this apparent FRET change was not sufficient to explain the much larger changes in FRET observed with the H2A/[end,+6] and H3/[end,+6] constructs during remodeling, indicating that our FRET assays using nucleosomes with Cy3 on the octamer and Cy5 on the DNA reported primarily on movement of the DNA relative to the nucleosome, rather than the movement of the H2A-H2B dimer relative to the rest of the octamer.

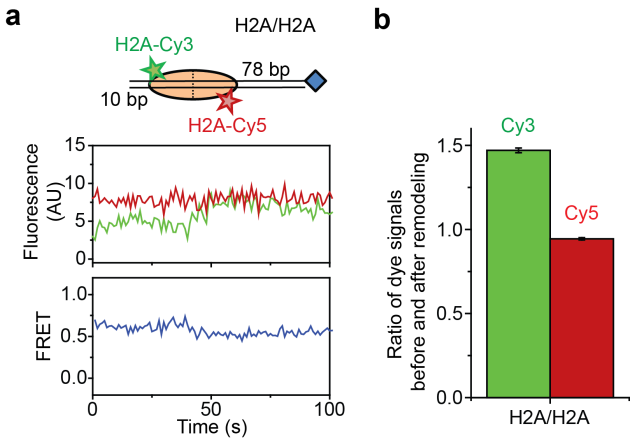


Figure 4.2.7: Remodeling of the H2A/H2A construct by RSC. **a,** Top: Cartoon of the H2A/H2A nucleosome construct. Bottom: Representative Cy3 intensity, Cy5 intensity, and FRET trace when H2A/H2A nucleosomes are incubated with 6 nM RSC and 5 µM ATP. **b,** The mean ratios of the Cy3 signals before and after remodeling (green) and the Cy5 signals before and after remodeling (red). Error bars represent the SEM from $>$ 100 nucleosomes.

4.2.3 Probing for potential DNA unwrapping at the edge of the nucleosome

Next, we asked if RSC induces large-amplitude unwrapping of DNA at the edges of the nucleosome. We separately considered this possibility for the two edges of the nucleosome where DNA enters or exits the nucleosome. At the nucleosomal edge where DNA enters the nucleosome, translocation of the DNA along its canonical path on the nucleosome would be expected to produce a FRET increase as the Cy5 at the end of the linker DNA moves toward the edge of nucleosome, followed by a FRET decrease as the Cy5 moves along the nucleosome surface toward the SHL–2 site. On the other hand, if RSC were to unpeel a substantial amount of DNA and lift the DNA off the nucleosomal surface by a large distance, we would expect the FRET traces to exhibit a FRET decrease as the Cy5 dye on the linker DNA moves away from the nucleosome, where the Cy3 dye resides. The entry-side movement traces that we observed showed a substantial increase in FRET followed by a FRET decrease for both H2A/[end, +6] and H3/[end, +6] constructs (Figures 4.2.1c, 4.2.4c, 4.2.5c, and 4.2.8a), which was consistent with movement along or near the canonical path. This phenomenon was also

seen for constructs when we increased the linker DNA length, as seen for the H2A/[end,+11] and H3/[end,+9] constructs, where the Cy5 dye-labeled DNA end was initially 11 bp and 9 bp away from the nucleosome edge (Figure 4.2.8b,c). These entry-side movement traces also showed a small apparent decrease in FRET (ΔFRET ~ 0.1) prior to the FRET increase, but this apparent FRET decrease resulted primarily from an increase in Cy3 intensity without a corresponding decrease in Cy5 intensity (Figure 4.2.8a–g). The magnitude of this Cy3 intensity change was consistent with the changes seen when we labeled the two H2A subunits on the octamer with a Cy3 and Cy5 (i.e. the H2A/H2A construct) (Figure 4.2.8g), suggesting that this FRET change likely resulted from a change in the photophysical properties of the Cy3 dye upon RSC binding as opposed to a change in the distance between the DNA and the octamer. Even if these initial FRET decreases reflected a real change in distance, the magnitude of the FRET change, ΔFRET ~ 0.1 (Figure 4.2.8d–f), was much smaller than both the FRET changes that we observed for RSC-induced DNA translocation around the nucleosome and the FRET changes previously observed for transcription factor-mediated DNA unwrapping from the edge of the nucleosome (ΔFRET ~ 0.6) on a similarly labeled nucleosome construct [192]. Assuming a previously measured Forster radius of 6 nm for the Cy3-Cy5 pair [22], these small initial FRET decreases that we observed would correspond to a distance change of only ~0.5 nm. Thus, we conclude the DNA entering into the nucleosome was not lifted by a large distance away from the nucleosomal surface and did not deviate substantially from its canonical wrapping path under our remodeling conditions.

In order to determine whether the DNA was unwrapped and lifted by a large distance off the nucleosomal surface at the nucleosomal edge where DNA exits, our current labeling scheme was not adequate since both lifting of DNA off the nucleosomal surface and translocation of DNA along the canonical path on the nucleosome would generate a decrease in FRET. We therefore moved the Cy5 label from the end of the DNA to a site 15 bp inside the edge of nucleosomes by incorporating the dye into the sugar-phosphate backbone to generate the H2A/[backbone, –15] nucleosomes, where H2A again stands for the position of the Cy3

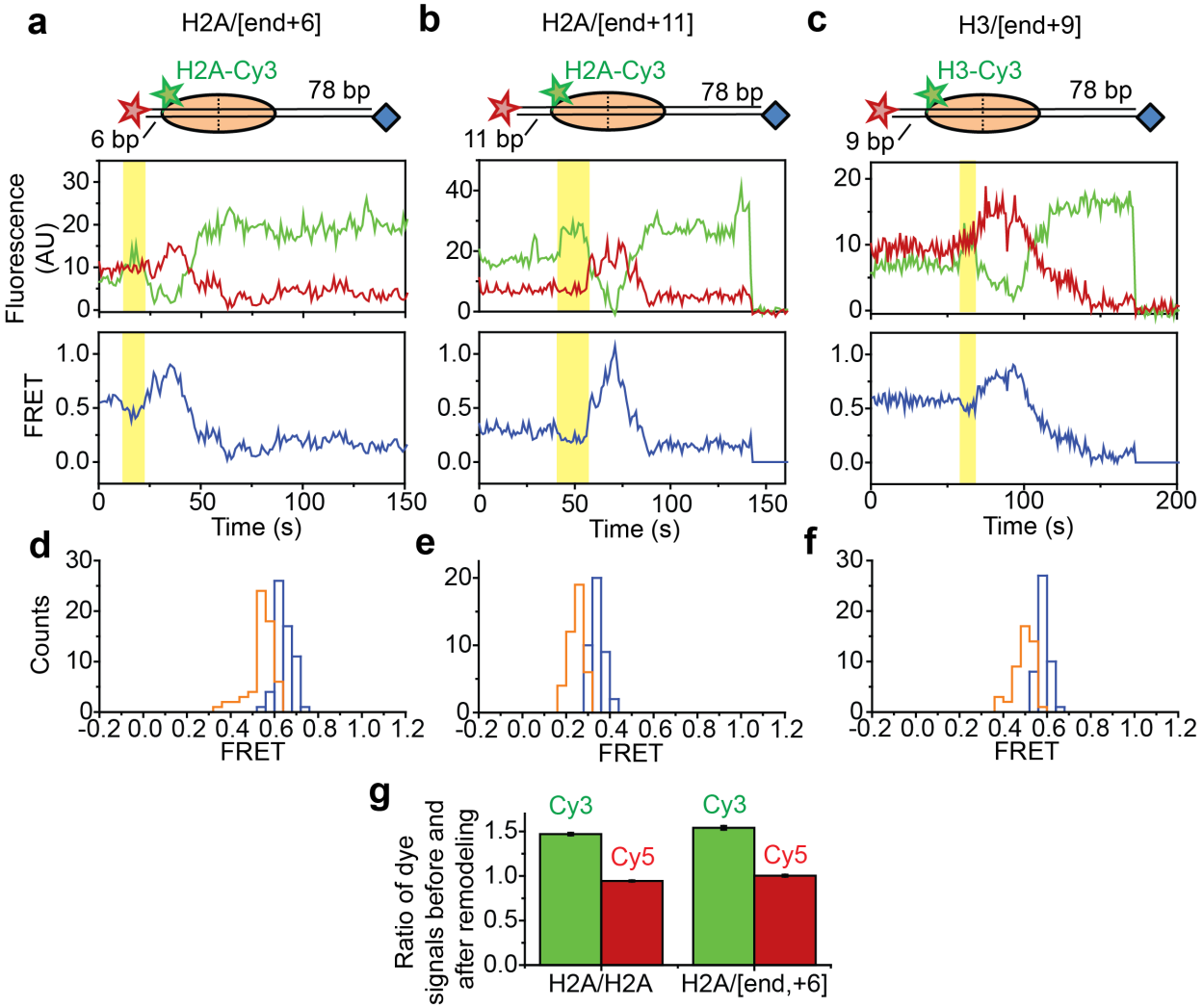


Figure 4.2.8: RSC-mediated nucleosome remodeling does not exhibit large-amplitude lifting of DNA away from the nucleosomal surface at the nucleosomal edge where DNA enters the nucleosome. **a–c,** Top: cartoons representing the H2A/[end,+6] (A), H2A/[end,+11] (b), and H3/[end,+9] (c) constructs used in each experiment. Bottom: Representative Cy3 intensity, Cy5 intensity, and FRET time traces showing entry-side movement. The region showing the transient FRET decrease is highlighted in yellow. [RSC] = 1 nM, [ATP] = 5μM. **d–f,** Histograms showing the starting FRET (blue) and FRET during the transient FRET decrease (orange) for the H2A/[end,+6] (d), H2A/[end,+11] (e), and H3/[end,+9] (f) constructs. **g,** The mean ratios of the Cy3 signals before and after the transient FRET decrease (green) and the Cy5 signals before and after the transient FRET decrease (red) for the H2A/[end,+6] constructs. The data for the H2A/H2A constructs from Figure 4.2.7b are replotted for comparison. Error bars represent the SEM from > 100 nucleosomes per construct.

label and [backbone, −15] stands for the position of the Cy5 label (Figure 4.2.9a). If RSC translocates the DNA around the nucleosome along its canonical path, the exit-side movement traces should show an initial FRET increase as the Cy5 moves toward the nucleosome edge, followed by a FRET decrease as the Cy5 exits the nucleosome. In contrast, substantial lifting of DNA away from the nucleosome surface should produce a substantial FRET decrease, as the lifting of the DNA away from the nucleosomal surface would move the Cy5 dye on the DNA farther away from the Cy3 dye on the octamer. The addition of RSC and ATP to this construct again generated two classes of traces: exit-side movement traces showing a transient increase in FRET followed by a decrease to zero FRET (Figure 4.2.9b) and entry-side movement traces showing a decrease in FRET to ~ 0.23 (Figure 4.2.9c). Without any ssDNA gap on the nucleosome, the FRET traces of the H2A/[backbone, −15] construct preferentially showed entry-side movement (Figure 4.2.9d), likely because placing the Cy5 inside the nucleosome slightly biased RSC to bind in the orientation that positioned the ATPase at the SHL−2 site. Nevertheless, when we introduced a 2-nt ssDNA gap at the SHL−2 site, nearly all of the traces showed the exit-side movement behavior with an initial increase in FRET followed the decrease to zero FRET; when a 2-nt ssDNA gap was introduced at the SHL+2 site, nearly all of the traces showed the entry-side movement behavior, displaying only a monotonic decrease in FRET (Figure 4.2.9d). These experiments confirm that the traces showing the initial FRET increase followed by a FRET decrease resulted from ATPase action at the SHL+2 site and thus represented exit-side movement. Since large-amplitude lifting of DNA away from the nucleosomal surface would have caused a FRET decrease instead of an initial FRET increase, we conclude that the DNA was not lifted away from the nucleosome surface by a large distance and did not deviate substantially from its canonical path as it exited the nucleosome.

Taken together, these observations suggest that our FRET data primarily reported on the movement of DNA along or near the canonical path at the both edges of the nucleosome. However, while our data are not consistent with RSC lifting DNA off the nucleosome surface

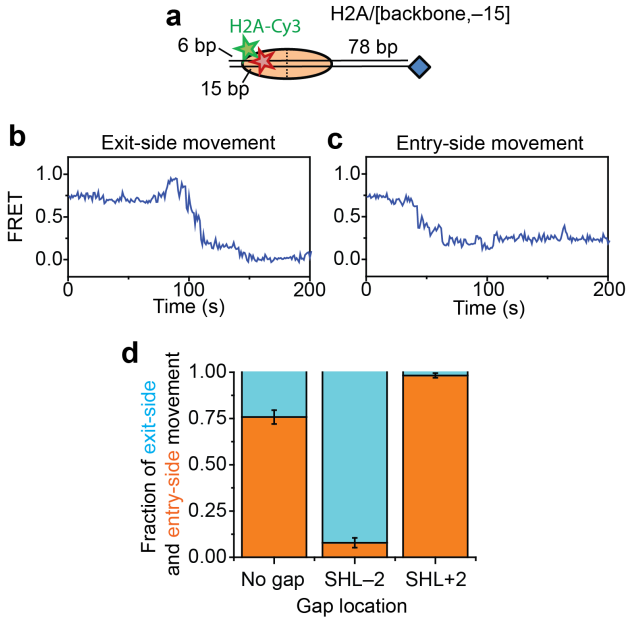


Figure 4.2.9: RSC-mediated nucleosome remodeling does not exhibit large-amplitude lifting of DNA away from the nucleosomal surface at the nucleosomal edge where DNA exits the nucleosome. **a,** Diagram of the H2A/[backbone,–15] nucleosome construct. **b,c,** Representative FRET traces showing exit-side movement (b) and entry-side movement (c). [RSC] = 1 nM, [ATP] = 5 µM. **d,** The fraction of entry-side movement (orange) and exit-side movement (cyan) traces observed with 1 nM RSC and 20 µM ATP for nucleosomes lacking any ssDNA gap, nucleosomes with a 2-nt ssDNA gap at the SHL–2 site, and nucleosomes with a 2-nt ssDNA gap at the SHL+2 site. Error bars represent the standard error from $>$ 100 nucleosomes per construct.

by large distances, these data do not exclude the possibility that the enzyme disrupts many of the histone-DNA contacts simultaneously during remodeling [106], but still holds the DNA near the surface of the nucleosome.

4.2.4 Characterization of the step size of DNA translocation

Next, we characterized the step size with which RSC translocated DNA into and out of the nucleosome. To determine the step size of DNA translocation at the nucleosomal edge where DNA exits the nucleosome, we monitored remodeling of the H2A/[end,+6] nucleosomes at 20°C (as opposed to 30°C in the previous experiments) with 2 µM ATP in order to slow the remodeling reaction and better resolve the translocation steps. The exit-side movement traces exhibited intermittent pauses interrupting the monotonic decrease to zero FRET (Figure 4.2.10a), and we applied a step-finding algorithm [183] to identify the location of the pauses and the sizes of the steps. The step size histogram showed a distribution of step sizes with a major peak centered at a ΔFRET of 0.1 and a tail extending to larger ΔFRET values (Figure 4.2.10b).

We then determined the step sizes of DNA translocation assuming that the DNA moved along the canonical path around the nucleosome. To calibrate the FRET changes associated with moving the DNA along its canonical path, we positioned Cy5 at the end of the linker DNA and varied the linker DNA length (Figure 4.2.11a, blue histograms). Consistent with previous results [26], these measurements showed an approximately linear relationship between the linker DNA length and the observed FRET value over the measured range with a slope of 0.053 ± 0.004 bp^{-1} (Figure 4.2.11b, blue). Because the photophysical effects associated with enzyme binding may affect this relationship, we also performed the calibration on enzyme-bound nucleosomes by measuring the FRET value of the first observed step during remodeling of these constructs (Figure 4.2.11a, red histograms) as a function of the initial linker DNA length. These experiments also showed a linear relationship with a similar slope of 0.047 ± 0.005 bp^{-1} (Figure 4.2.11b, red). From these slopes, we determined that the major

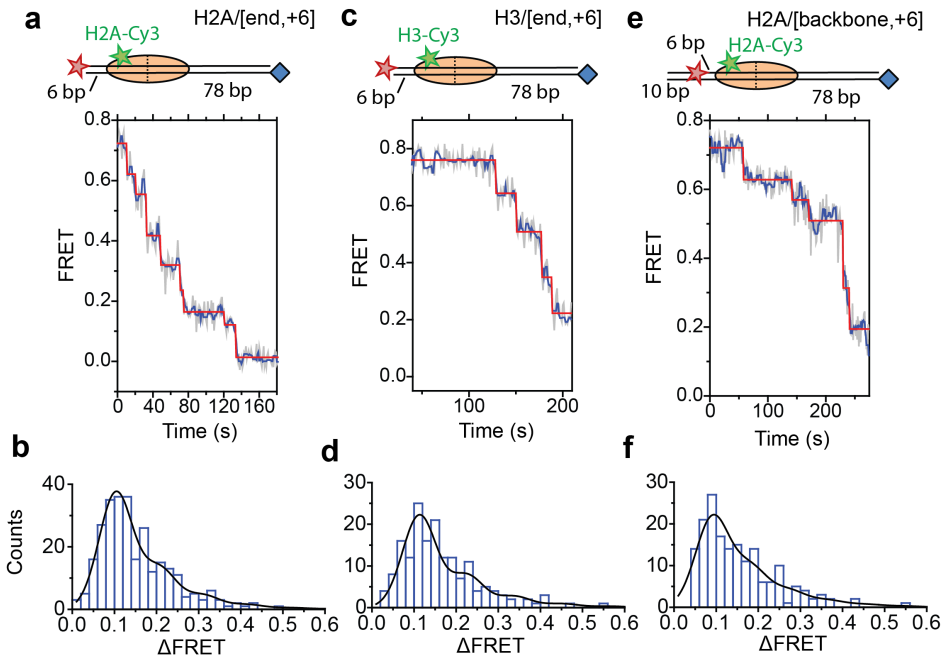


Figure 4.2.10: DNA exits the nucleosome in a stepwise manner with a step size of ~1-2 bp during RSC-mediated remodeling. Remodeling was monitored for H2A/[end,+6] (a and b), H3/[end,+6] (c and d), and H2A/[backbone,+6] (e and f) nucleosome constructs. **a,c,e,** Top: diagram of the nucleosome construct used. Bottom: Representative exit-side movement traces in the presence of 5 nM RSC and 2 µM ATP at 20°C. Light grey, raw FRET data; blue, 5-point median-filtered data; red, fit by a step-finding algorithm [183]. **b,d,f,** Histograms of the measured step sizes in FRET change (blue bars) and the fit to the modeled step size distribution shown in Eq. 4.4.1 (black line).

step size of DNA movement, corresponding to the peak of the ΔFRET distribution in Figure 4.2.10b (ΔFRET ~ 0.1) to be approximately 2 bp.

The step size histogram, however, showed many steps that were larger than expected for ~2 bp of DNA motion (Figure 4.2.10b). Although these data could reflect a heterogeneous step size, the distribution could also result from movement with a single step size, where the larger steps represented two or more steps where the intervening pause(s) were too short to detect. To test whether our data were consistent with such a model, we took advantage of the linear dependence of FRET on the linker DNA length and fit our data to a series of evenly spaced Gaussian peaks whose amplitude decreased by a factor of f from the previous peak, representing the probability of missing a step (see Section 4.4.6). The data are consistent with a step size of ΔFRET = 0.10 ± 0.04 with a probability of missing a step being f = 34% (Figure 4.2.10b and Figure 4.2.12a). This probability of missing a step is reasonable given that the steps occur stochastically, and many pauses could be too short to be observed. Based on the distribution of observed pause durations, we would expect a 25% probability of missing a step due to our limited time resolution (Figure 4.2.12b). The observed probability of missing a step was slightly larger than that expected from our time resolution limitation, suggesting the possibility of additional mechanisms for missing steps, for example, successive DNA distortions generated by the ATPase at the SHL2 site merging together while transiting to the edge of the nucleosome.

To test whether the step size determination was sensitive to the labeling scheme, we repeated the experiments with a second nucleosome construct where we moved the Cy3 dye from histone H2A to histone H3 (the H3/[end, +6] construct, Figure 4.2.10c,d) and a third construct where we moved the Cy5 to the DNA backbone (instead of the DNA end) 6 bp from the nucleosome edge (the H2A/[backbone, +6] construct, Figure 4.2.10e,f). Calibration curves for these two labeling schemes showed that the slopes of the FRET versus linker DNA length were 0.059 ± 0.005 bp^{-1} and 0.057 ± 0.002 bp^{-1}, respectively (Figure 4.2.11c–f, blue). Like for the H2A/[end, +6] construct, these values did not change appreciably in the

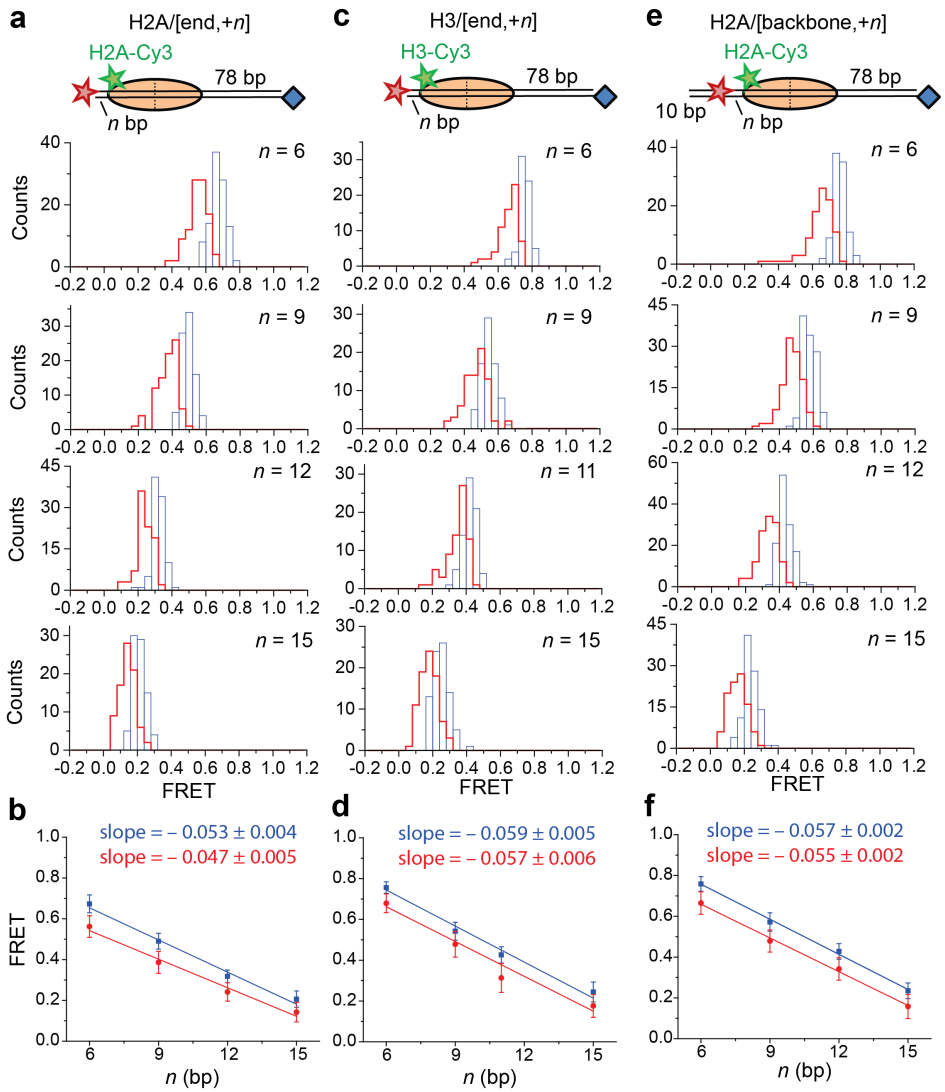


Figure 4.2.11: Calibration of FRET values as a function of the linker DNA length for constructs monitoring exit-side movement. **a,c,e,** Distribution of FRET values measured with 5 nM RSC and 2 µM ATP at 20°C for the H2A/[end,+n] (a), H3/[end,+n] (c), and H2A/[backbone,+n] (e) constructs before (blue) and after (red) the first observed remodeling step. **b,d,f,** Plots of the mean FRET value versus linker DNA length for the H2A/[end,+n] (b), H3/[end,+n] (d), and H2A/[backbone,+n] (f) constructs before (blue) and after (red) the first observed remodeling step. The mean and standard deviation of each point were obtained by fitting the FRET distributions to a Gaussian curve. The slopes were obtained by linear regression and the error is the standard error.

a

Construct	H2A/ [end,+6]	H3/ [end,+6]	H2A/ [backbone,+6]	H2A/ [end,+12]	H3/ [end,+9]
Fraction missed (f)	0.34	0.35	0.42	0.22	0.17
ΔFRET	0.10 ± 0.04	0.11 ± 0.04	0.09 ± 0.04	0.09 ± 0.03	0.08 ± 0.03
Slope (bp^{-1})	0.053 ± 0.004	0.059 ± 0.005	0.057 ± 0.002	0.055 ± 0.002	0.048 ± 0.004
Step size (bp)	1.9 ± 0.8	1.9 ± 0.7	1.6 ± 0.7	1.6 ± 0.5	1.7± 0.7

b

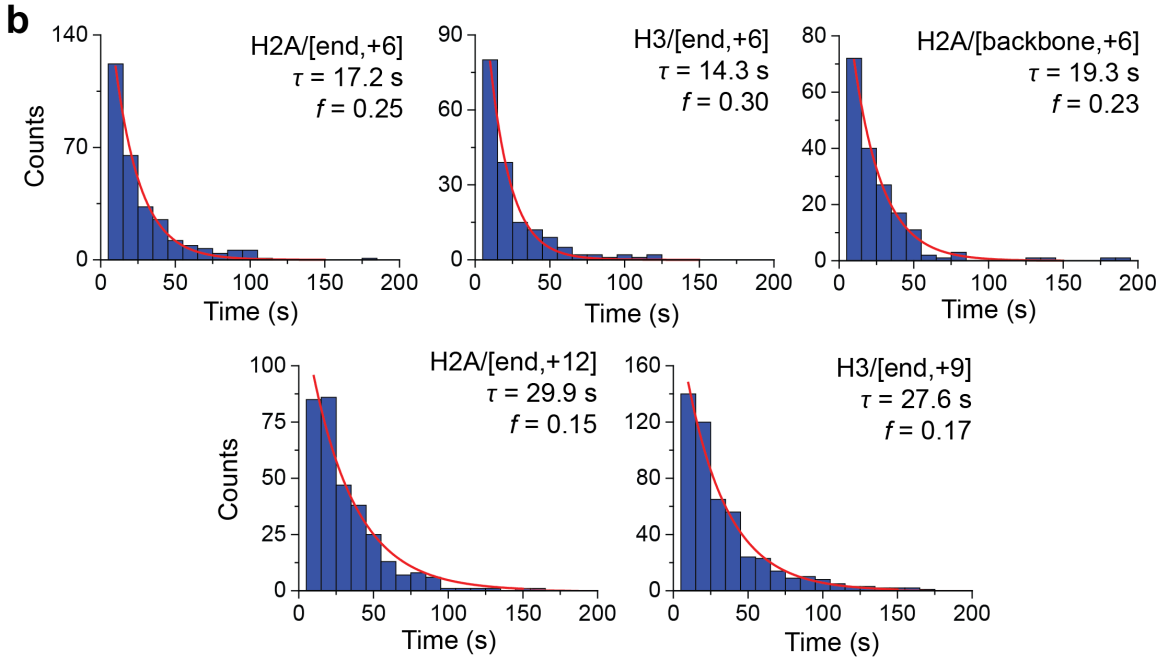


Figure 4.2.12: Analysis of DNA translocation step sizes of exit-side and entry-side movement using various labeling schemes. **a,** Table showing the estimated step size (± SD) and fraction of steps missed (f) from fits of the step size histograms to the stepping model in Eq. 4.4.1. These ΔFRET values are converted to step sizes in bp by dividing the FRET step sizes by slope of the calibration curves in Figure 4.2.11 for the experiments monitoring exit-side movement and Figure 4.2.14 for the experiments monitoring entry-side movement. **b,** Histograms of the observed pause lifetimes identified by the step-finding algorithm. The mean pause lifetime (τ) is estimated by fitting the distribution to an exponential, and the fraction of missed events (f) is the fraction of pauses expected to have a pause lifetime < 5 s (the threshold for detecting a step in the step-size analysis in Figures 4.2.10 and 4.2.13).

presence of a bound RSC enzyme (Figure 4.2.11c–f, red). The exit-side movement traces of these new constructs also showed that DNA translocation occurred in a stepwise manner (Figure 4.2.10c,e), and the step size distributions (Figure 4.2.10d,f) showed a major peak around a ΔFRET of 0.1 with a tail extending to larger values. Fitting these histograms to multiple Gaussian peaks as described earlier yielded a step size of ΔFRET = 0.11 ± 0.04 with a missed step probability of f = 35% for H3/[end, +6] and ΔFRET = 0.09 ± 0.04 and f = 42% for H2A/[backbone,+6] (Figure 4.2.12a,b). Comparing the ΔFRET values to the slopes of the calibration curves gave step size estimates of 1.9 ± 0.7 bp and 1.6 ± 0.7 bp for the H3/[end, +6] and H2A/[backbone, +6] constructs, respectively, consistent with the 1.9 ± 0.8 bp step size estimate for the H2A/[end, +6] construct.

Finally, we characterized the step size of DNA translocation at the nucleosomal edge where DNA enters the nucleosome. In the experiments that characterized the step sizes of exit-side motion, the dye-labeled linker DNA was "upstream" (to the left) of the 601 sequence (Figure 4.2.10a,c,e), so exit-side movement traces resulted from ATPase action at the SHL+2 site. Given that the DNA sequence is not symmetric around the nucleosome dyad, to maintain the same direction of DNA translocation around the nucleosome but move the dye locations to monitor the movement of DNA into the nucleosome, we made H2A/[end, +12] and H3/[end, +9] constructs where the dye-labeled linker DNA lied "downstream" (to the right) of the 601 sequence (Figure 4.2.13a,b). The entry-side movement traces primarily showed an increase in FRET as expected from the movement of the short linker DNA into the nucleosome, bringing the Cy5 dye closer to Cy3. As before, the entry-side movement traces showed an initial, small apparent decrease in FRET before the major FRET increase, likely due to the photophysical effect on the dye upon enzyme binding (Figure 4.2.8). The FRET increase occurred in a stepwise manner, indicating stepwise DNA translocation into the nucleosome (Figure 4.2.13a,b). The distributions of the ΔFRET step sizes again showed a major peak around ~0.1 with a tail extending to larger values (Figure 4.2.13c,d). Fitting to multiple evenly spaced Gaussian peaks as described earlier yielded a step size of ΔFRET

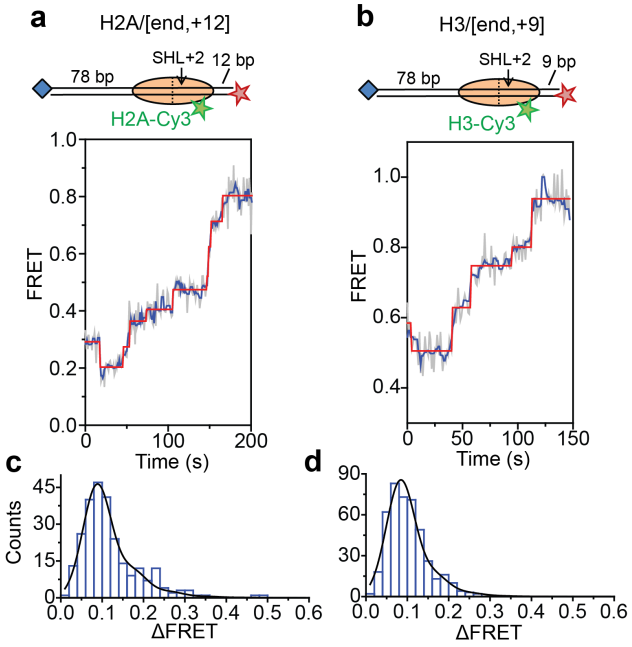


Figure 4.2.13: DNA enters the nucleosome in a stepwise manner with a step size of ~1-2 bp during RSC-mediated remodeling. Remodeling was monitored for H2A/[end,+12] (a and c), and H3/[end,+9] (b and d) constructs. **a,b,** Top: diagram of the nucleosome constructs used. Bottom: Representative entry-side movement traces in the presence of 5 nM RSC and 2 μM ATP at 20°C. Light grey, raw FRET data; blue, 5-point median-filtered data; red, fit by the step-finding algorithm. **c,d,** Histograms of the measured step sizes in FRET change (blue bars) and the fit to the modeled step size distribution shown in Eq. 4.4.1 (black line).

= 0.09 ± 0.03 with a missed step probability of f = 22% for the H2A/[end, +12] construct and ΔFRET = 0.08 ± 0.03 and f = 17% for the H3/[end, +9] construct (Figures 4.2.13c,d and 4.2.12). Assuming that DNA moved along the canonical nucleosomal wrapping path as it entered the nucleosome, we constructed calibration curves for the two labeling schemes by varying the initial linker DNA lengths and measured slopes of the FRET change versus linker DNA length to be 0.055 ± 0.002 bp^{-1} for the H2A/[end, +12] construct and 0.048 ± 0.004 bp^{-1} for the H3/[end, +9] construct, which again did not change appreciably in the presence of bound enzyme (Figure 4.2.14). Based on these slopes, we found the DNA translocation step sizes to be 1.6 ± 0.5 and 1.7 ± 0.7 bp, respectively, for these constructs. Therefore, RSC translocated DNA in 1-2 bp increments at both the entry side and exit side of the nucleosome.

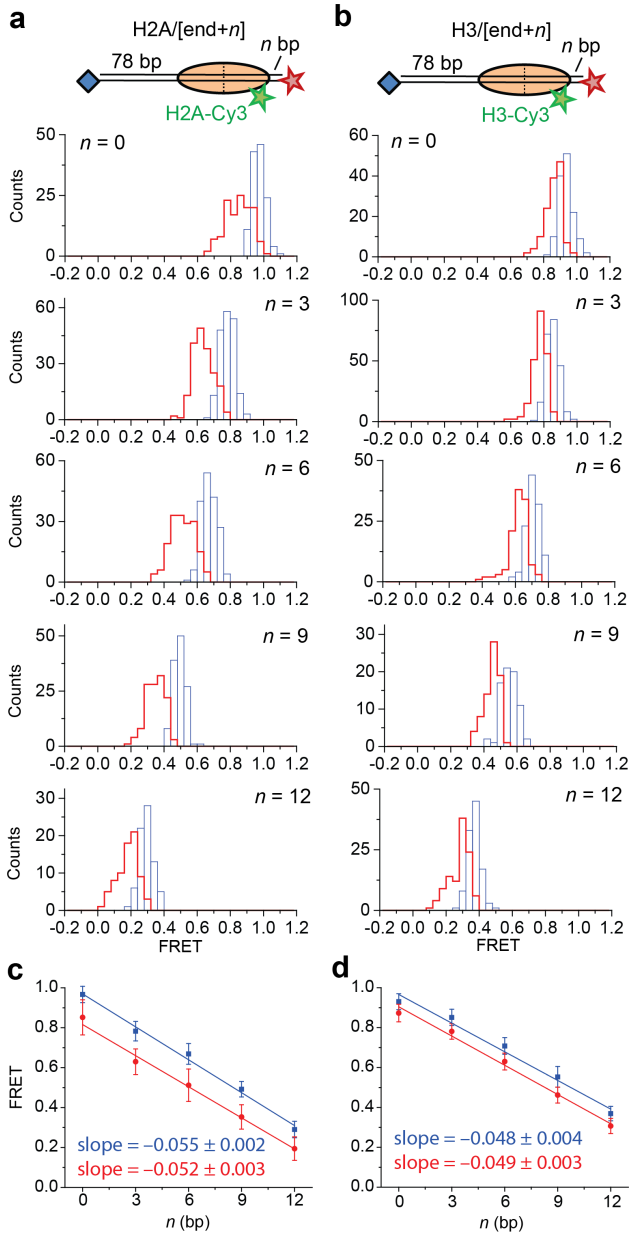


Figure 4.2.14: Calibration of FRET values as a function of the linker DNA length for constructs monitoring entry-side movement. **a,b,** Distribution of FRET values before (blue) and after (red) the first remodeling step in the presence of 5 nM RSC and 2µM ATP at 20°C for the H2A/[end,+n] (a) and H3/[end,+n] (b) constructs. **c,d,** Plots of the mean FRET versus linker DNA length for the H2A/[end,+n] (c) and H3/[end,+n] (d) constructs before (blue) and after (red) the first observed remodeling step. The mean and standard deviation for each point were obtained by fitting the FRET distributions to a Gaussian curve. The slopes were obtained by linear regression and the error is the standard error.

4.3 Discussion

In this study, we used single-molecule FRET to monitor nucleosome remodeling by a prototypical SWI/SNF-family remodeler, RSC. This approach enabled us to track the motion of DNA across individual nucleosome in real time, providing new insight into the mechanisms by which RSC repositions nucleosomes along DNA.

Our results showed that RSC remodeled mononucleosomes primarily by translocating DNA around the nucleosome. At the nucleosome edge, RSC did not lift the DNA by a large distance away from the nucleosomal surface under our remodeling conditions, and our data are consistent with translocation of the DNA along or close to the canonical nucleosomal wrapping path during RSC-mediated remodeling. We observed a small FRET decrease preceding DNA translocation (Figure 4.2.8) that was likely due to changes in the photophysical properties of Cy3 upon enzyme binding. Similarly, a construct designed to report on the dynamics of the H2A-H2B dimer also showed only a small apparent change in FRET likely due to photophysical changes of Cy3 upon enzyme binding. Even if these FRET changes were due to actual distance changes, rather than photophysical changes of the dye, they would represent minor deviations (~0.5 nm) of DNA from the canonical wrapping path around the octamer or small repositioning of the H2A-H2B dimer relative to the H3-H4 tetramer. Although we can exclude large-amplitude lifting of the DNA from the surface of the octamer, smaller distance changes could still disrupt many histone-DNA contacts, so we cannot exclude the possibility that the observed translocation of DNA around the nucleosome is associated with simultaneous disruption a substantial number of histone-DNA contacts. However, such a small-distance motion would have only a minimal effect on our measured FRET signal.

Thus, we characterized the step size of DNA translocation into and out of the nucleosome by assuming that the DNA moved along the canonical path around the nucleosome during RSC-mediated remodeling. We observed a step size of 1-2 bp both for entry-side and exit-side movement. Even if the motion of the DNA is not exactly along the canonical path,

the small ΔFRET step sizes of 0.08-0.11 that we observed in the single-molecule nucleosome remodeling traces corresponds to a distance change of ~0.3-0.6 nm (assuming a previously measured Förster radius of 6 nm for the Cy3-Cy5 FRET pair [22]), which is again consistent with translocation of 1-2 bp of dsDNA (0.34 nm/bp). Although changes to the photophysical properties of the dyes or the relative orientation of the dyes could potentially affect this distance calibration, many of these effects would be expected to depend on the specific labeling scheme. The fact that we observed these 1-2 bp step sizes with several different attachment sites of the Cy3 dye on the octamer and different attachment sites of the Cy5 dye on the DNA suggests that these numbers likely reflect the true translocation step size of DNA around the nucleosome during RSC-catalyzed remodeling. It has been shown previously that some SF2 ATPase-containing DNA translocases or helicases translocate DNA in 1 bp increments [27, 193, 194, 195, 196], so the SF2 ATPase domain of RSC might also translocate DNA at the SHL2 site of the nucleosome in 1 bp steps. The reason that the step size that we observed for RSC-mediated nucleosome translocation is not exactly 1bp could be due to the following: 1) missing steps due to our limited time resolution could cause two consecutive 1 bp steps appear as a 2 bp step, 2) two 1 bp DNA distortions generated at the SHL2 site could merge together while transiting to the exit site, and 3) the DNA translocation path could deviate slightly from the canonical wrapping path, making our calibration, which was determined based on canonically wrapped nucleosomes, slightly off. Because we track the motion of the DNA only at the nucleosomal edges, our data cannot determine whether intranucleosomal loops or bulges form, though our data agree with previous findings that if such loops do exist, they must accumulate and release gradually.

The small step size that we observed here is consistent with a previous study reporting that the isolated RSC catalytic subunit translocates naked DNA in ~2 bp increments [187]. However, in the context of remodeling intact nucleosomes, a previous biochemical study showed that SWI/SNF enzymes translocate DNA around the nucleosome in large, ~50 bp steps [96]. Our current observations suggest that these large ~50 bp translocation steps do

not reflect the fundamental step size of nucleosome translocation by RSC. Rather, they may be compound steps composed of many 1-2 bp steps and reflect relatively long kinetic pauses of DNA translocation imposed by the energy landscape of the nucleosomal substrates. Because the SF2 ATPase likely translocates DNA 1 bp at a time at the SHL2 site, our observation that RSC moves the DNA into and out of the nucleosome in 1-2 bp increments suggests that motion at the nucleosomal entry and exit sites are directly coupled to DNA translocation at the SHL2 site by the ATPase domain and that the other domains of the catalytic subunit or the accessory subunits of the enzyme do not play an active role in translocating DNA around the nucleosome.

4.4 Materials and methods

4.4.1 Histone purification and labeling

Recombinant *X. laevis* histones were expressed in BL21 (DE3) pLysS cells (Stratagene) and purified under denaturing conditions. Briefly, we isolated inclusion bodies and extracted the histones as described previously [181], then dialyzed the histones into buffer A (7 M urea, 20 mM Na-HEPES pH 8.0, 100 mM NaCl, 1 mM DTT and 1 mM Na-EDTA). This solution was then loaded onto HiTrap-Q cation exchange and ResourceS anion exchange columns (GE Healthcare) connected in series [197]. After washing with buffer A, the HiTrap-Q column was then removed before the protein was eluted by gradually increasing the concentration of NaCl. To generate histones site-specifically labeled with Cy3, plasmids for the expression of H2A K119C, and double mutant H3 (G33C+C110A) were created by site directed mutagenesis. These constructs were purified and labeled with sulfo-Cy3 maleimide or sulfo-Cy5 maleimide (GE Healthcare) under denaturing conditions [127]. These sites were chosen because they were used in previous FRET-based studies of nucleosome remodeling [114, 115]. Histone octamer was reconstituted with an ~ 1:1 ratio of labeled:unlabeled histone in order to maximize the yield of singly-labeled octamer and purified by gel filtration chromatogra-

phy as described previously [127, 181]. For the reconstitution of the H2A/H2A construct, histone octamer was reconstituted with a ~ 1:1 ratio of Cy3-labeled:Cy5-labeled H2A.

4.4.2 Preparation of DNA constructs

DNA constructs were made by PCR from a plasmid containing a modified 601 positioning sequence [180, 198]. The PCR primers contained 5' Cy5 or biotin-TEG modifications (IDT) to install these modifications at the indicated locations. For constructs containing gaps or backbone Cy5 labels, the DNA was constructed by annealing a set of overlapping oligonucleotides (IDT) and ligating them into a single double-stranded DNA construct [127]. Backbone Cy5 labels were inserted opposite guanosine residues. DNAs were purified by PAGE.

4.4.3 Preparation of nucleosome constructs

Mononucleosomes were assembled by the salt dialysis method after mixing labeled octamer and DNA at a 1.2:1 ratio and purified by glycerol gradient centrifugation as previously described [199].

4.4.4 RSC purification

RSC was purified from strain BCY211 that expresses a TAP-tagged Rsc2 subunit following published protocols [197]. Protein concentration was quantified by Sypro Red staining and comparison with BSA standards.

4.4.5 Single-molecule FRET measurements

Quartz slides were cleaned, functionalized with a 1:100 mixture of biotin-PEG:PEG (Laysan Bio), and assembled into flow chambers as previously described [20]. Nucleosomes were immobilized to biotin-PEG via streptavidin. The sample was excited with a 532 nm Nd:YAG

laser (CrystalLaser) on a custom-built total-internal-reflection fluorescence (TIRF) microscope [200]. Fluorescence emission was collected with a 60x water immersion objective (Olympus), filtered with at 550 nm long-pass filter (Chroma Technology), split with a 630 nm dichroic mirror (Chroma Technology), and imaged onto two halves of an Andor iXon+888 EM-CCD camera with a 1 Hz integration time. Imaging was done at 30°C (unless otherwise specified) in imaging buffer (40 mM tris, 12 mM HEPES, pH 7.5, 60 mM KCl, 3 mM $MgCl_2$, 10% (v/v) glycerol, 0.02% (v/v) Igepal CA-630, 10% (w/v) glucose, 2 mM trolox, 0.1 mg/mL acetylated BSA), supplemented with an oxygen scavenging system (800 µg/mL glucose oxidase, 50 µg/mL catalase) [26, 167]. In experiments where the FRET acceptor dye is on the DNA and the FRET donor is on the histone, we see three populations of labeled nucleosomes: singly-labeled nucleosomes where the dye resides on the histone subunit proximal to the short linker, singly-labeled nucleosomes where the dye resides on the histone subunit distal to the short linker, and the doubly-labeled species [26]. These nucleosomes can be distinguished on the basis of their FRET value, and we select only those nucleosomes with the Cy3 on the proximal histone for analysis, which gave the highest FRET values. For the H2A/H2A construct, we selected those showing FRET > 0.2 as nucleosomes lacking Cy5 show zero FRET and nucleosomes lacking Cy3 are not visible during 532 nm illumination.

4.4.6 Data analysis

Data were analyzed with IDL (ITT Visual Information Solutions) and MATLAB (MathWorks) using custom-written code. Single nucleosomes were identified by selecting traces that showed one-step photobleaching. The fluorescence intensity after photobleaching was used for background subtraction. To classify traces as entry-side or exit-side movement traces, we manually identified traces exhibiting remodeling, smoothed the data with a 3 pt median filter, then classified them as follows: If a trace spent 3 consecutive points 0.1 FRET above the initial FRET (0.07 FRET above the initial FRET for the H2A/[backbone,–15] data) before reaching the final FRET value, indicating an increase in FRET followed by

FRET decrease, it was classified as an entry-side movement trace for the H2A/[end,+6] and H3/[end,+6] constructs or an exit-side movement trace for the H2A/[backbone,–15] construct. If the trace reached the final FRET value without spending 3 consecutive frames 0.1 or 0.07 FRET above the initial FRET, indicating a monotonic FRET decrease, the trace was classified as an exit-side movement traces for the H2A/[end,+6] and H3/[end,+6] constructs or an entry-side movement trace for the H2A/[backbone,–15] construct.

In the experiments to determine the step size of remodeling, pauses in the FRET traces were identified using a previously developed step-finding algorithm [183], and pauses shorter than 5 s were ignored. We analyzed the step sizes only for those steps occurring within the range where FRET appears to vary linearly with distance from the nucleosome (as determined by the calibration curves in Figure 4.2.11 and Figure 4.2.14). To simulate the step size distribution resulting from a uniform step size and a fixed probability of missing short pauses, causing larger steps whose size is a multiple of the step size, the observed step size histograms were fit to the function:

$$y = \sum_{n=1}^{6} A\, f^{n-1} \exp\left(-\frac{(x - n\,c)^2}{2s^2}\right) \tag{4.4.1}$$

where f is the probability that a pause is missed, c is the step size, and s is the standard deviation of the step size. The s parameter was fit globally across all step size histograms showing the same type of motion (i.e. exit-side movement or entry-side movement).

4.4.7 Ensemble FRET measurements

Ensemble FRET measurements were performed by monitoring the Cy5 intensity at 670 nm under 532 nm excitation in a Cary Eclipse Fluorescence Spectrophotometer (Varian). Reactions were performed in imaging buffer, and initiated by the addition of RSC and ATP to a solution containing nucleosomes. Data were normalized by scaling the initial Cy5 intensity to 1.0 and the final steady state Cy5 intensity to 0.

Chapter 5

Conclusions and future directions

Single molecule FRET is a powerful technique for studying the structural dynamics of biomolecules, and in this dissertation, I have demonstrated its use in understanding the mechanism and regulation of two classes of molecular motor proteins: reverse transcriptases and chromatin remodeling enzymes.

In Chapter 2, we used a smFRET assay to study the initiation of reverse transcription in order to determine the mechanism of RT pausing during this process. The capability of smFRET to characterize heterogeneity within a sample allowed us to observe RT binding to its tRNA/vRNA substrate in two different orientations: a polymerase-competent mode that places the polymerase active site over the 3' end of the tRNA primer, and an orientation flipped by 180°, where the polymerase active site is far from the 3' end of the primer. The equilibrium between these two states correlates with the polymerase activity of the enzyme—on substrates where RT spends most of its time bound in the polymerase orientation, it exhibits its highest rate of nucleotide addition, and on substrates where RT spends most of its time bound in the flipped orientation, it exhibits the slowest rate of nucleotide addition—indicating that the binding orientation of the enzyme is a major factor influencing the pauses that occur during the initiation of reverse transcription. When we examined which structural features of the tRNA/vRNA substrate regulate the binding orientation of the enzyme, we found that a stem-loop structure in the vRNA forces RT to bind predominantly in the flipped orientation at the +3 position, causing the major pause at this position. The

stem-loop also defines the transition to a faster, more processive mode of synthesis after the addition of the 6th nucleotide as this stage is the point at which RT has synthesized enough DNA to disassemble the stem-loop. These results provide a mechanistic explanation for the changes of activity observed during the initiation of reverse transcription and show how the structural dynamics of a ribonucleoprotein complex can regulate the activity of a molecular motor enzyme.

In Chapter 3, we used smFRET to study the regulation of the human ISWI remodeling enzyme, ACF. smFRET enabled us to characterize substeps of the remodeling process in order to determine which of the stages of nucleosome remodeling—DNA translocation or exit from the pause phases interrupting DNA translocation—are regulated by the H4 tail and linker DNA. These experiments, plus experiments characterizing various mutants of the proteins involved, show that the AutoN region, a part of the Snf2h catalytic subunit that inhibits the ATPase activity of the remodeler; the histone H4 tail, a portion of the nucleosome that helps relieve AutoN inhibition; and the N-terminal region of the accessory subunit Acf1, a region involved in sequestering the H4 tail in a linker length dependent manner, all contribute to the regulation of ACF by linker DNA length. Together, these results lead to the following model for the regulation of ACF by linker DNA length: remodeling occurs in a series of alternating translocation and pause phases, and linker length sensing occurs during the pause phase, where it regulates the exit from the pause phase and the subsequent onset of DNA translocation. The N-terminal region of Acf1 senses linker DNA and allosterically communicates the absence or presence of the linker DNA to the ATPase by binding or releasing to the H4 tail. When the linker DNA is short, Acf1 sequesters the H4 tail, leaving AutoN to inhibit the ATPase and prevent exit from the pause phase. When the linker DNA is long, Acf1 instead binds the linker DNA, leaving the H4 tail to displace AutoN, activate the ATPase, and allow DNA translocation to occur. These results provide a mechanistic picture of the intra- and inter-molecular interactions that take place during the mechanochemical cycle of a motor enzyme to regulate the catalytic activity of the motor.

In Chapter 4, we applied smFRET to understand the mechanism of nucleosome remodeling by the SWI/SNF family of chromatin remodeling enzymes. Here, the ability of smFRET to distinguish between different reaction pathways was essential because remodeling by RSC could proceed in one of two directions. While experiments that rely on ensemble averaging would have been confounded by the two remodeling directions, smFRET enabled us to analyze the two types of movement (exit-side movement and entry-side movement) separately. Furthermore, we made use of the ability of smFRET to report on the structure of transient intermediates to characterize the motion of DNA into and out of the nucleosome. These experiments reveal that DNA moves around the nucleosome largely along its canonical path without large-scale lifting of the DNA away from the nucleosome or substantial displacement of the H2A-H2B dimer. DNA enters and exits the nucleosome in a stepwise manner, with step sizes of 1-2 bp, suggesting that DNA movement is coupled directly to DNA translocation by the ATPase domain. These results provide important insights into the mechanism of nucleosome remodeling by SWI/SNF family remodeling enzymes, and demonstrates the power of smFRET in characterizing the structure and dynamics of enzyme-bound intermediates.

I will conclude with a discussion of how to build upon the results presented in this dissertation, focusing on chromatin remodeling by the SWI/SNF family enzymes. In particular, I will discuss experiments to further elucidate the mechanism of DNA motion, examine the mechanism of nucleosome disassembly, and monitor more complex remodeler activities.

The data presented in Chapter 4 show DNA moving into and out of the nucleosome in 1-2 bp increments. These observations are inconsistent with previously proposed models in which the remodeler scoops in tens of base pairs at a time to generate a loop that propagates around the nucleosome [96]. However, these data do not rule out the formation of intranucleosomal loops as loops could still form if the onset of exit-side translocation is delayed relative to the onset of entry-side translocation. Therefore, determining the coordination between entry-side and exit-side DNA translocation is required to gain further insight into the mechanism of nucleosome remodeling. A three-color FRET labeling scheme [201], in which

the nucleosome is labeled with Cy3, Cy5 and Alexa750 (Figure 5.0.1), should enable the simultaneous monitoring three distances (the Cy3-Cy5 distance, the Cy3-Alexa750 distance, and the Cy5-Alexa750 distance) in a single experiment. In this labeling scheme, when the ATPase acts at the SHL+2 site, FRET between Cy3 and Cy5 would report on entry-side movement and FRET between the Cy5 and Alexa 750 would report on exit-side movement. Similarly, when the ATPase acts at the SHL–2 site, the Cy3-Cy5 FRET would report on exit-side movement and the Cy5-Alexa 750 FRET would report on entry-side movement. The Cy3-Alexa750 FRET could provide a more sensitive measure of whether the remodeler perturbs the canonical wrapping of the nucleosome. Determining the coordination between entry-side and exit-side movement, however, could be difficult if apparent FRET changes reflecting enzyme binding cannot be distinguished from real FRET changes due to DNA translocation. If such obstacles can be overcome, these experiments would enable determining whether remodeling occurs via a concerted motion model, where DNA motion at the entry and exit sides occurs simultaneously; a loop-propagation model, where entry-side movement precedes exit-side movement; or a tension-based model, where exit-side movement precedes entry-side movement (Figure 1.4.1).

In addition to the ability to reposition nucleosomes, RSC and other SWI/SNF enzymes have also been reported to engage in histone and octamer ejection [80, 81, 82, 83, 84, 85, 86]. We do not observe these activities in our assay because they require free nucleosomes or free DNA as acceptors, the activity of histone chaperones, or a dinucleosome construct. However, given that we do not observe substantial disruption to the canonical wrapping of the DNA or the conformation of the histone octamer, how DNA translocation activity gets coupled to the disassembly of nucleosomes remains incompletely understood. A number of factors have been reported to influence nucleosome disassembly by SWI/SNF family remodelers, including transcription factors [80, 202], histone chaperones [84, 86], histone modifications [203, 204], DNA sequence [205], and neighboring nucleosomes [85]. Our smFRET assays could be used to test whether any of these factors alter the nucleosome remodeling mechanism

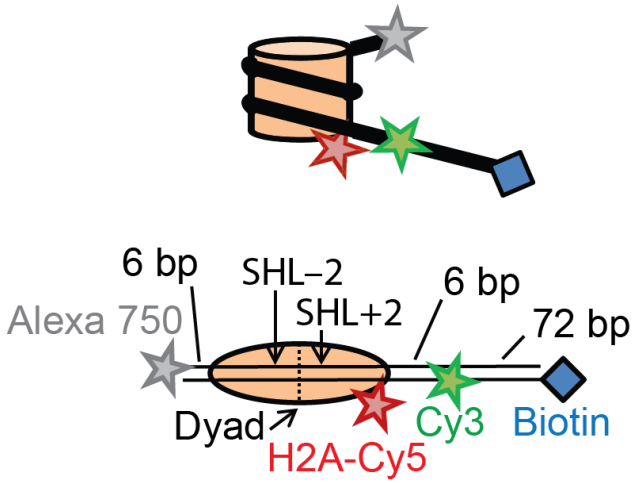


Figure 5.0.1: Construct design for the proposed three-color FRET experiments. Top: Cartoon of the nucleosome showing the labeling positions of the dyes. Bottom: Diagram of the footprint of the nucleosome (tan oval) on the double-stranded DNA (black lines). The Cy3, Cy5, and Alexa 750 dyes are shown as green, red, and grey stars, respectively. The biotin is shown as a blue square.

and induce SWI/SNF remodelers to unwrap DNA from the edges of the nucleosome or induce movement of the H2A-H2B dimer. Our observations that RSC does not substantially alter the canonical wrapping of the nucleosome or the structure of the octamer, however, is consistent with a model in which RSC does not disassemble the nucleosomes to which it binds, and instead ejects neighboring nucleosomes through a collision between the remodeler and the neighboring nucleosomes [85]. Regardless of the mechanism by which it occurs, reconstituting nucleosome disassembly in a smFRET assay would also allow us to resolve intermediates of the disassembly process. For example, studies of salt-induced octamer disassembly have found evidence of octamer conformational changes that precede H2A-H2B dimer loss [206], and experiments using the H2A/H2A construct (Figure 4.2.5e) could determine whether such intermediates occur during remodeler-induced octamer disassembly as well. Examining the lifetimes of these intermediate states under a number of different conditions could elucidate which steps of the disassembly process are affected by some of the factors, such as histone modifications, that are thought to regulate nucleosome disassembly.

Finally, *in vivo,* SWI/SNF remodelers are involved in transcriptional activation, the

creation of nucleosome free regions, and the formation of altered nucleosomal structures at specific loci [63, 64, 65, 66, 67]. These processes likely involve the cooperation of SWI/SNF enzymes with a number of other factors, including other chromatin remodeling enzymes, histone modifying enzymes, transcription factors, or histone chaperones. Reconstituting such activities using purified factors could be difficult given that the identities of the factors involved in these processes are incompletely understood. However, advancements have been made in performing single-molecule assays in whole cell lysates, for example, to study RNA splicing [207, 208] or examine biomolecular interactions and complex formation [209]. Using cell lysates from yeast may enable us to reconstitute some of these more complex remodeler activities and monitor them in a single molecule format. For example, RSC is thought to be involved in creating stable, partially unwrapped nucleosomes in the promoter region at the GAL1/10 locus in yeast [185]. We could create a fluorescently labeled nucleosomal substrate mimicking the features of this locus, immobilize it onto a microscope slide, apply whole cell lysate, and monitor the structure of these nucleosomes as the factors present in the lysate act. These studies could take advantage of the ease with which yeast can be genetically manipulated in order to introduce fluorescent labels on proteins of interest (such as the SWI/SNF and ISWI remodelers that have antagonistic functions at promoters [73]) or to knockout specific factors in order to assess their function. It is also possible to purify native chromatin from specific loci in yeast [210] in order to obtain a more physiological substrate for use in these experiments. Although labeling such native substrates would be difficult, it is possible to devise readouts based on single-molecule fluorescence to monitor processes such as transcription from the native loci [211]. Such experiments may extend the capability of single molecule fluorescence approaches to interrogate the function, mechanism, and regulation of molecular motor enzymes.

Appendix A

Unpublished data

A.1 Termination of reverse transcription

The (+) strand DNA of the HIV genome is synthesized in two parts, one that initiates at the 3' polypurine tract (3'PPT) sequence and the other that initiates at the central polypurine tract (cPPT) sequence (Figure 1.3.1). The DNA synthesized from the 3'PPT primer terminates at the central termination sequence (CTS), a sequence near the middle of the viral genome containing a dA_n-dT_n tract [48]. Mutation of the CTS is deleterious to viral replication, indicating the importance of termination to reverse transcription. Interestingly, the CTS stops DNA synthesis only in the context of strand displacement synthesis, and the CTS induces pausing, not termination, during DNA synthesis in the absence of a non-template strand. Although the narrow of the minor groove of the dA_n–dT_n tract is thought to be involved in causing pausing and termination at the CTS [212, 213], the mechanism for pausing and termination remain incompletely understood. Footprinting and kinetic data suggest that RT may interact with the CTS in a different conformation than typically observed when RT interacts with a double-stranded DNA [214, 215].

We hypothesized that the ability of RT to adopt multiple binding orientations or translational positions on its substrates may underlie the pausing and termination observed on the

CTS. To test this hypothesis, we examined the interaction of RT on 19 bp double-stranded DNA substrates containing either a wild-type CTS or a mutated CTS that does not cause pausing or termination [216] (Figure A.1.1a). The DNA substrates were labeled with the FRET acceptor dye Cy5 either near the 5' end of the primer (5' label) or near the 3' end of the primer (3' label) (Figure A.1.1a). We also employed two different labeling positions on RT, one in which the FRET acceptor dye Cy3 resides within the fingers domain at position 38 of the p66 subunit (F-label) and one where the Cy3 resides at the C-terminus of the protein near the RNase H domain (Figure A.1.1b) [24]. When RT interacted with a substrate containing the mutant sequence, the F-labeled RT gave a low FRET value on the 5'-labeled substrate and a high FRET value on the 3'-labeled substrate (Figure A.1.1c), indicating that the fingers domain and polymerase active site resided near the 3' end of the primer. The H-labeled RT gave a high FRET value on the 5'-labeled substrate and a low FRET value on the 3'-labeled substrate (Figure A.1.1c), indicating that the RNase H domain resided near the 5' end of the primer. These results are consistent with RT binding to the mutant substrate in the polymerase-competent binding mode, which is the expected binding mode for RT on a double-stranded DNA (dsDNA) substrate.

In contrast, when we examined RT's interactions with a substrate containing the wild-type CTS, we saw F-labeled RT binding to the substrate in such a way that both the 5'- and 3'-labeled substrates gave a high FRET value (Figure A.1.1d). Furthermore, H-labeled RT binding to both the 5'- and 3'-labeled substrates produced a low FRET value (Figure A.1.1d). These results seem to suggest that RT is binding to the wild-type CTS substrate in a different mode, which somehow places the fingers domain near the 5' and 3' ends of the substrate simultaneously. However, these two sites are situated ~ 6.5 nm away from each other and do not seem to move closer upon binding by RT (Figure A.1.1e). Thus, these results could indicate that the different labeling schemes may be causing RT to bind the 5'-labeled substrate in a different mode than it binds the 3'-labeled substrate.

We also examined RT's interactions with wild-type CTS and mutated CTS substrates

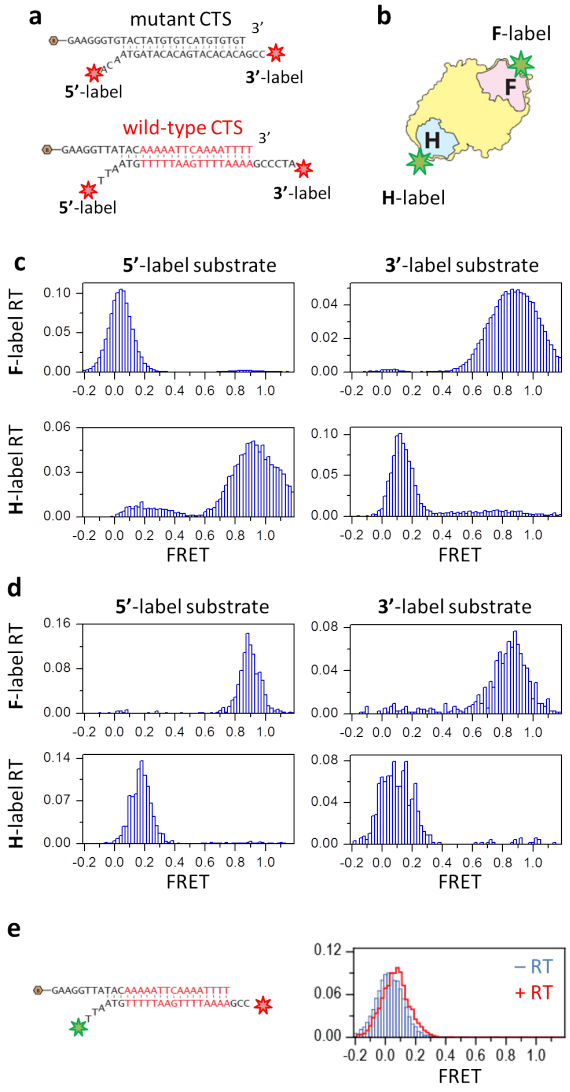


Figure A.1.1: Binding orientation of RT on substrates containing a wild-type or mutant CTS. **a,** Design of the mutant CTS (top) and wild-type CTS (bottom) constructs. The positions of Cy5 in the 5'- and 3'- labeled constructs are represented as red stars. The position of the biotin label is represented as a tan hexagon. The CTS is colored red. **b,** A cartoon of RT showing the position of the Cy3 label in the F-labeled and H-labeled constructs. The fingers domain (F) and RNase H domain (H) are colored pink and teal, respectively. **c,d,** Distribution of FRET values from the interaction of the mutant CTS substrate (c) or wild-type CTS substrate (d) with RT. Top Row: Experiments performed with F-labeled RT. Bottom Row: Experiments performed with H-labeled RT. Left Column: Experiments performed with the 5'-labeled substrate. Right Column: Experiments performed with the 3'-labeled substrate. **e,** Left: Diagram of the FRET-labeled CTS substrate to measure potential DNA bending. The positions of the Cy3 and Cy5 dyes are indicated by green and red stars, respectively. Right: Histogram showing the FRET values observed for the FRET-labeled CTS substrate in the absence (blue) or presence (red) of 1 µM unlabeled RT.

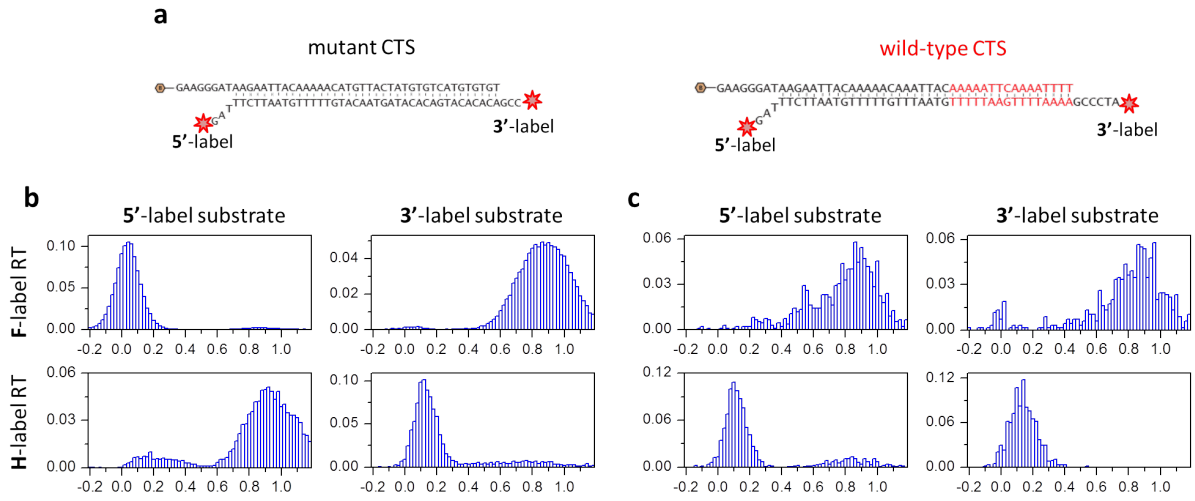


Figure A.1.2: Binding mode of RT on longer CTS substrates. **a,** Diagrams of the mutant (left) and wild-type (right) CTS substrates. The position of Cy5 in the 5'- and 3'- labeled constructs are represented as red stars. The position of the biotin label is represented as a tan hexagon. **b,c,** Distribution of FRET values from the interaction of the mutant CTS substrate (b) or wild-type CTS substrate (c) with RT. Top Row: Experiments performed with F-labeled RT. Bottom Row: Experiments performed with H-labeled RT. Left Column: Experiments performed with the 5'-labeled substrate. Right Column: Experiments performed with the 3'-labeled substrate.

containing a 38 bp duplex on which RT can display translational dynamics (Figure A.1.2a). On the mutated CTS, the F-labeled RT gives a high FRET value when interacting with 3'-labeled substrate and a zero FRET value when interacting with the 5'-labeled substrate (Figure A.1.2b). The H-labeled RT gives a low FRET value when interacting with both the 5'- and 3'-labeled substrates (Figure A.1.2b). This indicates that RT binds primarily in the polymerase mode at the front end of the duplex, as expected for RT's interaction with a normal dsDNA construct. RT does not spend a appreciable amount of time in the flipped orientation (where we would expect the H-labeled RT to give a high FRET value when interacting with the 3'-labeled substrate) or bound to the back end of the duplex (where we would expect the H-labeled RT to give a high FRET value when interacting with the 5' labeled substrate).

In contrast, RT interacted with the wild-type 38bp CTS substrate in a different binding mode. F-labeled RT showed a high FRET value on both the 5'-labeled and 3'-labeled

substrates while the H-labeled RT showed a low FRET value on both the 5'-labeled and 3'-labeled substrates (Figure A.1.2c). Again, these observations seem contradictory given the large (~ 13 nm) distance between the 3' and 5' ends of the DNA substrate, and may indicate that the labeling perturbs RT binding, causing RT to interact with the 5'- and 3'-labeled substrates in different ways. Thus, further experiments are required to determine RT's true binding mode on the CTS substrate.

A.2 Effect of DNA sequence on the remodeling direction of RSC

Because the nucleosome remodeling assays presented in Chapter 4 can determine the direction in which RSC repositions the octamer, we used our assay to examine the factors regulating the directionality of remodeling. RSC contains two subunits, Rsc3 and Rsc30, that bind to CGCG sequences [68]. In our initial experiments on RSC, we noticed that H2A/[end,+6] nucleosomes constructed from two different DNA sequences gave different yields of entry-side and exit-side movement traces (Figure A.2.1): on sequence A, we observed roughly 60% of traces showing entry-side motion and 40% of traces showing exit-side motion, but on sequence B, we observed 80% of traces showing entry-side motion and 20% of traces showing exit-side motion. This effect is largely an effect of the linker DNA as replacing the 78 bp linker DNA from sequence B with the linker DNA sequence from sequence A (to create the seq B+A_link construct) creates a construct where 60% of the remodeling traces show entry-side motion (Figure A.2.1). The linker DNA from sequence B contains a CGCG sequence near the edge of the 601 sequence, and to test whether this was involved in altering the remodeling direction of RSC, we mutated that region by replacing the CGCG sequence and the 5 bp preceding it with the corresponding sequence from sequence A (Figure A.2.1a). This construct showed a more equal number of entry- and exit-side movement traces (Figure A.2.1b), suggesting that the CGCG sequence is involved in biasing RSC binding in an

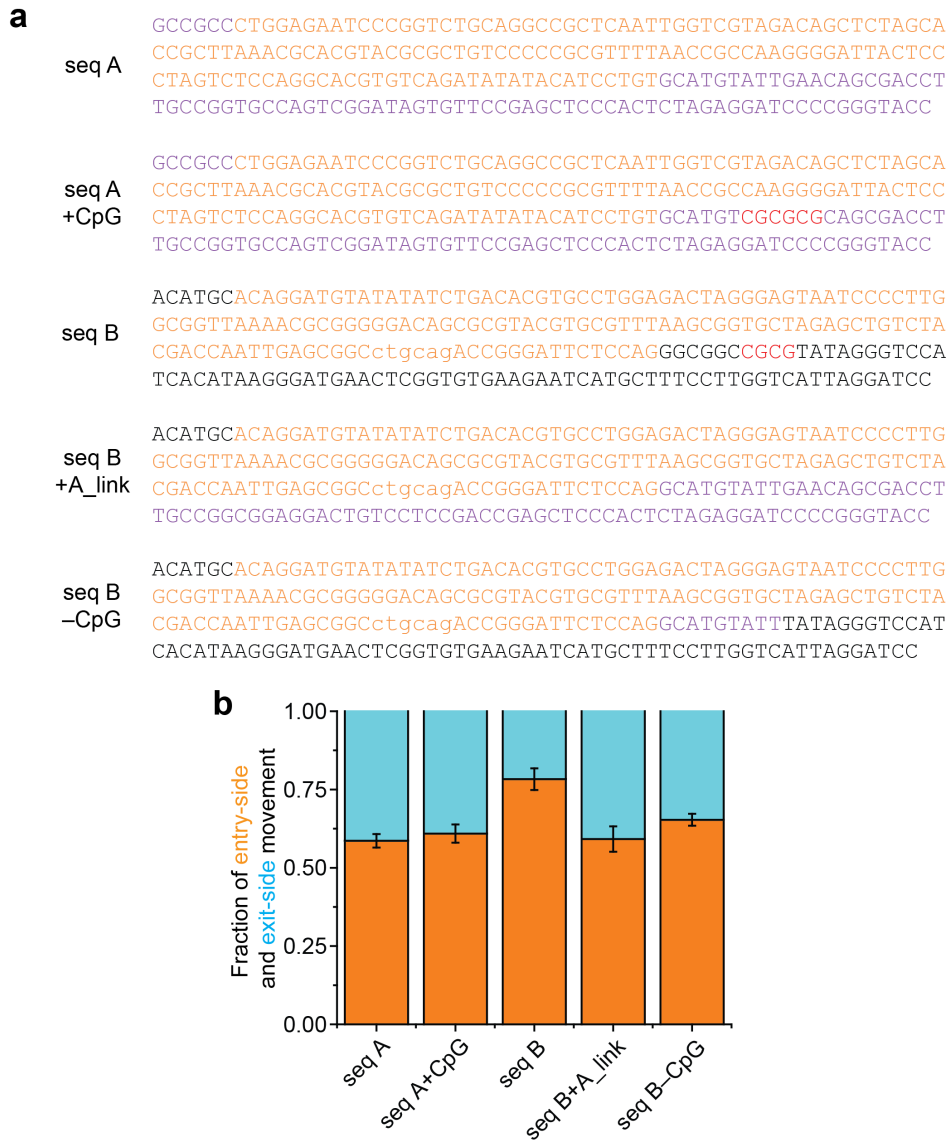


Figure A.2.1: Effects of DNA sequence on remodeling direction. **a,** The DNA sequences used in this study. The 601 positioning sequence is shown in orange. The linker DNA sequences from seq A are colored purple and the linker DNAs from seq B are colored black. CGCG sequences are highlighted in red. **b,** The fraction of remodeling traces showing entry-side motion (orange) and exit-side motion (cyan). Remodeling was measured in the presence of 1 nM RSC and 20 µM ATP at 30°C.

orientation that places the ATPase at the SHL–2 site. However, introducing a CGCGCG sequence into sequence A does not appreciably alter the preferred remodeling direction of the remodeler (Figure A.2.1), indicating that the CGCG may not be sufficient to alter the remodeling direction of RSC, and additional sequences may be required.

A.3 Measuring the path of DNA translocation during nucleosome remodeling by RSC

In this section, our goal was to determine whether the DNA moves along its canonical path during RSC remodeling or whether RSC disrupts the canonical wrapping of the nucleosome and moves the DNA along a different path. We previously addressed this question by using 2-nt ssDNA gaps to stall the remodeler after different amounts of translocation in order to measure the FRET of these intermediates and compare them to the FRET values of canonically wrapped nucleosomes with the dye repositioned by the same amount [26, 27]. To determine whether RSC translocates DNA along its canonical path as it enters the nucleosome, we constructed a series of H2A/[end,+12] nucleosomes with a 2-nt ssDNA gap placed at varying distances away from the SHL –2 site (Figure A.3.1a). The exit-side movement traces from these constructs showed a monotonic decrease in FRET to zero as expected, and the entry-side movement traces showed an increase in FRET to a higher value, where the remodeler stalls for some time before either the dyes photobleach or the FRET decreases again (Figure A.3.1b). For each of the gap locations we tested, we classified traces as showing entry-side or exit-side movement and plotted the fraction of entry- and exit-side movement traces observed (Figure A.3.1c). When the gap is at the SHL–2 site (n = 0), very few traces show entry-side movement, but as the gap moves away from the SHL–2 site, we saw an increase in the yield of entry-side movement traces (Figure A.3.1c). This result indicates that gaps closer than ~7 bp to the SHL–2 site can interfere with the onset ATPase action at the SHL–2 site.

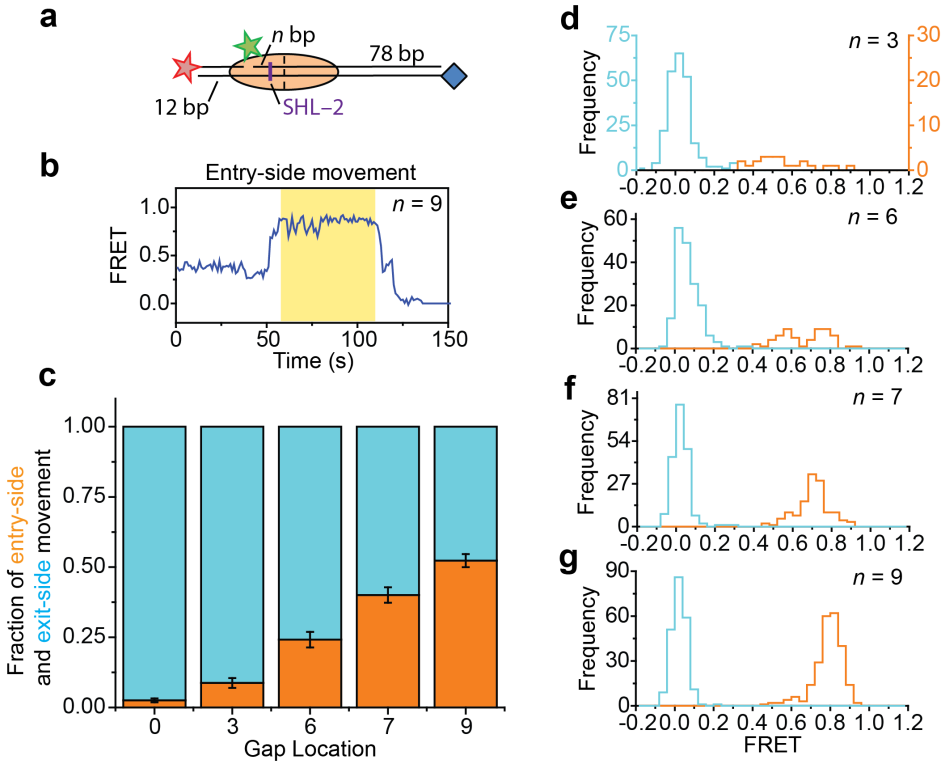


Figure A.3.1: Measuring the path of the DNA entering the nucleosome. **a,** Diagram of the H2A/[end,+12] constructs used. **b,** Representative entry-side movement trace from H2A/[end,+12] nucleosomes with a 2-nt ssDNA gap 9 bp away from the SHL–2 site. The region where the remodeler stalls at the gap, from which the final FRET values are calculated, is highlighted in yellow. **c,** The fraction of traces showing exit-side (cyan) and entry-side (orange) movement. **d-g,** Histograms showing the final FRET values from the exit-side (cyan) and entry-side (orange) movement traces on H2A/[end,+12] constructs where the gap is 3 (d), 6 (e), 7 (f), or 9 (g) bp away from the SHL–2 site.

We also looked at the final FRET values at reached by the remodeling reactions and plotted a histogram of the final FRET values for exit-side and entry-side movement (Figure A.3.1d-g). Although the exit-side movement traces show a homogeneous peak at a final FRET of zero, the entry-side movement traces from many of the constructs exhibited fairy heterogeneous final FRET values (Figure A.3.1d,e). Because many of the constructs exhibit heterogeneous stalling locations during entry-side movement, it is difficult to draw any conclusions from these data.

We performed a similar set of experiments to monitor the DNA exiting the nucleosome by constructing H2A/[end,+6] nucleosomes with 2-nt ssDNA gaps varying distances away from the SHL+2 site (Figure A.3.2a). The entry-side movement traces showed the expected transient FRET increase followed by a decrease to low FRET while the exit-side movement traces showed a monotonic decrease to low FRET where the remodeler pauses until photobleaching occurs or the FRET shows a subsequent increase (Figure A.3.2b). Quantifying the fraction of traces showing exit-side and entry-side movement showed a strong preference for generating entry-side movement traces that decreased as the gap was moved away from the SHL+2 site (Figure A.3.2c). This result confirms that gaps closer than ~7 bp to the SHL2 site can interfere with the onset of ATPase action at that site.

For each of constructs examined, we recorded the final FRET values from the exit-side and entry-side motion traces and plotted histograms of those values (Figure A.3.2d-g). The entry-side movement traces all show a consistent final FRET value ~ 0.2 FRET as expected. The entry-side movement traces, however, mostly showed a broad peak at low FRET whose position did not change appreciably as the position of the gap was moved on the DNA. In the case of the constructs where the gap is 3 or 5 bp away from the SHL+2 site, the final FRET values from the exit-side movement traces overlapped well with the distribution of final FRET values observed from the entry-side movement traces. Given that we observed very many more entry-side movement traces than exit-side movement traces, it is possible

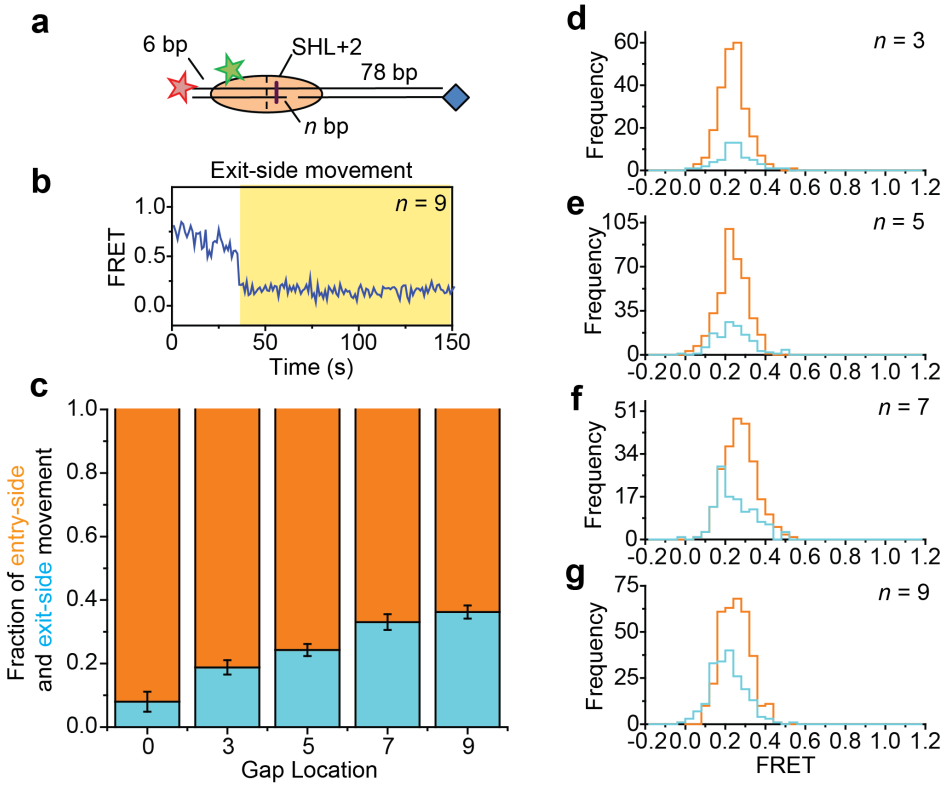


Figure A.3.2: Measuring the path of the DNA exiting the nucleosome. **a,** Diagram of the H2A/[end,+6] constructs used. **b,** Representative exit-side movement trace from H2A/[end,+6] nucleosomes with a 2-nt ssDNA gap 9 bp away from the SHL+2 site. The region where the remodeler stalls at the gap, from which the final FRET values are calculated, is highlighted in yellow. **c,** The fraction of traces showing exit-side (cyan) and entry-side (orange) movement. **d-g,** Histograms showing the final FRET values from the exit-side (cyan) and entry-side (orange) movement traces on H2A/[end,+6] constructs where the gap is 3 (d), 5 (e), 7 (f), or 9 (g) bp away from the SHL–2 site.

that some of the entry-side movement traces exhibited a transient FRET increase too short to be detected, and were misclassified as exit-side movement traces.

To avoid problems with trace misclassification, we adopted a different strategy by placing two consecutive methylphosphonate (MeP) linkages in the DNA in order to stall the remodeler. MeP groups are uncharged analogs of the phosphate groups that link nucleosides in DNA and have been used to interfere with the ability of DNA translocases to interact with the sugar-phosphate backbone of the DNA [217]. Whereas nucleosomes with 2-nt ssDNA gaps at both SHL2 sites are unstable (data not shown), nucleosomes with MeP linkages installed at both SHL2 sites are stable, perhaps because the MeP linkages perturb the DNA structure less than gaps. Thus, we created H2A/[end,+6] constructs with consecutive MeP linkages at the SHL–2 site, to block the generation of entry-side movement traces, and another set of consecutive MeP linkages either 5 or 9 bp from SHL+2, to limit the amount of DNA translocation at this site (Figure A.3.3a). We therefore expected these two constructs to show 5 or 9 bp of exit-side movement followed by stalling, that would leave the Cy5 dye 11 or 15 bp away from the edge of the nucleosome. As expected, remodeling of these constructs produced no entry-side movement traces, and the exit-side movement traces showed monotonic decreases in FRET to low FRET values. We determined the initial and final FRET values from many nucleosomes showing remodeling and plotted histograms of these values (Figure A.3.3b,c). The construct where the MeP linkages lie 9 bp away from the SHL+2 site produced uniform stalling at a FRET of 0.09 ± 0.05, which is consistent with ~11 bp of translocation along the canonical path (Figure A.3.3c). However, the final FRET value from the constructs where the MeP linkages lie 5 bp away from the SHL+2 site were very heterogeneous (Figure A.3.3b), making it difficult to interpret the data.

While the strategy of stalling the remodeler was successful for measuring the path of DNA translocation by the ISWI family remodelers, the SWI/SNF remodelers engage the SHL2 site much differently than the ISWI family remodelers [124], perhaps accounting for why remodeling of some of these constructs exhibited stalling at heterogeneous FRET val-

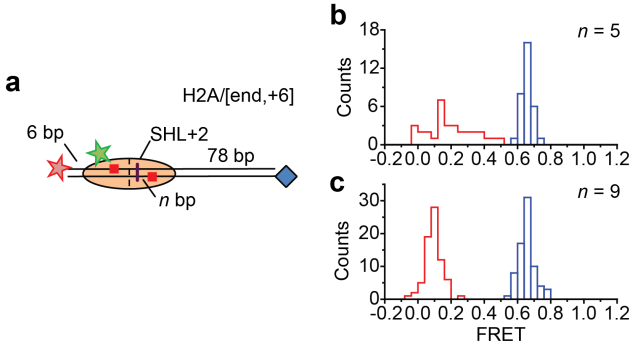


Figure A.3.3: Stalling DNA translocation with methylphosphonate linkages. **a,** Design of the H2A/[end,+6] constructs used. The location of the methylphosphonate linkages are shown as red squares. **b,c,** Histograms showing the average FRET before (blue) and after (red) remodeling with 1 nM RSC and 20 µM ATP.

ues (Figures A.3.1d,e and A.3.3b). Because we could not interpret the results of these experiments, we can only qualitatively assess whether the FRET changes we observe during nucleosome remodeling by RSC are consistent with the DNA moving along its canonical path.

A.4 RSC can reverse directions during remodeling

In our examination of RSC activity on constructs containing 2-nt ssDNA gaps, we noticed that many of the traces showed the enzyme apparently reversing direction after the ATPase stalls at the location of the gap (Figure A.4.1a,b). For example, when RSC remodeled H2A/[end,+12] nucleosomes with a gap 9 bp away from SHL–2, the traces showing entry-side movement result from ATPase action at the SHL–2 site and would therefore be blocked from translocating more than 9 bp of DNA. These traces showed an initial waiting period for remodeling to begin (region 1), a increase in FRET followed by stalling at a high FRET value (region 2), and a monotonic decrease to zero FRET (region 3) (Figure A.4.1a). This monotonic decrease to zero FRET in region 3 resembles exit-side DNA motion, indicating that the remodeler may have reversed the direction of DNA translocation after stalling. However, this monotonic decrease in FRET could result from other processes, such as unwrapping of the

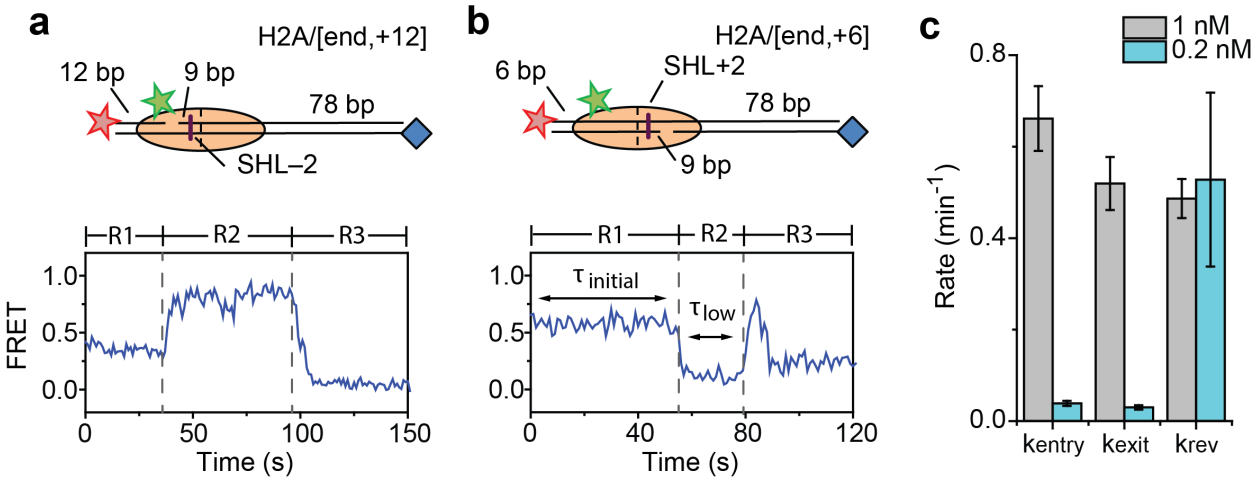


Figure A.4.1: Direction reversal by RSC. **a,b,** Entry-side movement of H2A/[end,+12] nucleosomes with a 2-nt ssDNA gap 9 bp away from SHL–2 (a) or exit-side movement of H2A/[end,+6] nucleosomes with a 2-nt ssDNA gap 9 bp away from SHL+2 (b). Top: Cartoon of the nucleosome constructs. Bottom: Example FRET traces. Region 1 (R1) highlights the waiting period before remodeling begins, region 2 (R2) highlights entry-side movement (a) or exit-side movement (b) followed by stalling of the ATPase, and region 3 (R3) highlights direction reversal. **c,** Rate constants from exit-side movement traces from the remodeling of the nucleosome constructs in panel (b) in the presence of 50 µM ATP and either 1 nM or 0.2 nM RSC.

DNA away from the edge of the nucleosome. However, when RSC remodeled H2A/[end,+6] nucleosomes with a gap 9 bp away from SHL+2, the exit-side movement traces exhibited an initial waiting period (region 1), a monotonic decrease in FRET to stall at a low FRET value (region 2), and a subsequent transient increase in FRET followed by a decrease to ~ 0.2 FRET (region 3) (Figure A.4.1b). The FRET changes observed in region 3 of these traces are characteristic of entry-side movement traces and inconsistent with processes like unwrapping, indicating that after stalling at the 2-nt ssDNA gap, the remodeler can reverse direction and translocate DNA in the opposite direction around the nucleosome.

These direction reversals could be due either to binding of a second remodeler after dissociation of the original remodeler, or it could be due to a single remodeler reversing direction without dissociating. In order to distinguish between these possibilities, we tested the dependence of the reversal kinetics on the concentration of enzyme during remodeling of the H2A/[end,+6] nucleosomes with a gap 9 bp away from the SHL+2 site. If the direction

reversal is due to enzyme dissociation followed by enzyme binding, the rate at which direction reversal occurs (k_{rev}) should be dependent on the concentration of enzyme. However, if the reversal represents a single enzyme dynamically reorienting itself, k_{rev} should be independent of the enzyme concentration. To determine k_{rev}, we performed remodeling reactions with either 1 nM or 0.2 nM RSC, examined all traces showing exit-side movement, measured the lifetime of the low FRET state (τ_{low}), and recorded the fraction of traces that showed direction reversal versus the fraction of traces that showed bleaching before reversal could occur. Because photobleaching and direction reversal compete in determining the lifetime of the low FRET state, k_{rev} is given by $P_{rev}/\langle\tau_{low}\rangle$, where $\langle\tau_{low}\rangle$ is the average lifetime of the low FRET state and P_{rev} is the fraction of traces showing direction reversal. These experiments demonstrate that k_{rev} is similar between the two enzyme concentrations (Figure A.4.1c). We also estimated the rate at which remodeling begins by similar means: we examined all traces, measured the lifetime of the initial FRET state until remodeling or bleaching occurred ($\tau_{initial}$), and recorded the fraction of traces where we observed exit-side motion, entry-side motion, or bleaching before either type of motion could be observed. Because enzyme binding is required for remodeling, we the rate at which remodeling begins should depend on the concentration of enzyme. Indeed, the rate at which exit-side movement begins (k_{exit}) and the rate at which entry-side movement begins (k_{entry}) both show a strong dependence on the concentration of enzyme (Figure A.4.1c). Furthermore, at 0.2 nM RSC, the k_{rev} is much faster than k_{exit} or k_{entry}, indicating that the reversal step cannot involve free enzyme binding from solution.

Thus, when the ATPase of RSC stalls at a 2-nt ssDNA gap, it is capable of reversing direction and this direction reversal does not require enzyme dissociation. This direction reversal could either be due to the ATPase disengaging from one SHL2 site and the entire complex reorienting itself to engage the other SHL2 site, or it could involve the ATPase translocating DNA in the opposite direction, which has been observed previously [109, 218].

However, given that the behavior of RSC when remodeling most constructs containing gaps is strange (see Section A.3), the physiological relevance of this observation is unclear.

A.5 The kinetic step size of RSC

Although ISWI remodelers translocate DNA 1 bp at a time, they exhibit kinetic pausing and show an initial step of 7 bp followed by subsequent steps of 3-4 bp in size [26, 27]. To determine whether RSC exhibits a similar kinetic step size, we examined the exit-side movement traces from a series of H2A/[end,+n] constructs , identified the location of the major pause sites, and constructed a histogram of these values across many individual nucleosome remodeling events (Figure A.5.1). These histograms showed a peak at an intermediate FRET value, and the average FRET value of this peak increased as the initial length of the DNA linker was changed from 3 bp to 0 bp to -3 bp (Figure A.5.1c). Assuming the DNA exits the nucleosome along its canonical path, we can estimate the kinetic step size to be 11.6 $\pm$ 1.2 bp, 12.8 $\pm$ 1.5 bp and 13.3 $\pm$ 1.5 bp on the n = 3, 0, and –3 bp constructs, respectively. Thus, overall these results suggest a kinetic step size of 12.6 $\pm$ 0.8 bp. These measurements suggest that nucleosome remodeling by RSC involves kinetic step sizes larger than those observed for ISWI remodelers.

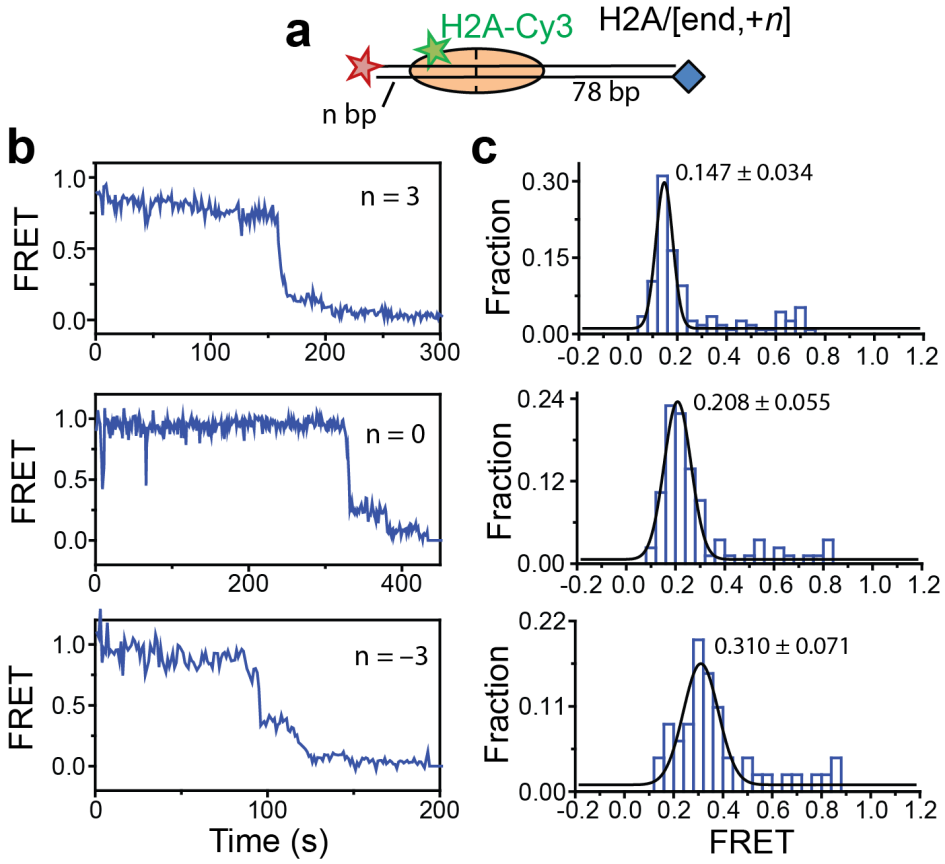


Figure A.5.1: Kinetic step size of RSC. **a,** Cartoon depicting the nucleosome construct used. **b,** Representative FRET traces showing exit-side movement on nucleosomes with n = 3, 0 or -3 bp of linker DNA. **c,** The FRET values at which the remodeling traces show pausing are plotted on a histogram and fit to a Gaussian curve (black). Histograms are constructed from >50 traces collected in at least three independent experiments. The reported error is the standard deviation of the Gaussian curve. [RSC] = 1-5 nM, [ATP] = 2-5 µM, 30°C.

Bibliography

[1] Boehr, D. D., Nussinov, R., and Wright, P. E. (2009) The role of dynamic conformational ensembles in biomolecular recognition. *Nature chemical biology 5*, 789–96.

[2] Cui, Q., and Karplus, M. (2008) Allostery and cooperativity revisited. *Protein science : a publication of the Protein Society 17*, 1295–307.

[3] Tzeng, S.-R., and Kalodimos, C. G. (2011) Protein dynamics and allostery: an NMR view. *Current opinion in structural biology 21*, 62–7.

[4] Villali, J., and Kern, D. (2010) Choreographing an enzyme's dance. *Current opinion in chemical biology 14*, 636–43.

[5] Hammes, G. G., Benkovic, S. J., and Hammes-Schiffer, S. (2011) Flexibility, diversity, and cooperativity: pillars of enzyme catalysis. *Biochemistry 50*, 10422–30.

[6] Tokuriki, N., and Tawfik, D. S. (2009) Protein dynamism and evolvability. *Science (New York, N.Y.) 324*, 203–7.

[7] Leschziner, A. E., and Nogales, E. (2007) Visualizing flexibility at molecular resolution: analysis of heterogeneity in single-particle electron microscopy reconstructions. *Annual review of biophysics and biomolecular structure 36*, 43–62.

[8] Dror, R. O., Dirks, R. M., Grossman, J. P., Xu, H., and Shaw, D. E. (2012) Biomolecular simulation: a computational microscope for molecular biology. *Annual review of biophysics 41*, 429–52.

[9] Neutze, R., and Moffat, K. (2012) Time-resolved structural studies at synchrotrons and X-ray free electron lasers: opportunities and challenges. *Current opinion in structural biology 22*, 651–9.

[10] Rosenzweig, R., and Kay, L. E. (2014) Bringing dynamic molecular machines into focus by methyl-TROSY NMR. *Annual review of biochemistry 83*, 291–315.

[11] Greenleaf, W. J., Woodside, M. T., and Block, S. M. (2007) High-Resolution, Single-Molecule Measurements of Biomolecular Motion. *Annual Review of Biophysics and Biomolecular Structure 36*, 171–190.

[12] Toprak, E., and Selvin, P. R. (2007) New fluorescent tools for watching nanometer-scale conformational changes of single molecules. *Annual review of biophysics and biomolecular structure 36*, 349–69.

[13] Myong, S., and Ha, T. (2010) Stepwise translocation of nucleic acid motors. *Current opinion in structural biology 20*, 121–7.

[14] Bustamante, C., Cheng, W., Mejia, Y. X., and Meija, Y. X. (2011) Revisiting the central dogma one molecule at a time. *Cell 144*, 480–97.

[15] Ha, T., Enderle, T., Ogletree, D. F., Chemla, D. S., Selvin, P. R., and Weiss, S. (1996) Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc Natl Acad Sci USA 93*, 6264–6268.

[16] Stryer, L., and Haugland, R. P. (1967) Energy Transfer - a Spectroscopic Ruler. *Proc Natl Acad Sci USA 58*, 719–&.

[17] Xie, X. S., and Lu, H. P. (1999) Single-molecule Enzymology. *Journal of Biological Chemistry 274*, 15967–15970.

[18] Moffitt, J. R., Chemla, Y. R., and Bustamante, C. (2010) Methods in statistical kinetics. *Methods in enzymology 475*, 221–57.

[19] Roy, R., Hohng, S., and Ha, T. (2008) A practical guide to single-molecule FRET. *Nature methods 5*, 507–16.

[20] Joo, C., and Ha, T. In *Single-Molecule Techniques: A Laboratory Manual*; Selvin, P. R., and Ha, T., Eds.; Cold Spring Harboar Laboratory Press: Cold Spring Harbor, N.Y., 2008; Chapter 2, pp 3–36.

[21] Lakowicz, J. R. *Principles of fluorescence spectroscopy*, 3rd ed.; Springer: New York, 2006.

[22] Murphy, M. C., Rasnik, I., Cheng, W., Lohman, T. M., and Ha, T. (2004) Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. *Biophys J 86*, 2530–2537.

[23] Deniz, A. A., Laurence, T. A., Dahan, M., Chemla, D. S., Schultz, P. G., and Weiss, S. (2003) Ratiometric single-molecule studies of freely diffusing biomolecules.

[24] Abbondanzieri, E. A., Bokinsky, G., Rausch, J. W., Zhang, J. X., Le Grice, S. F., and Zhuang, X. (2008) Dynamic binding orientations direct activity of HIV reverse transcriptase. *Nature 453*, 184–189.

[25] Liu, S., Abbondanzieri, E. A., Rausch, J. W., Le Grice, S. F., and Zhuang, X. (2008) Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. *Science 322*, 1092–1097.

[26] Blosser, T. R., Yang, J. G., Stone, M. D., Narlikar, G. J., and Zhuang, X. (2009) Dynamics of nucleosome remodelling by individual ACF complexes. *Nature 462*, 1022–1027.

[27] Deindl, S., Hwang, W. L., Hota, S. K., Blosser, T. R., Prasad, P., Bartholomew, B., and Zhuang, X. (2013) ISWI Remodelers Slide Nucleosomes with Coordinated Multi-Base-Pair Entry Steps and Single-Base-Pair Exit Steps. *Cell 152*, 442–452.

[28] Baltimore, D. (1970) RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature 226*, 1209–11.

[29] Temin, H. M., and Mizutani, S. (1970) RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature 226*, 1211–3.

[30] Telesnitsky, A., and Goff, S. P. In *Retroviruses*; Coffin, J. M., Hughes, S. H., and Varmus, H. E., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1997; Chapter 4, pp 121–160.

[31] Sarafianos, S. G., Marchand, B., Das, K., Himmel, D. M., Parniak, M. A., Hughes, S. H., and Arnold, E. (2009) Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *Journal of molecular biology 385*, 693–713.

[32] Le Grice, S. F. J. (2012) Human immunodeficiency virus reverse transcriptase: 25 years of research, drug discovery, and promise. *The Journal of biological chemistry 287*, 40850–7.

[33] Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1992) Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science (New York, N.Y.) 256*, 1783–90.

[34] Hansen, J., Schulze, T., Mellert, W., and Moelling, K. (1988) Identification and characterization of HIV-specific RNase H by monoclonal antibody. *The EMBO journal 7*, 239–43.

[35] Metzger, W., Hermann, T., Schatz, O., Le Grice, S. F., and Heumann, H. (1993) Hydroxyl radical footprint analysis of human immunodeficiency virus reverse transcriptase-template.primer complexes. *Proceedings of the National Academy of Sciences of the United States of America 90*, 5909–13.

[36] Ding, J., Das, K., Hsiou, Y., Sarafianos, S. G., Clark, A. D., Jacobo-Molina, A., Tantillo, C., Hughes, S. H., and Arnold, E. (1998) Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA template-primer and an antibody Fab fragment at 2.8 A resolution. *Journal of molecular biology 284*, 1095–111.

[37] Sarafianos, S. G., Das, K., Tantillo, C., Clark Jr., A. D., Ding, J., Whitcomb, J. M., Boyer, P. L., Hughes, S. H., and Arnold, E. (2001) Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J 20*, 1449–1461.

[38] Harada, F., Sawyer, R. C., and Dahlberg, J. E. (1975) A primer ribonucleic acid for initiation of in vitro Rous sarcarcoma virus deoxyribonucleic acid synthesis. *The Journal of biological chemistry 250*, 3487–97.

[39] Mölling, K., Bolognesi, D. P., Bauer, H., Büsen, W., Plassmann, H. W., and Hausen, P. (1971) Association of viral reverse transcriptase with an enzyme degrading the RNA moiety of RNA-DNA hybrids. *Nature: New biology 234*, 240–3.

[40] Baltimore, D., and Smoler, D. F. (1972) Association of an endoribonuclease with the avian myeloblastosis virus deoxyribonucleic acid polymerase. *The Journal of biological chemistry 247*, 7282–7.

[41] Keller, W., and Crouch, R. (1972) Degradation of DNA RNA hybrids by ribonuclease H and DNA polymerases of cellular and viral origin. *Proceedings of the National Academy of Sciences of the United States of America 69*, 3360–4.

[42] Leis, J. P., Berkower, I., and Hurwitz, J. (1973) Mechanism of action of ribonuclease H isolated from avian myeloblastosis virus and Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America 70*, 466–70.

[43] Smith, J. K., Cywinski, A., and Taylor, J. M. (1984) Initiation of plus-strand DNA synthesis during reverse transcription of an avian retrovirus genome. *Journal of virology 49*, 200–4.

[44] Omer, C. A., Resnick, R., and Faras, A. J. (1984) Evidence for involvement of an RNA primer in initiation of strong-stop plus DNA synthesis during reverse transcription in vitro. *Journal of virology 50*, 465–70.

[45] Finston, W. I., and Champoux, J. J. (1984) RNA-primed initiation of Moloney murine leukemia virus plus strands by reverse transcriptase in vitro. *Journal of virology 51*, 26–33.

[46] Huber, H. E., and Richardson, C. C. (1990) Processing of the primer for plus strand DNA synthesis by human immunodeficiency virus 1 reverse transcriptase. *The Journal of biological chemistry 265*, 10565–73.

[47] Hu, W. S., and Temin, H. M. (1990) Retroviral recombination and reverse transcription. *Science (New York, N.Y.) 250*, 1227–33.

[48] Charneau, P., Mirambeau, G., Roux, P., Paulous, S., Buc, H., and Clavel, F. (1994) HIV-1 reverse transcription. A termination step at the center of the genome. *Journal of molecular biology 241*, 651–62.

[49] Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science 282*, 1669–1675.

[50] Sarafianos, S. G., Clark, A. D., Das, K., Tuske, S., Birktoft, J. J., Ilankumaran, P., Ramesha, A. R., Sayer, J. M., Jerina, D. M., Boyer, P. L., Hughes, S. H., and Arnold, E. (2002) Structures of HIV-1 reverse transcriptase with pre- and post-translocation AZTMP-terminated DNA. *The EMBO journal 21*, 6614–24.

[51] Rausch, J. W., and Le Grice, S. F. (2004) 'Binding, bending and bonding': polypurine tract-primed initiation of plus-strand DNA synthesis in human immunodeficiency virus. *The International Journal of Biochemistry & Cell Biology 36*, 1752–1766.

[52] Schultz, S. J., Zhang, M., Kelleher, C. D., and Champoux, J. J. (2000) Analysis of plus-strand primer selection, removal, and reutilization by retroviral reverse transcriptases. *The Journal of biological chemistry 275*, 32299–309.

[53] Gotte, M., Li, X., and Wainberg, M. A. (1999) HIV-1 reverse transcription: a brief overview focused on structure-function relationships among molecules involved in initiation of the reaction. *Arch Biochem Biophys 365*, 199–210.

[54] Le Grice, S. F. (2003) "In the beginning": initiation of minus strand DNA synthesis in retroviruses and LTR-containing retrotransposons. *Biochemistry 42*, 14349–14355.

[55] Tisne, C. (2005) Structural bases of the annealing of primer tRNA(3Lys) to the HIV-1 viral RNA. *Curr HIV Res 3*, 147–156.

[56] Abbink, T. E., and Berkhout, B. (2008) HIV-1 reverse transcription initiation: a potential target for novel antivirals? *Virus Res 134*, 4–18.

[57] Isel, C., Ehresmann, C., and Marquet, R. (2010) Initiation of HIV reverse transcription. *Viruses 2*, 213–243.

[58] Kornberg, R. D. (1974) Chromatin structure: a repeating unit of histones and DNA. *Science (New York, N.Y.) 184*, 868–71.

[59] Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature 389*, 251–260.

[60] Richmond, T. J., and Davey, C. A. (2003) The structure of DNA in the nucleosome core. *Nature 423*, 145–150.

[61] Luger, K., Dechassa, M. L., and Tremethick, D. J. (2012) New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat Rev Mol Cell Biol 13*, 436–447.

[62] Swygert, S. G., and Peterson, C. L. (2014) Chromatin dynamics: interplay between remodeling enzymes and histone modifications. *Biochim Biophys Acta 1839*, 728–736.

[63] Clapier, C. R., and Cairns, B. R. (2009) The biology of chromatin remodeling complexes. *Annu Rev Biochem 78*, 273–304.

[64] Hargreaves, D. C., and Crabtree, G. R. (2011) ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell Res 21*, 396–420.

[65] Narlikar, G. J., Sundaramoorthy, R., and Owen-Hughes, T. (2013) Mechanisms and Functions of ATP-Dependent Chromatin-Remodeling Enzymes. *Cell 154*, 490–503.

[66] Becker, P. B., and Workman, J. L. (2013) Nucleosome remodeling and epigenetics. *Cold Spring Harb Perspect Biol 5*.

[67] Bartholomew, B. (2014) Regulating the chromatin landscape: structural and mechanistic perspectives. *Annual review of biochemistry 83*, 671–96.

[68] Badis, G. et al. (2008) A library of yeast transcription factor motifs reveals a widespread function for Rsc3 in targeting nucleosome exclusion at promoters. *Mol Cell 32*, 878–887.

[69] Hartley, P. D., and Madhani, H. D. (2009) Mechanisms that specify promoter nucleosome location and identity. *Cell 137*, 445–458.

[70] Wippo, C. J., Israel, L., Watanabe, S., Hochheimer, A., Peterson, C. L., and Korber, P. (2011) The RSC chromatin remodelling enzyme has a unique role in directing the accurate positioning of nucleosomes. *EMBO J 30*, 1277–1288.

[71] Whitehouse, I., Rando, O. J., Delrow, J., and Tsukiyama, T. (2007) Chromatin remodelling at promoters suppresses antisense transcription. *Nature 450*, 1031–1035.

[72] Yadon, A. N., Van de Mark, D., Basom, R., Delrow, J., Whitehouse, I., and Tsukiyama, T. (2010) Chromatin remodeling around nucleosome-free regions leads to repression of noncoding RNA transcription. *Mol Cell Biol 30*, 5110–5122.

[73] Parnell, T. J., Schlichter, A., Wilson, B. G., and Cairns, B. R. (2015) The chromatin remodelers RSC and ISW1 display functional and chromatin-based promoter antagonism. *eLife 4*, e06073.

[74] Ho, L., and Crabtree, G. R. (2010) Chromatin remodelling during development. *Nature 463*, 474–484.

[75] Wilson, B. G., and Roberts, C. W. (2011) SWI/SNF nucleosome remodellers and cancer. *Nat Rev Cancer 11*, 481–492.

[76] Ronan, J. L., Wu, W., and Crabtree, G. R. (2013) From neural development to cognition: unexpected roles for chromatin. *Nat Rev Genet 14*, 347–359.

[77] Hamiche, A., Sandaltzopoulos, R., Gdula, D. A., and Wu, C. (1999) ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF. *Cell 97*, 833–42.

[78] Langst, G., Bonte, E. J., Corona, D. F. V., and Becker, P. B. (1999) Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. *Cell 97*, 843–852.

[79] Whitehouse, I., Flaus, A., Cairns, B. R., White, M. F., Workman, J. L., and Owen-Hughes, T. (1999) Nucleosome mobilization catalysed by the yeast SWI/SNF complex. *Nature 400*, 784–787.

[80] Owen-Hughes, T., Utley, R. T., Cote, J., Peterson, C. L., and Workman, J. L. (1996) Persistent site-specific remodeling of a nucleosome array by transient action of the SWI/SNF complex. *Science 273*, 513–516.

[81] Lorch, Y., Zhang, M., and Kornberg, R. D. (1999) Histone octamer transfer by a chromatin-remodeling complex. *Cell 96*, 389–392.

[82] Bruno, M., Flaus, A., Stockdale, C., Rencurel, C., Ferreira, H., and Owen-Hughes, T. (2003) Histone H2A/H2B dimer exchange by ATP-dependent chromatin remodeling activities. *Mol Cell 12*, 1599–1606.

[83] Boeger, H., Griesenbeck, J., Strattan, J. S., and Kornberg, R. D. (2003) Nucleosomes unfold completely at a transcriptionally active promoter. *Mol Cell 11*, 1587–1598.

[84] Lorch, Y., Maier-Davis, B., and Kornberg, R. D. (2006) Chromatin remodeling by nucleosome disassembly in vitro. *Proc Natl Acad Sci U S A 103*, 3090–3093.

[85] Dechassa, M. L., Sabri, A., Pondugula, S., Kassabov, S. R., Chatterjee, N., Kladde, M. P., and Bartholomew, B. (2010) SWI/SNF has intrinsic nucleosome disassembly activity that is dependent on adjacent nucleosomes. *Mol Cell 38*, 590–602.

[86] Kuryan, B. G., Kim, J., Tran, N. N., Lombardo, S. R., Venkatesh, S., Workman, J. L., and Carey, M. (2012) Histone density is maintained during transcription mediated by the chromatin remodeler RSC and histone chaperone NAP1 in vitro. *Proc Natl Acad Sci U S A 109*, 1931–1936.

[87] Ito, T., Bulger, M., Pazin, M. J., Kobayashi, R., and Kadonaga, J. T. (1997) ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell 90*, 145–155.

[88] Ito, T., Levenstein, M. E., Fyodorov, D. V., Kutach, A. K., Kobayashi, R., and Kadonaga, J. T. (1999) ACF consists of two subunits, Acf1 and ISWI, that function cooperatively in the ATP-dependent catalysis of chromatin assembly. *Genes Dev 13*, 1529–1539.

[89] Loyola, A., LeRoy, G., Wang, Y. H., and Reinberg, D. (2001) Reconstitution of recombinant chromatin establishes a requirement for histone-tail modifications during chromatin assembly and transcription. *Genes & development 15*, 2837–51.

[90] Kukimoto, I., Elderkin, S., Grimaldi, M., Oelgeschläger, T., and Varga-Weisz, P. D. (2004) The histone-fold protein complex CHRAC-15/17 enhances nucleosome sliding and assembly mediated by ACF. *Molecular cell 13*, 265–77.

[91] Saha, A., Wittmeyer, J., and Cairns, B. R. (2002) Chromatin remodeling by RSC involves ATP-dependent DNA translocation. *Genes Dev 16*, 2120–2134.

[92] Whitehouse, I., Stockdale, C., Flaus, A., Szczelkun, M. D., and Owen-Hughes, T. (2003) Evidence for DNA translocation by the ISWI chromatin-remodeling enzyme. *Molecular and cellular biology 23*, 1935–45.

[93] Kagalwala, M. N., Glaus, B. J., Dang, W., Zofall, M., and Bartholomew, B. (2004) Topography of the ISW2-nucleosome complex: insights into nucleosome spacing and chromatin remodeling. *EMBO J 23*, 2092–2104.

[94] Schwanbeck, R., Xiao, H., and Wu, C. (2004) Spatial contacts and nucleosome step movements induced by the NURF chromatin remodeling complex. *The Journal of biological chemistry 279*, 39933–41.

[95] Saha, A., Wittmeyer, J., and Cairns, B. R. (2005) Chromatin remodeling through directional DNA translocation from an internal nucleosomal site. *Nat Struct Mol Biol 12*, 747–755.

[96] Zofall, M., Persinger, J., Kassabov, S. R., and Bartholomew, B. (2006) Chromatin remodeling by ISW2 and SWI/SNF requires DNA translocation inside the nucleosome. *Nat Struct Mol Biol 13*, 339–346.

[97] Dang, W., and Bartholomew, B. (2007) Domain architecture of the catalytic subunit in the ISW2-nucleosome complex. *Molecular and cellular biology 27*, 8306–17.

[98] Dechassa, M. L., Zhang, B., Horowitz-Scherer, R., Persinger, J., Woodcock, C. L., Peterson, C. L., and Bartholomew, B. (2008) Architecture of the SWI/SNF-nucleosome complex. *Mol Cell Biol 28*, 6010–6021.

[99] Bowman, G. D. (2010) Mechanisms of ATP-dependent nucleosome sliding. *Curr Opin Struct Biol 20*, 73–81.

[100] Längst, G., and Becker, P. B. (2001) ISWI induces nucleosome sliding on nicked DNA. *Molecular cell 8*, 1085–92.

[101] Aoyagi, S., and Hayes, J. J. (2002) hSWI/SNF-catalyzed nucleosome sliding does not occur solely via a twist-diffusion mechanism. *Mol Cell Biol 22*, 7484–7490.

[102] Lorch, Y., Davis, B., and Kornberg, R. D. (2005) Chromatin remodeling by DNA bending, not twisting. *Proc Natl Acad Sci U S A 102*, 1329–1332.

[103] Strohner, R., Wachsmuth, M., Dachauer, K., Mazurkiewicz, J., Hochstatter, J., Rippe, K., and Langst, G. (2005) A 'loop recapture' mechanism for ACF-dependent nucleosome remodeling. *Nat Struct Mol Biol 12*, 683–690.

[104] Narlikar, G. J., Phelan, M. L., and Kingston, R. E. (2001) Generation and interconversion of multiple distinct nucleosomal states as a mechanism for catalyzing chromatin fluidity. *Mol Cell 8*, 1219–1230.

[105] Aoyagi, S., Narlikar, G., Zheng, C., Sif, S., Kingston, R. E., and Hayes, J. J. (2002) Nucleosome remodeling by the human SWI/SNF complex requires transient global disruption of histone-DNA interactions. *Mol Cell Biol 22*, 3653–3662.

[106] Lorch, Y., Maier-Davis, B., and Kornberg, R. D. (2010) Mechanism of chromatin remodeling. *Proc Natl Acad Sci U S A 107*, 3458–3462.

[107] Shukla, M. S., Syed, S. H., Montel, F., Faivre-Moskalenko, C., Bednar, J., Travers, A., Angelov, D., and Dimitrov, S. (2010) Remosomes: RSC generated non-mobilized particles with approximately 180 bp DNA loosely associated with the histone octamer. *Proc Natl Acad Sci U S A 107*, 1936–1941.

[108] Kassabov, S. R., Zhang, B., Persinger, J., and Bartholomew, B. (2003) SWI/SNF unwraps, slides, and rewraps the nucleosome. *Mol Cell 11*, 391–403.

[109] Zhang, Y., Smith, C. L., Saha, A., Grill, S. W., Mihardja, S., Smith, S. B., Cairns, B. R., Peterson, C. L., and Bustamante, C. (2006) DNA translocation and loop formation mechanism of chromatin remodeling by SWI/SNF and RSC. *Mol Cell 24*, 559–568.

[110] Chaban, Y., Ezeokonkwo, C., Chung, W.-H., Zhang, F., Kornberg, R. D., Maier-Davis, B., Lorch, Y., and Asturias, F. J. (2008) Structure of a RSC-nucleosome complex and insights into chromatin remodeling. *Nature structural & molecular biology 15*, 1272–1277.

[111] Ramachandran, S., Zentner, G. E., and Henikoff, S. (2015) Asymmetric nucleosomes flank promoters in the budding yeast genome. *Genome Res 25*, 381–390.

[112] Kassabov, S. R., Henry, N. M., Zofall, M., Tsukiyama, T., and Bartholomew, B. (2002) High-resolution mapping of changes in histone-DNA contacts of nucleosomes remodeled by ISW2. *Molecular and cellular biology 22*, 7524–34.

[113] Fan, H. Y., He, X., Kingston, R. E., and Narlikar, G. J. (2003) Distinct strategies to make nucleosomal DNA accessible. *Mol Cell 11*, 1311–1322.

[114] Yang, J. G., Madrid, T. S., Sevastopoulos, E., and Narlikar, G. J. (2006) The chromatin-remodeling enzyme ACF is an ATP-dependent DNA length sensor that regulates nucleosome spacing. *Nat Struct Mol Biol 13*, 1078–1083.

[115] Rowe, C. E., and Narlikar, G. J. (2010) The ATP-dependent remodeler RSC transfers histone dimers and octamers through the rapid formation of an unstable encounter intermediate. *Biochemistry 49*, 9882–9890.

[116] Clapier, C. R., Langst, G., Corona, D. F., Becker, P. B., and Nightingale, K. P. (2001) Critical role for the histone H4 N terminus in nucleosome remodeling by ISWI. *Mol Cell Biol 21*, 875–883.

[117] Hamiche, A., Kang, J. G., Dennis, C., Xiao, H., and Wu, C. (2001) Histone tails modulate nucleosome mobility and regulate ATP-dependent nucleosome sliding by NURF. *Proc Natl Acad Sci U S A 98*, 14316–14321.

[118] Dang, W., Kagalwala, M. N., and Bartholomew, B. (2006) Regulation of ISW2 by concerted action of histone H4 tail and extranucleosomal DNA. *Mol Cell Biol 26*, 7388–7396.

[119] Shogren-Knaak, M., Ishii, H., Sun, J. M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science 311*, 844–847.

[120] Ferreira, H., Flaus, A., and Owen-Hughes, T. (2007) Histone modifications influence the action of Snf2 family remodelling enzymes by different mechanisms. *J Mol Biol 374*, 563–579.

[121] He, X., Fan, H. Y., Narlikar, G. J., and Kingston, R. E. (2006) Human ACF1 alters the remodeling strategy of SNF2h. *Journal of Biological Chemistry 281*, 28636–28647.

[122] Stockdale, C., Flaus, A., Ferreira, H., and Owen-Hughes, T. (2006) Analysis of nucleosome repositioning by yeast ISWI and Chd1 chromatin remodeling complexes. *Journal of Biological Chemistry 281*, 16279–16288.

[123] He, X., Fan, H. Y., Garlick, J. D., and Kingston, R. E. (2008) Diverse regulation of SNF2h chromatin remodeling by noncatalytic subunits. *Biochemistry 47*, 7025–7033.

[124] Dechassa, M. L., Hota, S. K., Sen, P., Chatterjee, N., Prasad, P., and Bartholomew, B. (2012) Disparity in the DNA translocase domains of SWI/SNF and ISW2. *Nucleic Acids Res*

[125] Hota, S. K., Bhardwaj, S. K., Deindl, S., Lin, Y. C., Zhuang, X., and Bartholomew, B. (2013) Nucleosome mobilization by ISW2 requires the concerted action of the ATPase and SLIDE domains. *Nat Struct Mol Biol 20*, 222–229.

[126] Liu, S., Harada, B. T., Miller, J. T., Le Grice, S. F. J., and Zhuang, X. (2010) Initiation complex dynamics direct the transitions between distinct phases of early HIV reverse transcription. *Nature structural & molecular biology 17*, 1453–60.

[127] Hwang, W. L., Deindl, S., Harada, B. T., and Zhuang, X. (2014) Histone H4 tail mediates allosteric regulation of nucleosome remodelling by linker DNA. *Nature 512*, 213.

[128] Marquet, R., Isel, C., Ehresmann, C., and Ehresmann, B. (1995) tRNAs as primer of reverse transcriptases. *Biochimie 77*, 113–124.

[129] Cordell, B., Swanstrom, R., Goodman, H. M., and Bishop, J. M. (1979) tRNATrp as primer for RNA-directed DNA polymerase: structural determinants of function. *J Biol Chem 254*, 1866–1874.

[130] Cobrinik, D., Soskey, L., and Leis, J. (1988) A retroviral RNA secondary structure required for efficient initiation of reverse transcription. *J Virol 62*, 3622–3630.

[131] Harrich, D., and Hooker, B. (2002) Mechanistic aspects of HIV-1 reverse transcription initiation. *Rev Med Virol 12*, 31–45.

[132] Isel, C., Lanchy, J. M., Le Grice, S. F., Ehresmann, C., Ehresmann, B., and Marquet, R. (1996) Specific initiation and switch to elongation of human immunodeficiency virus type 1 reverse transcription require the post-transcriptional modifications of primer tRNA3Lys. *EMBO J 15*, 917–924.

[133] Lanchy, J. M., Ehresmann, C., Le Grice, S. F., Ehresmann, B., and Marquet, R. (1996) Binding and kinetic properties of HIV-1 reverse transcriptase markedly differ during initiation and elongation of reverse transcription. *EMBO J 15*, 7178–7187.

[134] Lanchy, J. M., Keith, G., Le Grice, S. F., Ehresmann, B., Ehresmann, C., and Marquet, R. (1998) Contacts between reverse transcriptase and the primer strand govern the transition from initiation to elongation of HIV-1 reverse transcription. *J Biol Chem 273*, 24425–24432.

[135] Goldschmidt, V., Paillart, J. C., Rigourd, M., Ehresmann, B., Aubertin, A. M., Ehresmann, C., and Marquet, R. (2004) Structural variability of the initiation complex of HIV-1 reverse transcription. *J Biol Chem 279*, 35923–35931.

[136] Wilkinson, K. A., Gorelick, R. J., Vasa, S. M., Guex, N., Rein, A., Mathews, D. H., Giddings, M. C., and Weeks, K. M. (2008) High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biol 6*, e96.

[137] Liang, C., Rong, L., Gotte, M., Li, X., Quan, Y., Kleiman, L., and Wainberg, M. A. (1998) Mechanistic studies of early pausing events during initiation of HIV-1 reverse transcription. *J Biol Chem 273*, 21309–21315.

[138] Thrall, S. H., Krebs, R., Wohrl, B. M., Cellai, L., Goody, R. S., and Restle, T. (1998) Pre-steady-state kinetic characterization of RNA-primed initiation of transcription by HIV-1 reverse transcriptase and analysis of the transition to a processive DNA-primed polymerization mode. *Biochemistry 37*, 13349–13358.

[139] Lanchy, J. M., Isel, C., Keith, G., Le Grice, S. F., Ehresmann, C., Ehresmann, B., and Marquet, R. (2000) Dynamics of the HIV-1 reverse transcription complex during initiation of DNA synthesis. *J Biol Chem 275*, 12306–12312.

[140] Rong, L., Liang, C., Hsu, M., Guo, X., Roques, B. P., and Wainberg, M. A. (2001) HIV-1 nucleocapsid protein and the secondary structure of the binary complex formed between tRNA(Lys.3) and viral RNA template play different roles during initiation of (-) strand DNA reverse transcription. *J Biol Chem 276*, 47725–47732.

[141] DeStefano, J. J., Mallaber, L. M., Fay, P. J., and Bambara, R. A. (1993) Determinants of the RNase H cleavage specificity of human immunodeficiency virus reverse transcriptase. *Nucleic Acids Res 21*, 4330–4338.

[142] Iwatani, Y., Rosen, A. E., Guo, J., Musier-Forsyth, K., and Levin, J. G. (2003) Efficient initiation of HIV-1 reverse transcription in vitro. Requirement for RNA sequences downstream of the primer binding site abrogated by nucleocapsid protein-dependent primer-template interactions. *J Biol Chem 278*, 14185–14195.

[143] Isel, C., Westhof, E., Massire, C., Le Grice, S. F., Ehresmann, B., Ehresmann, C., and Marquet, R. (1999) Structural basis for the specificity of the initiation of HIV-1 reverse transcription. *EMBO J 18*, 1038–1048.

[144] Paillart, J. C., Dettenhofer, M., Yu, X. F., Ehresmann, C., Ehresmann, B., and Marquet, R. (2004) First snapshots of the HIV-1 RNA structure in infected cells and in virions. *J Biol Chem 279*, 48397–48403.

[145] Peliska, J. A., and Benkovic, S. J. (1992) Mechanism of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase. *Science (New York, N.Y.) 258*, 1112–8.

[146] Suo, Z., and Johnson, K. A. (1997) Effect of RNA secondary structure on the kinetics of DNA synthesis catalyzed by HIV-1 reverse transcriptase. *Biochemistry 36*, 12459–12467.

[147] Rein, A., Henderson, L. E., and Levin, J. G. (1998) Nucleic-acid-chaperon activity of retroviral nucleocapsid proteins: significance for viral replication. *Trends Biochem Sci 23*, 297–301.

[148] Thomas, J. A., and Gorelick, R. J. (2008) Nucleocapsid protein function in early infection processes. *Virus Res 134*, 39–63.

[149] Herschlag, D. (1995) RNA chaperones and the RNA folding problem. *J Biol Chem 270*, 20871–20874.

[150] Rodriguez-Rodriguez, L., Tsuchihashi, Z., Fuentes, G. M., Bambara, R. A., and Fay, P. J. (1995) Influence of human immunodeficiency virus nucleocapsid protein on synthesis and strand transfer by the reverse transcriptase in vitro. *J Biol Chem 270*, 15005–15011.

[151] Ji, X., Klarmann, G. J., and Preston, B. D. (1996) Effect of human immunodeficiency virus type 1 (HIV-1) nucleocapsid protein on HIV-1 reverse transcriptase activity in vitro. *Biochemistry 35*, 132–143.

[152] Wu, T., Datta, S. A., Mitra, M., Gorelick, R. J., Rein, A., and Levin, J. G. (2010) Fundamental differences between the nucleic acid chaperone activities of HIV-1 nucleocapsid protein and Gag or Gag-derived proteins: Biological implications. *Virology 405*, 556–567.

[153] Grohmann, D., Godet, J., Mely, Y., Darlix, J. L., and Restle, T. (2008) HIV-1 nucleocapsid traps reverse transcriptase on nucleic acid substrates. *Biochemistry 47*, 12230–12240.

[154] Wu, W., Henderson, L. E., Copeland, T. D., Gorelick, R. J., Bosche, W. J., Rein, A., and Levin, J. G. (1996) Human immunodeficiency virus type 1 nucleocapsid protein reduces reverse transcriptase pausing at a secondary structure near the murine leukemia virus polypurine tract. *J Virol 70*, 7132–7142.

[155] Tanchou, V., Gabus, C., Rogemond, V., and Darlix, J. L. (1995) Formation of stable and functional HIV-1 nucleocapsid complexes in vitro. *J Mol Biol 252*, 563–571.

[156] Lori, F., di Marzo Veronese, F., de Vico, A. L., Lusso, P., Reitz Jr., M. S., and Gallo, R. C. (1992) Viral DNA carried by human immunodeficiency virus type 1 virions. *J Virol 66*, 5067–5074.

[157] Trono, D. (1992) Partial reverse transcripts in virions from human immunodeficiency and murine leukemia viruses. *J Virol 66*, 4893–4900.

[158] Zhang, H., Dornadula, G., Orenstein, J., and Pomerantz, R. J. (2000) Morphologic changes in human immunodeficiency virus type 1 virions secondary to intravirion reverse transcription: evidence indicating that reverse transcription may not take place within the intact viral core. *J Hum Virol 3*, 165–172.

[159] Beerens, N., Groot, F., and Berkhout, B. (2001) Initiation of HIV-1 reverse transcription is regulated by a primer activation signal. *J Biol Chem 276*, 31247–31256.

[160] Houzet, L., Morichaud, Z., Didierlaurent, L., Muriaux, D., Darlix, J. L., and Mougel, M. (2008) Nucleocapsid mutations turn HIV-1 into a DNA-containing virus. *Nucleic Acids Res 36*, 2311–2319.

[161] Thomas, J. A., Bosche, W. J., Shatzer, T. L., Johnson, D. G., and Gorelick, R. J. (2008) Mutations in human immunodeficiency virus type 1 nucleocapsid protein zinc fingers cause premature reverse transcription. *J Virol 82*, 9318–9328.

[162] Rigourd, M., Lanchy, J. M., Le Grice, S. F., Ehresmann, B., Ehresmann, C., and Marquet, R. (2000) Inhibition of the initiation of HIV-1 reverse transcription by 3'-azido-3'-deoxythymidine. Comparison with elongation. *J Biol Chem 275*, 26944–26951.

[163] Hooker, C. W., Lott, W. B., and Harrich, D. (2001) Inhibitors of human immunodeficiency virus type 1 reverse transcriptase target distinct phases of early reverse transcription. *J Virol 75*, 3095–3104.

[164] Le Grice, S. F. J., Cameron, C. E., and Benkovic, S. J. (1995) Purification and characterization of human immunodeficiency virus type 1 reverse transcriptase. *Methods Enzymol 262*, 130–147.

[165] Rausch, J. W., Sathyanarayana, B. K., Bona, M. K., and Le Grice, S. F. (2000) Probing contacts between the ribonuclease H domain of HIV-1 reverse transcriptase and nucleic acid by site-specific photocross-linking. *J Biol Chem 275*, 16015–16022.

[166] Lapham, J., and Crothers, D. M. (1996) RNase H cleavage for processing of in vitro transcribed RNA for NMR studies and RNA ligation. *RNA 2*, 289–296.

[167] Rasnik, I., McKinney, S. A., and Ha, T. (2006) Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat Methods 3*, 891–893.

[168] Venezia, C. F., Howard, K. J., Ignatov, M. E., Holladay, L. A., and Barkley, M. D. (2006) Effects of efavirenz binding on the subunit equilibria of HIV-1 reverse transcriptase. *Biochemistry 45*, 2779–2789.

[169] Kapanidis, A. N., Lee, N. K., Laurence, T. A., Doose, S., Margeat, E., and Weiss, S. (2004) Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc Natl Acad Sci U S A 101*, 8936–8941.

[170] Thompson, R. E., Larson, D. R., and Webb, W. W. (2002) Precise Nanometer Localization Analysis for Individual Fluorescent Probes. *Biophys. J. 82*, 2775–2783.

[171] Yildiz, A., Forkey, J. N., McKinney, S. A., Ha, T., Goldman, Y. E., and Selvin, P. R. (2003) Myosin V walks hand-over-hand single fluorophore imaging with 1.5-nm localization. *Science 300*, 2061–2065.

[172] Varga-Weisz, P. D., Wilm, M., Bonte, E., Dumas, K., Mann, M., and Becker, P. B. (1997) Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature 388*, 598–602.

[173] Tsukiyama, T., Palmer, J., Landel, C. C., Shiloach, J., and Wu, C. (1999) Characterization of the Imitation Switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. *Genes Dev 13*, 686–697.

[174] Yamada, K., Frouws, T. D., Angst, B., Fitzgerald, D. J., DeLuca, C., Schimmele, K., Sargent, D. F., and Richmond, T. J. (2011) Structure and mechanism of the chromatin remodelling factor ISW1a. *Nature 472*, 448–453.

[175] Clapier, C. R., and Cairns, B. R. (2012) Regulation of ISWI involves inhibitory modules antagonized by nucleosomal epitopes. *Nature 492*, 280–284.

[176] LeRoy, G., Loyola, A., Lane, W. S., and Reinberg, D. (2000) Purification and characterization of a human factor that assembles and remodels chromatin. *J Biol Chem 275*, 14787–14790.

[177] Poot, R. A., Dellaire, G., Hulsmann, B. B., Grimaldi, M. A., Corona, D. F., Becker, P. B., Bickmore, W. A., and Varga-Weisz, P. D. (2000) HuCHRAC, a human ISWI chromatin remodelling complex contains hACF1 and two novel histone-fold proteins. *EMBO J 19*, 3377–3387.

[178] Fyodorov, D. V., and Kadonaga, J. T. (2002) Binding of Acf1 to DNA involves a WAC motif and is important for ACF-mediated chromatin assembly. *Mol Cell Biol 22*, 6344–6353.

[179] Gelbart, M. E., Rechsteiner, T., Richmond, T. J., and Tsukiyama, T. (2001) Interactions of Isw2 chromatin remodeling complex with nucleosomal arrays: analyses using recombinant yeast histones and immobilized templates. *Mol Cell Biol 21*, 2098–2106.

[180] Lowary, P. T., and Widom, J. (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol 276*, 19–42.

[181] Luger, K., Rechsteiner, T. J., and Richmond, T. J. (1999) Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol 304*, 3–19.

[182] Aitken, C. E., Marshall, R. A., and Puglisi, J. D. (2008) An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys J 94*, 1826–1835.

[183] Kerssemakers, J. W., Munteanu, E. L., Laan, L., Noetzel, T. L., Janson, M. E., and Dogterom, M. (2006) Assembly dynamics of microtubules at molecular resolution. *Nature 442*, 709–712.

[184] Bazett-Jones, D. P., Cote, J., Landel, C. C., Peterson, C. L., and Workman, J. L. (1999) The SWI/SNF complex creates loop domains in DNA and polynucleosome arrays and can disrupt DNA-histone contacts within these domains. *Mol Cell Biol 19*, 1470–1478.

[185] Floer, M., Wang, X., Prabhu, V., Berrozpe, G., Narayan, S., Spagna, D., Alvarez, D., Kendall, J., Krasnitz, A., Stepansky, A., Hicks, J., Bryant, G. O., and Ptashne, M. (2010) A RSC/nucleosome complex determines chromatin architecture and facilitates activator binding. *Cell 141*, 407–418.

[186] Liu, N., Peterson, C. L., and Hayes, J. J. (2011) SWI/SNF- and RSC-catalyzed nucleosome mobilization requires internal DNA loop translocation within nucleosomes. *Mol Cell Biol 31*, 4165–4175.

[187] Sirinakis, G., Clapier, C. R., Gao, Y., Viswanathan, R., Cairns, B. R., and Zhang, Y. (2011) The RSC chromatin remodelling ATPase translocates DNA with high force and small step size. *EMBO J 30*, 2364–2372.

[188] Hota, S. K., and Bartholomew, B. (2011) Diversity of operation in ATP-dependent chromatin remodelers. *Biochimica et biophysica acta 1809*, 476–87.

[189] Cairns, B. R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B., and Kornberg, R. D. (1996) RSC, an essential, abundant chromatin-remodeling complex. *Cell 87*, 1249–1260.

[190] Flaus, A., and Owen-Hughes, T. (2003) Dynamic properties of nucleosomes during thermal and ATP-driven mobilization. *Mol Cell Biol 23*, 7767–7779.

[191] Shundrovsky, A., Smith, C. L., Lis, J. T., Peterson, C. L., and Wang, M. D. (2006) Probing SWI/SNF remodeling of the nucleosome by unzipping single DNA molecules. *Nat Struct Mol Biol 13*, 549–554.

[192] Li, G., and Widom, J. (2004) Nucleosomes facilitate their own invasion. *Nature structural & molecular biology 11*, 763–9.

[193] Myong, S., Bruno, M. M., Pyle, A. M., and Ha, T. (2007) Spring-loaded mechanism of DNA unwinding by hepatitis C virus NS3 helicase. *Science 317*, 513–516.

[194] Gu, M., and Rice, C. M. (2010) Three conformational snapshots of the hepatitis C virus NS3 helicase reveal a ratchet translocation mechanism. *Proc Natl Acad Sci U S A 107*, 521–528.

[195] Rajagopal, V., Gurjar, M., Levin, M. K., and Patel, S. S. (2010) The protease domain increases the translocation stepping efficiency of the hepatitis C virus NS3-4A helicase. *J Biol Chem 285*, 17821–17832.

[196] Cheng, W., Arunajadai, S. G., Moffitt, J. R., Tinoco Jr., I., and Bustamante, C. (2011) Single-base pair unwinding and asynchronous RNA release by the hepatitis C virus NS3 helicase. *Science 333*, 1746–1749.

[197] Wittmeyer, J., Saha, A., and Cairns, B. (2004) DNA translocation and nucleosome remodeling assays by the RSC chromatin remodeling complex. *Methods Enzymol 377*, 322–343.

[198] Partensky, P. D., and Narlikar, G. J. (2009) Chromatin remodelers act globally, sequence positions nucleosomes locally. *J Mol Biol 391*, 12–25.

[199] Lee, K. M., and Narlikar, G. (2001) Assembly of nucleosomal templates by salt dialysis. *Curr Protoc Mol Biol Chapter 21*, Unit 21 6.

[200] Deindl, S., and Zhuang, X. (2012) Monitoring conformational dynamics with single-molecule fluorescence energy transfer: applications in nucleosome remodeling. *Methods in enzymology 513*, 59–86.

[201] Lee, S., Lee, J., and Hohng, S. (2010) Single-molecule three-color FRET with both negligible spectral overlap and long observation time. *PloS one 5*, e12270.

[202] Gutiérrez, J. L., Chandy, M., Carrozza, M. J., and Workman, J. L. (2007) Activation domains drive nucleosome eviction by SWI/SNF. *The EMBO journal 26*, 730–40.

[203] Chandy, M., Gutiérrez, J. L., Prochasson, P., and Workman, J. L. (2006) SWI/SNF displaces SAGA-acetylated nucleosomes. *Eukaryotic cell 5*, 1738–47.

[204] Lorch, Y., Griesenbeck, J., Boeger, H., Maier-Davis, B., and Kornberg, R. D. (2011) Selective removal of promoter nucleosomes by the RSC chromatin-remodeling complex. *Nature structural & molecular biology 18*, 881–5.

[205] Lorch, Y., Maier-Davis, B., and Kornberg, R. D. (2014) Role of DNA sequence in chromatin remodeling and the formation of nucleosome-free regions. *Genes & development 28*, 2492–7.

[206] Bohm, V., Hieb, A. R., Andrews, A. J., Gansen, A., Rocker, A., Toth, K., Luger, K., and Langowski, J. (2011) Nucleosome accessibility governed by the dimer/tetramer interface. *Nucleic Acids Res 39*, 3093–3102.

[207] Crawford, D. J., Hoskins, A. A., Friedman, L. J., Gelles, J., and Moore, M. J. (2007) Visualizing the splicing of single pre-mRNA molecules in whole cell extract. *RNA 14*, 170–179.

[208] Hoskins, A. A., Friedman, L. J., Gallagher, S. S., Crawford, D. J., Anderson, E. G., Wombacher, R., Ramirez, N., Cornish, V. W., Gelles, J., and Moore, M. J. (2011) Ordered and dynamic assembly of single spliceosomes. *Science (New York, N.Y.) 331*, 1289–95.

[209] Jain, A., Liu, R., Ramani, B., Arauz, E., Ishitsuka, Y., Ragunathan, K., Park, J., Chen, J., Xiang, Y. K., and Ha, T. (2011) Probing cellular protein complexes using single-molecule pull-down. *Nature 473*, 484–8.

[210] Griesenbeck, J., Boeger, H., Strattan, J. S., and Kornberg, R. D. (2003) Affinity purification of specific chromatin segments from chromosomal loci in yeast. *Molecular and cellular biology 23*, 9275–82.

[211] Zhang, Z., Revyakin, A., Grimm, J. B., Lavis, L. D., and Tjian, R. (2014) Single-molecule tracking of the transcription cycle by sub-second RNA detection. *eLife 3*, e01775.

[212] Lavigne, M., Roux, P., Buc, H., and Schaeffer, F. (1997) DNA curvature controls termination of plus strand DNA synthesis at the centre of HIV-1 genome. *Journal of molecular biology 266*, 507–24.

[213] Lavigne, M., and Buc, H. (1999) Compression of the DNA minor groove is responsible for termination of DNA synthesis by HIV-1 reverse transcriptase. *Journal of molecular biology 285*, 977–95.

[214] Lavigne, M., Polomack, L., and Buc, H. (2001) Structures of complexes formed by HIV-1 reverse transcriptase at a termination site of DNA synthesis. *The Journal of biological chemistry 276*, 31439–48.

[215] Berdis, A. J., Stetor, S. R., LeGrice, S. F., and Barkley, M. D. (2001) Molecular mechanism of sequence-specific termination of lentiviral replication. *Biochemistry 40*, 12140–9.

[216] Lavigne, M., Polomack, L., and Buc, H. (2001) DNA synthesis by HIV-1 reverse transcriptase at the central termination site: a kinetic study. *The Journal of biological chemistry 276*, 31429–38.

[217] Aathavan, K., Politzer, A. T., Kaplan, A., Moffitt, J. R., Chemla, Y. R., Grimes, S., Jardine, P. J., Anderson, D. L., and Bustamante, C. (2009) Substrate interactions and promiscuity in a viral DNA packaging motor. *Nature 461*, 669–73.

[218] Lia, G., Praly, E., Ferreira, H., Stockdale, C., Tse-Dinh, Y. C., Dunlap, D., Croquette, V., Bensimon, D., and Owen-Hughes, T. (2006) Direct observation of DNA distortion by the RSC complex. *Mol Cell 21*, 417–425.