Published in final edited form as:

Curr Opin Struct Biol. 2010 February ; 20(1): 121–127. doi:10.1016/j.sbi.2009.12.008.

Stepwise translocation of nucleic acid motors

Sua Myong^{1,3} and Taekjip Ha^{2,3,4}

¹Department of Bioengineering, University of Illinois at Urbana-Champaign

²Department of Physics, University of Illinois at Urbana-Champaign

³Institute for Genomic Biology, University of Illinois at Urbana-Champaign

⁴Howard Hughes Medical Institute, Urbana, Illinois 61801 USA

Summary

Recent single molecule studies have made a significant contribution to the understanding of the molecular mechanism involved in the movement of motor proteins which process DNA and RNA. Measurement of stepsize in two disparate motors, NS3 helicase and ribosome both revealed three basepair steps which consist of three hidden substeps. Combined with previous structural studies, NS3 is likely taking a single nucleotide step of translocation coupled to one ATP binding event and this mode may be conserved in multitude of helicases. Such a stepwise translocation movement appears to occur through main contacts with the phosphate backbone. Double stranded RNA and DNA motor, RIG-I and Φ 29 respectively showed translocation on a duplex while tracking exclusively a single stranded RNA/DNA in a directional manner, 5' to 3' in both cases. Spontaneous dynamics displayed by ribosome ratcheting and SSB (single stranded DNA binding protein) diffusing on DNA were rectified by interacting cofactors and proteins, EF-G and RecA respectively.

Introduction

Molecular motors are much like cars driving on a highway. They are often directional with variable speed and their processivity is determined by how long they stay on track before trailing off to an exit. Many motors consume energy to translocate on their track and thus slow down or stop when lacking sufficient amount of fuel as in the case of motor vehicles. Such behaviors have been examined through the glasses of highly sensitive single molecular binoculars and structural eye pieces to reveal hidden steps and mechanistic details involved in the movement of motor proteins. The new insights from these studies imply a mechanism similar to the gears controlling a mechanical clock where the movement of the minute and hour hands are under the control of the pendulum which converts the swinging motion into rotational movement of the corresponding wheels for the two hands. Similarly, the mechanical translocation of a motor protein is governed by ATP hydrolysis which modulates their conformation, thereby converting chemical energy into mechanical movement in a stepwise manner. This review attempts to articulate some key findings reported in the past two years by single molecule studies on proteins which process nucleic acids in various contexts. The proteins under review include helicases, a viral nucleic acid packaging motor, reverse transcriptase, single stranded DNA binding proteins and the ribosome, all of which are well

© 2010 Elsevier Ltd. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

known proteins whose motion on the DNA or RNA lattice are important for the processing of nucleic acids.

Step size as small as a single nucleotide

The step size of a motor protein can be defined as the physical distance moved by the protein on the track during a biochemical cycle in which chemical energy consumption is coupled to a mechanical cycle of advancing the protein. Previous biochemical studies have revealed variable stepsizes from DNA motors ranging from a single nucleotide (nt) to 23 nt (1-4), indicating that the intrinsic stepsize of such motors can be as tiny as a single base of DNA or RNA. Proteins involved in replication and transcription are likely to process nucleotides as a chemical entity one at a time to ensure an accurate information transfer (5-6). In contrast, proteins such as a virus packaging motor may treat more than one nucleotide as a physical unit, thus process it with less rigorous measures (7).

A recent study demonstrated that a single nucleotide is likely the smallest step possible for proteins which participate in genome processing. Using single molecule FRET, the unwinding steps of NS3, a helicase found in Hepatitis C virus, were detected. NS3 unwound DNA in discrete steps of about three basepairs. Interestingly, the dwell time in each plateau was not exponential, as would be expected if there is a single rate-limiting step. Instead, the dwell time histogram showed a rising phase and a decaying phase, most consistent with three irreversible hidden steps before each step of 3 bp unwinding (8). Combining with the structural data on NS3 without ATP (9) and related RNA helicases with ATP (10-11), it was proposed that 3 bp are unwound in a burst after accumulation of elastic energy in the protein-DNA complex during three substeps, one nt per ATP, of NS3 translocation on the DNA lattice (Figure 1). The identical set of the structural arrangement was also present in other helicases including Rep, PcrA, UvrD and Hel308, strongly suggesting that a single nucleotide step may be a universal mechanism shared amongst numerous helicases (Figure 2).

A strikingly similar observation was made for the ribosome during translation of mRNA, visualized by an optical tweezers technique (12). The mRNA formed a hairpin, and ribosome motion on the RNA codon-by-codon was captured upon unwinding of the hairpin in 3 bp increments. The dwell time histogram was again best explained by three hidden steps occurring before ribosome translocation over a single codon (Figure 3). Presently, the origin of these hidden steps is unclear but it will be interesting to test if similar kinetics can be observed even when the ribosome does not need to unwind duplexes for translocation.

Phosphate backbone is the main contact

Most DNA and RNA motor proteins do not bind in a sequence specific manner and the structure of double stranded DNA and RNA reveals an apparent difficulty for a protein to reach into the grooves for approaching bases. It is perhaps for this reason that a multitude of proteins including helicases appear to translocate primarily on the phosphate backbone. The aforementioned single nucleotide stepping of NS3 was based on the structural data which presented two phosphate contacting threonine (Thr) residues on two domains of NS3 which are highly conserved and indispensable for the ATPase and unwinding activity of NS3 (9). These Thr residues at the elbows formed by a beta strand and an alpha helix each are three nucleotides apart in the absence of ATP, but shows two nucleotide distance in ATP bound structure of its close homologs, Vasa and eIF4AIII (10-11) (Figure 4). This single nucleotide distance change mediated by two Thr-phosphate contacts are conserved in various helicases as discussed above and may represent a unifying mechanism for motor proteins to walk on the phosphate backbone as a primary translocating track (Figure 1).

The significance of the phosphate contact was also demonstrated by Aathavan et al who investigated the interaction between a viral packaging motor from $\Phi 29$ and its substrate, double stranded DNA (dsDNA) using the optical tweezers technique (13). The directional movement of dsDNA through the ring shaped ATPase was greatly hindered when methylphosphonate modified bases were interjected in its tracking strand. Methylphosphonate replacement of phosphate moiety removes negative charges from the backbone, hence the phosphate contacts with the neighboring amino acid of the packaging motor may lose their grip on the phosphates, leading to stalling.

Single strand tracking by double stranded DNA and RNA motors

The concept of directionality becomes uncertain when it comes to motor proteins which translocate on double stranded DNA or RNA. In most cases, they are multimeric complexes composed of two or more units of the proteins which make contact with both strands by encompassing the entire double strand. Recently, we discovered that RIG-I, an antiviral sensor and an ATPase is a robust translocase on double stranded RNA (dsRNA) (14). RIG-I recognizes and binds double stranded RNA of viruses as a pathogenic signature and elicits an antiviral signaling cascade. When subjected to various substrates, RIG-I displayed a strong preference to bind dsRNA as a dimer whereas it did not bind or translocate on dsDNA. On RNA/DNA heteroduplexes, however RIG-I showed translocation activity comparable to dsRNA, indicating that single strand of RNA is sufficient to serve as a tracking strand and to support RIG-I translocation. Directionality of RIG-I translocation was obtained by observing the translocation on RNA:DNA heteroduplex of a truncation mutant of RIG-I without the N-terminal signaling domain. The fluorescence signal changes were consistent with 5' to 3' directional translocation on the RNA strand. A similar bias to track one strand of a duplex DNA was shown in the study of the viral packaging motor, $\Phi 29$. The translocation of double stranded DNA through a pentameric ring of ATPases was not disrupted when the methyl phosphonate modification was placed on 3'-5' directional strand. The same degree of modification (10 nucleotide), however perturbed the translocation dramatically on the counter strand in 5' to 3' in the direction of viral packaging, suggesting a primary contact with the single strand DNA in the 5' to 3' direction (13).

DNA vs. RNA motors

Most motor proteins are highly selective to either DNA or RNA as a preferred substrate although there are exceptions such as NS3 helicase which unwinds both RNA and DNA (15).

Liu et al presented an interesting case of a motor protein, reverse transcriptase (RT) from HIV which transcribes DNA from viral ssRNA template. When given a RNA:DNA heteroduplex, RT showed a random sliding behavior which is stabilized by the presence of a cognate dNTP, promoting the positioning of RT at the 3' termini of the DNA where the dNTP needs to be incorporated. Using a fluorescently labeled RT, they also showed that RT undergoes a rapid conformational change to engage into a state where it can carry out polymerization. When subjected to a dsDNA construct, the same sliding behavior and the dNTP driven positioning occurred, indicating that the RT sliding is allowed in both types of duplex compositions, consistent with its in vivo functions of reverse transcription and plus-strand DNA synthesis. Hence, the substrate selectivity of RT is toward RNA:DNA and DNA:DNA, yet the polymerization mode ensues only at the trailing end of 3' DNA thereby achieving its catalytic specificity (16). In terms of a primary contact, RT is likely to engage more tightly to ssDNA strand in search for the 3' terminal polymerization site while loosely gripping the complementary ssRNA. Based on its bidirectional movement, it would be of interest to compare the ssDNA contact surface of RT and other directional ssDNA translocases to understand the molecular basis of directional motors.

A similar substrate specificity is also revealed in the RIG-I study mentioned above. RIG-I forms a dimer (17) and recognizes double stranded RNA, a molecular pattern present in viruses (18). While RIG-I did not bind or translocate on dsDNA, it bound and translocated on dsRNA as well as the RNA:DNA heteroduplexes although the binding affinity was lower (14). The activity of RIG-I on heteroduplex construct indicates that RIG-I tracks primarily one ssRNA even in the context of dsRNA, hence ssRNA is the minimum requirement of RIG-I translocation. The ability of RIG-I to track ssRNA, yet not unwind the duplex may be an intriguing attribute exclusive to this family of proteins. Structural studies of RIG-I ATPase domain bound to RNA substrate should illuminate the distinct features that allow this translocation mode.

Spontaneous conformational dynamics utilized for function

While catalytic reactions such as unwinding by a helicase are strictly coupled to ATP (or NTP) hydrolysis, some motor proteins exhibit dynamic movements even in the absence of an energy input. Often these movements are programmed to maximize its activity upon encountering an ideal active site or a correct substrate. The Ribosome is a universal translational machinery which consists of 30S and 50S subunits. Although there has been much prediction about the intersubunit rotations, also known as 'ratcheting', correlated with the translation cycle, in particular the translocation of the ribosome on the mRNA lattice (19-24) the recent report by Cornish et al marks the first real-time observation of this phenomenon (25). The two ribosome subunits rotate spontaneously between the two conformational states referred to as classical and hybrid states (Figure 3). The two subunits modulated their bias to either state depending on the applied reaction conditions, confirming the two states as the intermediate steps involved in translation. Remarkably, removing the amino-acid from the tRNA bound to the so called 'P-site' led to the 100 fold acceleration of the ratcheting dynamics, showing that a single amino-acid change could lubricate the 2.6 Mega Dalton machine! It is likely that this ratcheting dynamics is equivalent to the inchworming motion of helicases (Figure 2). Just like ATP binding and hydrolysis modulate the opening and closing of two RecA like domains on the helicases to cause a directional movement, binding and GTP hydrolysis of an elongation factor known as EF-G may rectify the spontaneous ribosome dynamics to enable it to translocate on the mRNA track in a directional motor. In this sense, we may view the ribosome as an inchworming Brownian ratchet (Figure 3).

Another example of spontaneous dynamics was manifested by SSB (Single Stranded DNA Binding protein). It was shown that that *E. coli* SSB diffuses on ssDNA with an estimated diffusion coefficient of $\sim 300 \text{ (nt)}^2/\text{s}$ and with an average step size of about 3 nt (26). The diffusion occurred at least of 65 nt and could be used to resolve secondary structures dynamically, and to enhance RecA filament growth. This first example of any protein diffusing on single stranded DNA provides a new model with which many proteins that bind to single stranded DNA or RNA via multiple binding interactions may redistribute themselves after initial binding and coordinate with other proteins for access to the DNA. Similar to the Brownian inchworm model for the ribosome movement, the spontaneous motion of SSB protein on the DNA can be rectified via the directional 3' to 5' extension of the RecA filament. Of note, RecA filament grows and shrinks in units of single monomers (27) and the binding site size of RecA monomer is 3 nt (28) which matches the estimated diffusion step size of SSB. Because recent single molecule analysis suggested that the eukaryotic homolog Rad51 also extends and shrinks its filament in single monomer units (29-30), it is plausible that the mechanism discovered from the studies of *E. coli* SSB and RecA is also applicable to the eukaryotic systems. The intersubunit rotation of the ribosome and the diffusion of SSB on single stranded DNA are two different modes of spontaneous dynamics, yet both motions are likely to offer an advantage for their function.

Conclusion

DNA and RNA are much like molecular highways on which motor proteins run. They have a distinct polarity with different chemistry which attracts specific motors to latch on and travel. The motor proteins can be classified into two categories for the purpose of our discussion. The first class is the genome processing proteins including helicases, ribosomes and reverse transcriptases which partake in gene transfer reactions. Second class is the proteins which process RNA and DNA in a non-genomic manner such as the viral packaging motor and the antiviral receptor, RIG-I. Single molecule approaches have been successfully employed to extract the fundamental mechanisms underlying the function of both types of proteins. As the first class proteins, NS3 and ribosome revealed highly regular stepsize of three base pairs, each of which are composed of three substeps representing single nucleotide steps of translocation. For the second class proteins, RIG-I and the viral packaging motor, Φ 29 are both double stranded RNA and DNA translocases respectively, yet both translocate on one of the double strands in a directional manner. Other characteristics such as phosphate contact and DNA vs. RNA selectivity are attributed to both classes of proteins. Future studies would reveal more exciting details about the motor protein mechanism, which will bring about further understanding of these proteins.

References

1. Dillingham MS, Wigley DB, Webb MR. Biochemistry Jan 11;2000 39:205. [PubMed: 10625495]
2. Bianco PR, Kowalczykowski SC. Nature May 18;2000 405:368. [PubMed: 10830968]
3. Serebrov V, Pyle AM. Nature Jul 22;2004 430:476. [PubMed: 15269774]
4. Ali JA, Lohman TM. Science Jan 17;1997 275:377. [PubMed: 8994032]
5. Christian TD, Romano LJ, Rueda D. Proc Natl Acad Sci U S A. Dec 2;2009
6. Greenleaf WJ, Block SM. Science Aug 11;2006 313:801. [PubMed: 16902131]
- “7. Moffitt JR, et al. Observation of 10 basepair stepsize of a viral packaging motor, Φ 29 using an optical tweezers measurement. Nature Jan 22;2009 457:446. [PubMed: 19129763]
- “8. Myong S, Bruno MM, Pyle AM, Ha T. Observation of three base pair unwinding step of NS3 helicase which is composed of three single nucleotide substeps, using single molecule FRET assay. Science Jul 27;2007 317:513. [PubMed: 17656723]
9. Kim JL, et al. Structure Jan 15;1998 6:89. [PubMed: 9493270]
10. Andersen CB, et al. Science Sep 29;2006 313:1968. [PubMed: 16931718]
11. Sengoku T, Nureki O, Nakamura A, Kobayashi S, Yokoyama S. Cell Apr 21;2006 125:287. [PubMed: 16630817]
- “12. Wen JD, et al. Observation of three nucleotide step (one codon) of a ribosome which consists of three substeps monitored by optical tweezers technique. Nature Apr 3;2008 452:598. [PubMed: 18327250]
- “13. Aathavan K, et al. Report on the significance of phosphate backbone of DNA for directional translocation of Φ 29 using an optical tweezers measurement. Nature Oct 1;2009 461:669. [PubMed: 19794496]
- “14. Myong S, et al. Discovery of a robust translocational activity of an antiviral receptor, RIG-I on double stranded RNA using single molecule protein induced fluorescence enhancement (PIFE) assay. Science Feb 20;2009 323:1070. [PubMed: 19119185]
15. Pang PS, Jankowsky E, Planet PJ, Pyle AM. EMBO J Mar 1;2002 21:1168. [PubMed: 11867545]
- “16. Liu S, Abbondanzieri EA, Rausch JW, Le Grice SF, Zhuang X. Observation of spontaneous sliding and dynamic shuttling of reverse transcriptase of HIV using single molecule FRET measurement. Science Nov 14;2008 322:1092. [PubMed: 19008444]
17. Cui S, et al. Mol Cell Feb 1;2008 29:169. [PubMed: 18243112]
18. Hornung V, et al. Science Nov 10;2006 314:994. [PubMed: 17038590]
19. Cukras AR, Southworth DR, Brunelle JL, Culver GM, Green R. Mol Cell Aug;2003 12:321. [PubMed: 14536072]

20. Fredrick K, Noller HF. Science May 16;2003 300:1159. [PubMed: 12750524]
21. Frank J, Agrawal RK. Nature Jul 20;2000 406:318. [PubMed: 10917535]
22. Munro JB, Altman RB, O'Connor N, Blanchard SC. Mol Cell Feb 23;2007 25:505. [PubMed: 17317624]
23. Kim HD, Puglisi JD, Chu S. Biophys J Nov 15;2007 93:3575. [PubMed: 17693476]
24. Fei J, Kosuri P, MacDougall DD, Gonzalez RL Jr. Mol Cell May 9;2008 30:348. [PubMed: 18471980]
25. Cornish PV, Ermolenko DN, Noller HF, Ha T. Observation of spontaneous intersubunit rotation of a ribosome using single molecule FRET measurement. Mol Cell Jun 6;2008 30:578. [PubMed: 18538656]
26. Roy R, Kozlov AG, Lohman TM, Ha T. Discovery of diffusional movement of single stranded DNA binding protein (SSB) by single molecule FRET visualization. Nature Oct 22;2009 461:1092. [PubMed: 19820696]
27. Joo C, et al. Cell Aug 11;2006 126:515. [PubMed: 16901785]
28. Chen Z, Yang H, Pavletich NP. Nature May 22;2008 453:489. [PubMed: 18497818]
29. Arata H, et al. Proc Natl Acad Sci U S A Nov 17;2009 106:19239. [PubMed: 19884492]
30. van Mameren J, et al. Nature Feb 5;2009 457:745. [PubMed: 19060884]

Figure 1.

A. NS3 unwinding was observed by single molecule FRET measurement where high FRET from two dyes placed at the duplex junction decreased stepwise during unwinding. Six steps obtained from 18 bp unwinding reveals 3 bp stepsize. B. Dwell time analysis revealed a nonexponential decay which indicates three hidden substeps within 3 bp unwinding steps. C. NS3 takes three steps of single nucleotide translocation coupled to ATP binding, which builds strain on the DNA-protein interface. The strain is released about every three basepairs, leading to a burst of 3bp unwinding observed in A.

A Inchworm movement of NS3 translocation

B NS3 helicase structure

Figure 2.

A. NS3 and related helicases use two RecA like domains to translocate on the nucleic acid track, one nucleotide at a time in an inchworm like manner. The two domains 'walk' on the nucleic acid track by alternating affinities, mediated by ATP binding and ADP release. Dark and fair colored units represent tight and loose binding of two RecA like domains respectively. B. Two RecA like domains are depicted in the NS3 crystal structure (PDB: 1A1V) (9) with the same color designation as in A.

Figure 3.

A. In the presence of EF-G elongation factor, the two rotating subunits of the ribosome translocate on mRNA taking 3 nucleotide steps corresponding to the size of one codon. B. The 30S and 50S subunits are represented on the ribosome structure. C. The ribosome undergoes a spontaneous ratcheting motion in the absence of EF-G.

Structural basis for single nucleotide step

Figure 4.

A. NS3 crystallized without ATP shows two well conserved threonine residues (T269, T411 in blue) contacting the backbone phosphate 3 nts apart (PDB: 1A1V) (9). W501 (yellow), also known as “gatekeeper” is base stacked at the 3' end of the bound nucleotide. B, C. Both *Drosophila* Vasa (PDB:2DB3)(11) and eIF4AIII (PDB:2HYI) (10), two other SF2 helicases were co-crystallized with ATP analogs, AMPPNP and ADPNP, respectively. The structurally equivalent threonine residues (T375, T546 in Vasa, T164, T334 in eIF4AIII) show phosphate contacts in 2 nts distance, suggesting 1 nt movement coupled to one ATP binding in between domain 1 and 2, bringing the two RecA-like domains together.