Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

CHAPTER SIX

Molecular Mechanics of RNA Translocases

Steve C. Ding* *and* Anna Marie Pyle[†,‡]

Contents

1. Introduction 132
   1.1. The protein domains of an RNA translocase: The RecA motor fold and companion domains 133
2. Major Examples of Monomeric RNA Translocases 134
   2.1. The NS3/NPH-II family of SF2 proteins: $3' \rightarrow 5'$ translocation enzymes 134
   2.2. The UPf1 family of SF1 proteins: $5' \rightarrow 3'$ translocation enzymes 138
   2.3. Double-stranded translocation by monomeric RNA helicases: The RIG-I/Dicer family 140
3. Concluding Remarks 144
Acknowledgments 145
References 145

Abstract

Historically, research on RNA helicase and translocation enzymes has seemed like a footnote to the extraordinary progress in studies on DNA-remodeling enzymes. However, during the past decade, the rising wave of activity in RNA science has engendered intense interest in the behaviors of specialized motor enzymes that remodel RNA molecules. Functional, mechanistic, and structural investigations of these RNA enzymes have begun to reveal the molecular basis for their key roles in RNA metabolism and signaling. In this chapter, we highlight the structural and mechanistic similarities among monomeric RNA translocase enzymes, while emphasizing the many divergent characteristics that have caused this enzyme family to become one of the most important in metabolism and gene expression.

* Infectious and Inflammatory Disease Center, Sanford-Burnham Medical Research Institute, La Jolla, California, USA
† Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, Connecticut, USA
‡ Howard Hughes Medical Institute, Yale University, New Haven, Connecticut, USA

Methods in Enzymology, Volume 511
ISSN 0076-6879, DOI: 10.1016/B978-0-12-396546-2.00006-1
© 2012 Elsevier Inc.
All rights reserved.

1. Introduction

Progress in the field of RNA remodeling has largely resulted from the multidisciplinary nature of experimentation in this area. A combination of genetics, biochemistry, and biophysical studies in many different laboratories has established the biological functions of many RNA remodeling enzymes, the unique structural units upon which the enzymes are built, and the specific chemical and structural determinants by which RNA translocases and helicases identify their nucleic acid targets. As a result, several unifying themes and structural classes have emerged. We have limited the scope of this review to focus on monomeric RNA translocases that are phylogenetically classified as belonging to "helicase superfamily 1" (SF1) or "helicase superfamily 2" (SF2), which are nonetheless quite similar in sequence and morphology.

To date, enzymes in these families translocate unidirectionally along RNA in a processive manner. In order to exemplify the three major types of translocation behaviors that are observed for these enzymes, we focus on three different enzymes that share structural similarities while displaying divergent functional behavior (Fig. 6.1). Specifically, we discuss the following: Nonstructural protein 3 from hepatitis C virus (HCV NS3), a $3' \rightarrow 5'$

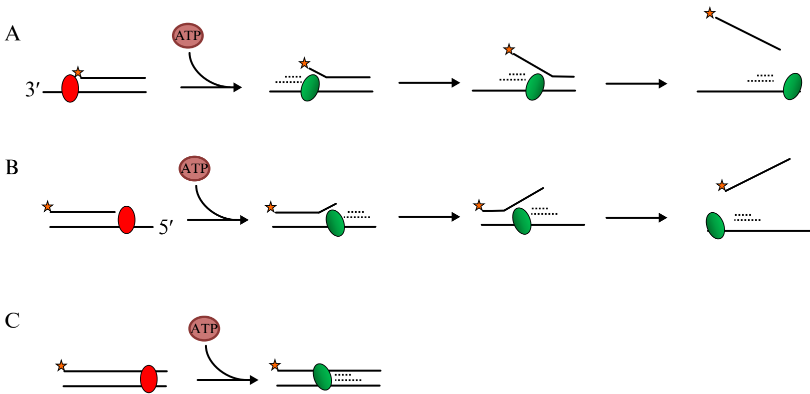

Figure 6.1 Cartoon representation of translocase activities. Depictions of diverse translocase activities are shown. These activities include (A) $3' \rightarrow 5'$ translocation on ssRNA, (B) $5' \rightarrow 3'$ translocation on ssRNA, and (C) translocation on dsRNA. In the absence of ATP, the helicases are static molecules (red ovals). Upon the addition of ATP, helicases begin to exert mechanical and dynamic movement (green ovals). Translocative behavior can lead to unwinding of RNA duplexes ((A) and (B)). However, it is worth noting that there are occasions in which translocation can occur in the absence of unwinding (C).

ssRNA translocase; Upf1, a $5' \rightarrow 3'$ ssRNA translocase; and retinoic acid inducible gene I (RIG-I), a dsRNA translocase. For each enzyme, we describe structural features of the translocative motor domains, and the contributions of additional domains that aid directional movement. We also discuss the molecular determinants for substrate recognition and translocation. These findings will be interpreted in light of the fact that most of these proteins function as components of larger macromolecular machines rather than in isolation (i.e., in the exon junction complex or within replication complexes). We highlight the examples where the presence or absence of additional cofactors alters enzymatic behavior, thereby providing insight into how the cell may use different strategies to regulate helicase activity *in vivo*.

1.1. The protein domains of an RNA translocase: The RecA motor fold and companion domains

Like their DNA-remodeling cousins, the core of RNA translocase enzymes is composed of two RecA-like domains (α/β domain; a.k.a. Rossman fold) that are linked in tandem (Fairman-Williams *et al.*, 2010; Pyle, 2011; Singleton *et al.*, 2007). ATP binds in the cleft formed between these two domains and conserved amino acid motifs responsible for ATP binding and hydrolysis line up along each RecA-like domain facing inward toward the cleft. Atop these two domains is the platform where RNA binds. Depending upon the structural features of additional domains outside the two RecA-like domains, the geometry of this platform can accommodate either ssRNA or dsRNA (Fig. 6.2).

Many helicase proteins belonging to the nonprocessive DEAD-box classification, which also hydrolyze ATP upon RNA binding, only have two RecA-like domains (Linder and Jankowsky, 2011). However, most translocative and processive enzymes have additional domains that potentiate important roles in substrate recognition and helicase function. Some of these domains play positive roles in conferring processive behavior (Domain 3 of NS3), some play autoinhibitory or regulatory roles to prevent processive behavior (Domains CH and 1B in Upf1), and some mediate complex functions such as the transmission of mechanical information over long distances (pincer domain of RIG-I) (Fig. 6.2). Further, some helicases, such as *E. coli* DbpA, use these extra domains to recognize a specific RNA tertiary structure in order to activate its ATPase activity (Tsu *et al.*, 2001), whereas other helicases lacking these additional domains, such as *S. cerevisiae* Dbp5, require the binding of auxiliary cofactors (Gle1) and small molecules (inositol hexakisphosphate) to stimulate RNA release (Montpetit *et al.*, 2011). Functional dissection of these additional domains or binding cofactors represents an active area of investigation by many laboratories.

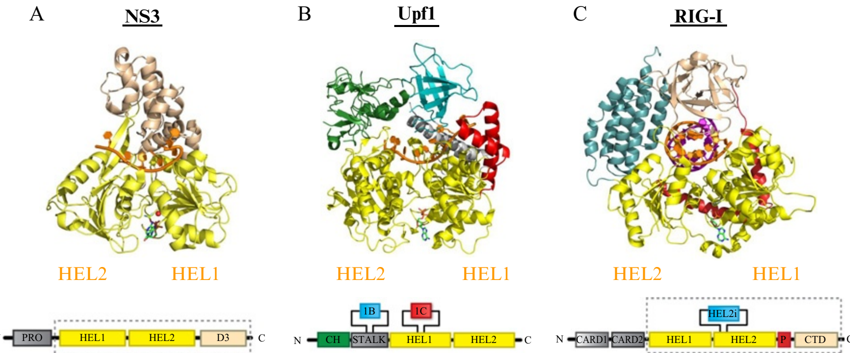

Figure 6.2 Structural attributes of translocating helicases. Structural snapshots of helicases on their respective RNA substrates are shown. These helicases are (A) HCV NS3, a $3' \rightarrow 5'$ ssRNA translocase, (B) Upf1, a $5' \rightarrow 3'$ ssRNA translocase, and (C) RIG-I, a dsRNA translocase. Boxed regions indicate the illustrated portion of the helicase. Domains outside the boxed regions were omitted for clarity. The two RecA-like domains are colored in yellow, and additional domains are colored according to the diagrams beneath each structure. ATP binds in the cleft between the two RecA-like domains. Regardless of the translocation direction, the ssRNAs (orange) lie in the same orientation atop the helicase motor: the $5'$ end rests above HEL2 and the $3'$ end rests above HEL1. This strand is called the tracking strand. For RIG-I, the tracking strand of the duplex is also colored in orange, and the activating strand is colored in purple.

2. Major Examples of Monomeric RNA Translocases

2.1. The NS3/NPH-II family of SF2 proteins: $3' \rightarrow 5'$ translocation enzymes

Most of our knowledge about these enzymes stems from pioneering work on SF2 helicases of viral origins, primarily from work on the NS3 proteins from the *Flaviviridae* viruses, Dengue Virus and HCV, and NPH-II from Vaccinia Virus. The NS3 helicases from HCV and Dengue are believed to be important during replication, where they may play a role in unraveling the highly structured single-stranded genomes of these viruses. However, HCV NS3 is now known to play other roles in the life of the virus, including a major function in virus particle packaging and assembly (Jones *et al.*, 2011; Ma *et al.*, 2008). Thus, despite the apparent helicase activity of NS3 enzymes, their actual roles in the viral lifecycles are not firmly established. Similarly, the NPH-II helicases of pox viruses are thought to be required for efficient transcription and inhibition of R-loop formation (Gross and Shuman, 1996). However, there is increasing evidence that they may play a role in release of the polymerase.

The robust unwinding activities of NS3 and NPH-II enzymes have helped reveal the specific structural and chemical features of RNA that are important for translocation and unwinding. Through a combination of ensemble studies in which unwinding was used as a proxy for translocation, together with direct single-molecule studies of translocation, it has been established that these enzymes specifically engage backbone residues on single-stranded regions of RNA (Beran *et al.*, 2006; Dumont *et al.*, 2006; Kawaoka *et al.*, 2004). ATP hydrolysis then promotes directional movement along this bound strand with concomitant displacement of the annealed strands. Indeed, robust helicases like NPH-II strip away annealed strands and even bound proteins with little to no regard for their chemical composition (Jankowsky *et al.*, 2000, 2001). Importantly, translocation by these enzymes requires physical continuity of phosphodiester linkages on this tracking strand, and a basic site along the tracking strand are tolerated, thereby indicating that direct interactions with nucleobases are not essential for unwinding activity (Beran *et al.*, 2006; Kawaoka *et al.*, 2004).

HCV NS3 is a bifunctional enzyme that contains a helicase domain, and an appended serine protease domain that is important during early stages of viral protein processing (Raney *et al.*, 2010). Although the protease and helicase domains are not known to function during the same stages of the viral lifecycle, they strongly influence each other's functions. That is, the serine protease domain is essential for conferring strong binding to the helicase domain, and the helicase domain enhances proteolysis by the protease domain (Beran and Pyle, 2008; Beran *et al.*, 2007). Thus, as in many helicases, NS3 has evolved a functional dependence on appended domains. The helicase domain of HCV NS3, as well as those for all other *Flaviviridae* NS3 proteins, is a three-lobed structure; two lobes are the RecA-like domains (HEL1 and HEL2) and the third lobe (Domain 3) is an all α-helical structure (Luo *et al.*, 2008a; Yao *et al.*, 1997; Fig. 6.2A). ssRNA binds in the cavity formed between HEL1 and HEL2 on the bottom and Domain 3 on the top (Appleby *et al.*, 2011; Luo *et al.*, 2008b). The geometry of the RNA binding cavity is such that it can only accommodate ssRNA, a finding that is consistent with direct binding studies demonstrating that NS3 binds tightly to single-stranded substrates and weakly to blunt, duplex substrates (Tai *et al.*, 1996).

In the context of a model unwinding substrate, which usually consists of a "tailed duplex," NS3 binds the single-stranded region, which serves as a launch pad for ATP-dependent tracking through the duplex portion of the substrate. NS3 binds ssRNA in an orientation whereby the phosphodiester backbone faces HEL1 and HEL2, and the nucleobases are displayed in the opposite direction toward Domain 3 (Fig. 6.2A). Enzymological studies and crystallographic investigations show that NS3 tracks along the phosphodiester backbone by forming two types of interactions: hydrogen bonds mediated by main-chain amide groups and contacts formed through polar

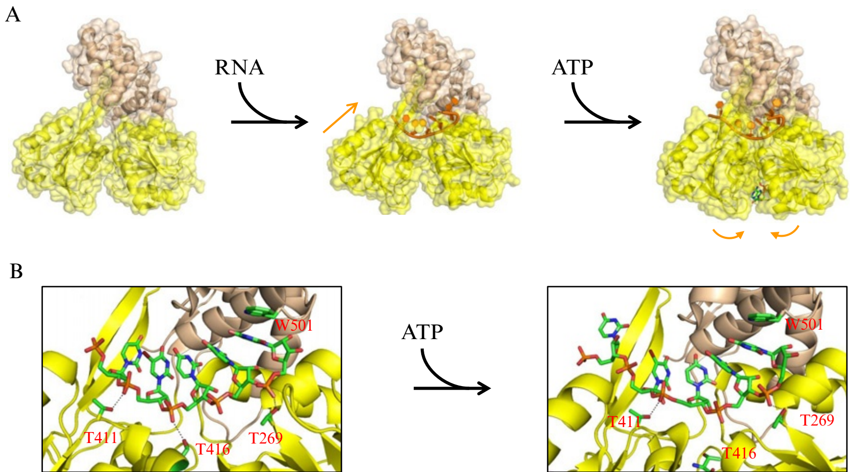

Figure 6.3 Dynamic motions of HCV NS3, a $3' \rightarrow 5'$ ssRNA translocase. (A) Domain motions of HCV NS3 in response to RNA and ATP binding. Upon binding ssRNA, HEL2 undergoes a small conformational rearrangement that allows for optimal positioning. In the presence of ATP, HEL1, and HEL2 move closer in proximity to each other to enclose the cleft. Domain motions were interpreted based upon aligning all structures to HEL1. (B) In the absence of ssRNA, the side chains of three threonine residues hydrogen bonds form direct contacts with the phosphodiester backbone of the ssRNA, whereas a sole tryptophan residue makes a stacking interaction with a nucleobase. Note that Thr269 and Thr411 are positioned three nucleotides apart. Upon binding of ATP and closure of the two RecA-like domains, Thr416 disengages from the backbone, and Thr269 and Thr411 are now spaced by only one nucleotide. Trp501 remains stacked with the nucleobase.

side chains (Appleby *et al.*, 2011; Beran *et al.*, 2006). Examples of interactions between the RNA backbone and main-chain amide groups include Val232 and Gly255 that extend from HEL1, and Lys371 and Arg393 that extend from HEL2. The use of polar side chains to form contacts is best illustrated through two evolutionarily conserved threonine residues (Thr269 from HEL1 and Thr411 from HEL2). For SF2 helicases, these two residues project from each RecA-like domain and act as dynamic pincers by forming hydrogen bonds with the nonbridging phosphoryl oxygen atoms (Fig. 6.3B). An additional hydrogen bond is established with the RNA backbone using the side chain of Thr416, a residue within motif V. However, this residue only engages the ssRNA backbone in the absence of ATP and disengages from the ssRNA in the presence of ATP. In the absence of ATP, Thr269 and Thr411 are spaced three nucleotides apart, allowing the protein to adopt a more open conformation. ATP binding induces closure of HEL1 and HEL2 by bringing them together in closer proximity such that the two threonine residues are now separated by two

nucleotides (Fig. 6.3A). Upon ATP hydrolysis, the products ADP and inorganic phosphate dissociate from the ATP binding cavity, and HEL2 relaxes to move away from HEL1, thus effectively advancing the helicase by one nucleotide toward the 5′ direction of the ssRNA.

Cycles of ATP binding and hydrolysis directly influence the number of hydrogen bonds formed between the helicase and RNA substrate, thus resulting in alternating high- and low-affinity states. Most interactions are observed in the ATP-free state, which represents the high-affinity state. Upon binding ATP, the number of contacts between helicase and the RNA decreases (notably, Thr416 disengages from the RNA backbone) and resulting in a low-affinity state, which is consistent with previous direct binding studies demonstrating the helicase has a lower affinity for RNA in the presence of nucleotide.

The 3′→5′ directionality of movement is achieved in part through conserved hydrophobic residue Trp501, which extends from Domain 3 (Fig. 6.3B). This residue stacks upon the nucleobase at the 3′ end of the ssRNA and effectively serves as an anchor point for the helicase to move unidirectionally toward the 5′ end of the ssRNA rather than slipping backward. Interestingly, through the cycles of ATP binding and hydrolysis, the orientation of Trp501 remains stacked with the nucleobase (Appleby *et al.*, 2011). Maintaining this interaction with the nucleobase is functionally significant for the translocation and possibly for unwinding mechanisms of NS3. Indeed, single-molecule experiments have suggested that Trp501 acts as a plowshare, dragging behind the translocating helicase (which moves along the tracking strand in one nucleotide increments as described above) (Myong *et al.*, 2007). Specifically, the data supports a model where Trp501 maintains its interactions with the same nucleobase, while the helicase undergoes two forward cycles of ATP binding and hydrolysis; during the third cycle, the helicase accumulates enough tension in the system and causes Trp501 to spring forward along the tracking strand by three nucleotides concomitant unwinding of three base pairs. Importantly, no sequence–specific interactions are formed between the nucleobases on the ssRNA strand and residues from Domain 3. This lack of specificity likely contributes to the ability of the helicase to unwind substrates in a sequence-independent fashion.

Currently, there are no crystal structures available of *bona fide* helicases in complex with duplex RNA substrates (there are structures of RIG-I with RNA duplexes, but it is not a helicase, *vide infra*). However, unwinding is likely achieved by the steric pressure from a β-hairpin that projects from HEL2. This conserved structural element is present in other related helicases and is proposed to function like a wedge to splay the two composite strands apart (Buttner *et al.*, 2007; Luo *et al.*, 2008b). While the tracking strand is fed through the ssRNA channel described above, the displaced strand is directed away. However, the trajectory of the displaced strand is currently unknown and represents a potential area of research.

2.2. The UPf1 family of SF1 proteins: $5' \rightarrow 3'$ translocation enzymes

Much less is known about the translocation and unwinding mechanisms of RNA helicases that move with 5′–3′ directionality. This is unfortunate because members of this class include helicases of viral origin such as SARS coronavirus, which had been shown to be unusually processive. The SARS helicase can efficiently unwind RNA duplexes of several hundred base pairs under single-cycle conditions (Ivanov and Ziebuhr, 2004). However, the most structurally and functionally characterized $5' \rightarrow 3'$ translocation enzymes of the SF1 protein Upf1. Upf1 is a core component of the nonsense-mediated decay (NMD) machinery that detects and rapidly degrades aberrant mRNA transcripts (Conti and Izaurralde, 2005). Intriguingly, Upf1 is phylogenetically related to well-characterized DNA helicases such as Rep and PcrA, rather than to other SF2 RNA helicases.

Upf1 exhibits low ATPase and unwinding activities in isolation and requires the binding of additional NMD factors Upf2 and Upf3 to stimulate these activities (Chamieh *et al.*, 2008). The structural basis for these functional observations was recently illustrated in a series of crystal structures of Upf1 in isolation and with various cofactors that alter its enzymatic activities (Chakrabarti *et al.*, 2011; Cheng *et al.*, 2007; Clerici *et al.*, 2009). Like the NS3 helicase, Upf1 contains a core motor domain composed of two RecA-like domains (HEL1 and HEL2) (Fig. 6.2B). The N-terminus of Upf1 contains a zinc-knuckle domain (CH domain) that exerts an inhibitory effect *in cis* on the enzymatic activities of the protein. Between the CH domain and HEL1 is a domain composed of six antiparallel β-strands (Domain 1B), and a stalk domain composed of two α-helices. Projecting from HEL1 is an additional domain also composed of two α-helices (Domain 1C). Such modular domains embedded within the RecA-like domains are unique features that typify members belonging to SF1. Interestingly, ssRNA binds in an orientation atop the HEL1 and HEL2 platform of Upf1 similar to that seen for NS3. This is a surprising finding, given that one would expect that a helicase that translocates and unwinds with an opposite polarity as NS3 to position itself on the RNA in a reversed orientation. However, this similarity in binding polarity, despite opposite translocation directionality, has also been observed in comparisons of the structures of DNA helicases, RecD2 and PcrA (Saikrishnan *et al.*, 2009; Velankar *et al.*, 1999), and in the hexameric helicases, Rho and E1 (Enemark and Joshua-Tor, 2006; Thomsen and Berger, 2009). Results with Upf1 are therefore consistent with the notion that translocases do not "turn around" to run backward. Rather, they reverse their gears (Pyle, 2009).

At a molecular level, Upf1 contacts the ssRNA using main-chain amide groups and side-chain interactions that project from HEL1, HEL2, the stalk,

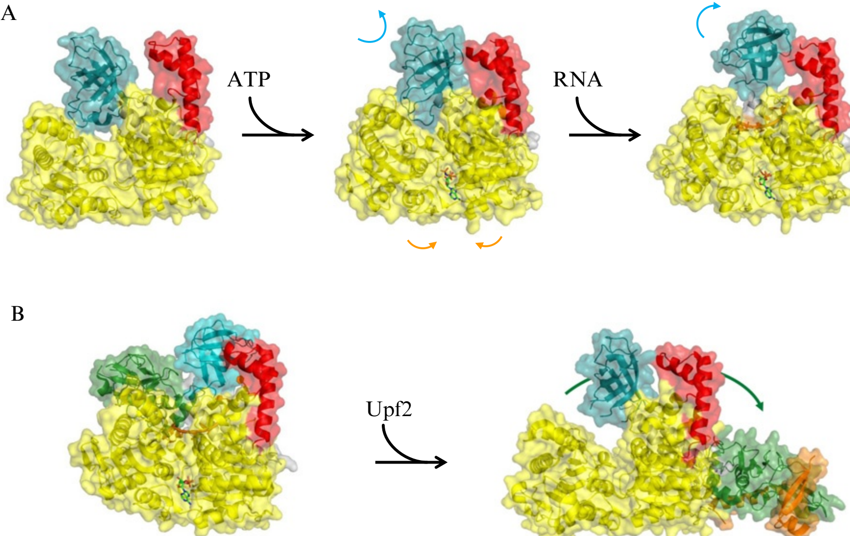

Figure 6.4 Dynamic motions of Upf1, a $5' \rightarrow 3'$ ssRNA translocase. (A) Similar to HCV NS3, the two RecA-like domains of Upf1 close upon binding ATP. Domain 1B undergoes a small conformational change and rotates away from the interface of the two RecA-like domains. Upon binding ssRNA, Domain1B makes another conformational change to position itself away from the 3′ end of the nucleic acid. This rearrangement allows Upf1 to translocate processively with $5' \rightarrow 3'$ directionality. Domain motions were interpreted based upon aligning all structures to HEL1. (B) In the presence of the CH domain (green), Upf1 adopts a compact, globular conformation that prevents processive translocation activity. Inhibition can be partially relieved by the binding of the cofactor Upf2 (orange). Binding leads to a large conformational rearrangement of the CH domain. Whereas the CH domain was once positioned above HEL2, binding of Upf2 repositions the CH domain behind HEL1.

Domain 1B, and Domain 1C. The CH domain does not directly interact with the RNA but rather exerts an allosteric effect on RNA binding by impinging upon Domain 1B to contact two additional nucleotides at the 3′ end of the RNA (Fig. 6.4B). Impinging upon Domain 1B likely prevents forward translocative movement and may account for the intrinsically low ATPase and unwinding activities of the full-length protein. Interestingly, the crystal structure of a Upf1 construct lacking its CH domain shows that Domain 1B rotates away from the 3′ end of the RNA such that Domain 1B is no longer locked in a restricted position (Fig. 6.4A). By allowing Domain 1B to disengage from the RNA, Upf1 displays higher ATPase and unwinding activities. Therefore, the presence or absence of Domain 1B interacting with the 3′ end of the RNA may represent an autoinhibition mechanism occurring *in cis* to modulate the helicase activities of Upf1.

In a cellular context, preventing the CH domain from impinging on Domain 1B is accomplished by the addition of Upf2 (Fig. 6.4B). The C-terminus of Upf2 binds directly to the CH domain of Upf1, and this interaction is sufficient to enhance the ATPase and unwinding activities of Upf1. From a structural viewpoint, the binding of Upf2 to Upf1 induces a large conformational change in the CH domain of Upf1. Whereas the CH domain was originally positioned atop the HEL2 domain, binding of Upf2 causes the CH domain to occupy a new position distal to HEL1. By occupying this new position, the CH domain can no longer impinge upon Domain 1B to interact with the 3′ end of the ssRNA. The repositioning of the CH domain is likely the structural basis for the functional observation that Upf2 enhances the biochemical properties of Upf1.

For future studies, it will be interesting to understand the coordinated domain motions of Upf1 in the presence of full-length Upf2 and Upf3 proteins or with an RNA unwinding substrate. Having such a global view may help uncover the biological significance of translocation and unwinding by Upf1 in the context of the NMD pathway. Further, it would be valuable to conduct experiments similar to those performed using NPH-II and NS3 proteins probing mechanistic features of substrate recognition. These studies would provide a basis to compare the similarities and differences between helicases that operate on the similar nucleic acid substrates but move in different directions.

2.3. Double-stranded translocation by monomeric RNA helicases: The RIG-I/Dicer family

While members of this class have been characterized biochemically, a structural view of any member of this class had long been elusive until recently. This was especially unfortunate given that members constituting this class of enzymes are involved in important functions such as innate immune signaling (RIG-I, and related proteins MDA5 and LGP2) and RNA interference (Dicer) (Bernstein *et al.*, 2001; Yoneyama *et al.*, 2004, 2005). In the absence of a crystal structure, it was difficult to understand how a helicase recognizes dsRNA, which remained a lingering question in the field. For example, does the dsRNA rest on top of the two RecA-like domains in an orientation similar to those seen for helicases that bind ssRNA? If so, is the binding groove sufficiently deep to discriminate dsRNA from ssRNA? For recognizing and binding a dsRNA substrate, is one strand preferred over the other strand? These questions were especially important to answer given that the RIG-I and Dicer proteins are phylogenetically related SF2 helicases, but function in strikingly different pathways: RIG-I translocates along dsRNA and functions in innate immune signaling, whereas Dicer binds and cleaves long double-stranded pre-microRNAs into short, double-stranded siRNAs. Though both proteins

are powered by the same helicase core motor, their divergent functions are conferred by accessory domains flanking the N- and C-termini of the respective protein. Fortunately, the structural basis for recognizing and binding dsRNA by RIG-I in the presence and absence of bound nucleotide was recently shown in a series of publications from numerous groups (Civril *et al.*, 2011; Jiang *et al.*, 2011; Kowalinski *et al.*, 2011; Luo *et al.*, 2011).

RIG-I recognizes and binds dsRNA using its central SF2 helicase motor core and displays dsRNA-dependent ATPase activity (Gack *et al.*, 2008). At the N-terminus, RIG-I contains two tandem *c*aspase *a*ctivation and *r*ecruitment *d*omains (CARDs; CARD1 and CARD2) for signaling, and at the C-terminus, it contains a domain important for recognizing 5′-triphosphate moieties (the CTD) (Cui *et al.*, 2008) (Fig. 6.2C). The CARD fold is typically found in proteins that function in apoptosis and inflammatory signaling pathways (Hofmann *et al.*, 1997). Therefore, it was highly unusual for a helicase to have such an accessory domain that functions outside the boundaries of the Central Dogma. In the context of RIG-I, it was later determined that the CARDs forms a platform to interact with other CARD-containing proteins to propagate a signaling cascade (Gack *et al.*, 2008). The CTD recognizes the 5′ end of either ssRNA or dsRNA, though it binds with the tightest affinity to duplexed substrates bearing a triphosphate group (Wang *et al.*, 2010).

The recognition of triphosphorylated substrates is a defining feature of RIG-I (Hornung *et al.*, 2006; Pichlmair *et al.*, 2006). As most cellular RNAs are capped or modified at their 5′ ends immediately following transcription, the chemical nature of free 5′ triphosphate groups are immediately recognized by RIG-I as "nonself" RNA. It is generally accepted that such "nonself" RNAs are generated when viruses hijack the cellular machinery to create new strands of viral genomic RNA *de novo* using its viral polymerase. This recognition event activates RIG-I to initiate a signaling cascade through the CARDs to ward off the invading virus to maintain an antiviral cellular state.

Translocation along RNA duplex by the helicase domain was demonstrated using a single-molecule fluorescence approach (Myong *et al.*, 2009). Using a wide spectrum of protein and RNA constructs, the authors demonstrated that RIG-I exhibits robust and repetitive translocation along duplex RNA substrates. Notably, translocation activity occurred in the absence of unwinding. This was consistent with previous studies showing that neither RIG-I, nor its cousin MDA5 display strong helicase activity. Both of these enzymes display ATPase activity that is specifically activated by duplex RNA rather than single strands (Gee *et al.*, 2008; Kang *et al.*, 2002). Functional dissection of the role for the CARD domains indicates that they play an inhibitory role by negatively regulating translocation activity (Myong *et al.*, 2009) (Fig. 6.5B). Binding of duplex RNA partially alleviates this inhibitory role to allow translocation. Interestingly, maximal translocation

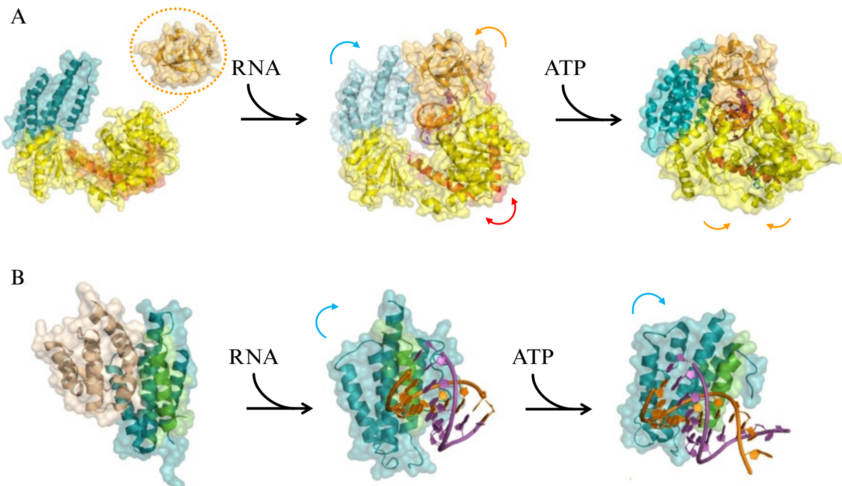

Figure 6.5 Dynamic motions of RIG-I, a dsRNA translocase. (A) As determined by SAXS measurements, RIG-I adopts an open conformation in the absence of dsRNA substrate. For clarity, we omitted the CARDs from the first structure. We model in a proposed position of the CTD in this figure (dotted circle) to account for the open conformation needed to allow dsRNA binding. HEL2i and the CTD surround the dsRNA substrate upon binding, which is likely facilitated by the V-shaped pincer domain. Additionally, the RecA-like domains close upon binding ATP, and HEL2 establishes additional contacts with the dsRNA backbone. Domain motions were interpreted based upon aligning all structures to HEL1 and dsRNA. (B) Major domain motions by HEL2i in response to dsRNA and ATP. In the *apo* conformation, the CARDs form a large hydrophobic interface with HEL2i and prevent HEL2i from interacting with the dsRNA. Likewise, the presence of the CARDs is inhibitory for processive translocative behavior. Upon binding dsRNA, HEL2i establishes a single, weak interaction with the backbone of the substrate using the face of a specific α-helix (green). Upon binding ATP, HEL2i moves further toward the dsRNA and establishes more contacts using this α-helix. This alternation between strong and weak interactions in response to ATP is likely the structural basis for dsRNA translocation.

activity was observed using triphosphorylated duplex RNA, thereby highlighting the specific structural and chemical features of the nucleic acid substrate requirement for RIG-I. Taken together, RIG-I uses its helicase domain to power translocation along dsRNA, potentially disrupting the viral replication machinery or assisting in the loading of additional RIG-I molecules.

From the collection of available recent RIG-I structures, we can now directly observe the coordinated motions of its multiple domains, which has led to new and unexpected structural findings. In addition to the CARDs, the canonical motor domains HEL1 and HEL2, and the CTD, RIG-I contains two domains that had never been appreciated previously. The

first domain is composed of five antiparallel α-helices and is termed as the insertion domain (HEL2i). HEL2i exists as an independent structure projecting above from HEL2 and forms direct contacts with the backbone dsRNA. The second domain is composed of two α-helices that form a V-shape, which we refer to as the pincer domain, which wraps around multiple domains in order to transduce mechanical signals throughout the protein.

In the absence of dsRNA, RIG-I adopts an open, extended conformation. CARD1 and CARD2 stack upon one another, and CARD2 forms extensive interactions with HEL2i that are stabilized by salt bridges and hydrophobic contacts (Kowalinski *et al.*, 2011) (Fig. 6.5B). The interactions between CARD2 and HEL2i may partially explain the inhibitory role of the CARDs observed in the single-molecule fluorescence studies. In the presence of dsRNA, however, RIG-I collapses upon and encircles the dsRNA (Jiang *et al.*, 2011; Luo *et al.*, 2011). This compaction event is consistent with small-angle X-ray scattering (SAXS) data measuring the dimensions of these molecules in solution in the presence and absence of dsRNA. Though a structure of the full-length RIG-I protein with dsRNA is unavailable, SAXS measurements indicate that such a complex would place the CARDs away in solution and distal to the dsRNA binding interface. In fact, such a conformation might even be optimal for the CARDs to interact with other protein cofactors.

When binding an RNA duplex, it is important to distinguish one strand from the other. We refer to the RNA strand that lies across the HEL1 and HEL2 motor domains as the tracking strand (Fig. 6.2C). Interestingly, the tracking strand is in the same polarity to those seen for NS3 and Upf1, whereby the 5′ end of the tracking strand is located above HEL2, while the 3′ end is located above HEL1. We refer to the complementary sister strand as the activating strand since the 5′ end of this strand directly interacts with the CTD and would carry the triphosphate group that activates RIG-I translocation activity. The duplex RNA is oriented perpendicular to the plane of HEL1 and HEL2, and numerous polar interactions are formed between the 2′-hydroxyl groups of the ribose sugars and the phosphodiester backbone using main-chain amides and side-chain residues; hydrophobic interactions are only observed between the CTD and the 5′ end of the activating strand. The lack of base specificity likely allows RIG-I to translocate on diverse dsRNAs without sequence bias.

The bound nucleotide state directly influences the number of interactions made between RIG-I and the dsRNA. When ATP occupies the cleft formed by HEL1 and HEL2, these two domains move in close proximity to one another in order to prime the catalytic residues lining the cleft for ATP binding and hydrolysis. In this closed conformation, HEL1 and HEL2 form contacts with both the tracking and activating strands. However, in the absence of nucleotide, RIG-I adopts a more open conformation where

HEL2 moves away from HEL1 and disengages from contacting the dsRNA. Intriguingly, HEL1 maintains its grip with the dsRNA regardless of the nucleotide bound state. Similar domain movements are also observed with HEL2i. Using the face of one of its helices, HEL2i forms extensive contacts with both strands of the dsRNA in the presence of nucleotide; these contacts are lost, except one, in the absence of bound nucleotide (Fig. 6.5B).

The act of engaging and then disengaging from the dsRNA by HEL2 and HEL2i in response to the bound nucleotide state may directly reflect the RIG-I translocation mechanism. By maintaining a constant point of contact with the dsRNA through HEL1, RIG-I may use HEL2 and HEL2i to inch the protein forward along one strand of the dsRNA. Upon nucleotide binding, RIG-I maintains a tight grip on the dsRNA by forming direct contacts through HEL1, HEL2, and HEL2i. Following, ATP hydrolysis may permit RIG-I to loosen its grip on the dsRNA by allowing HEL2 and HEL2i to disengage from the dsRNA. Each cycle of ATP binding and hydrolysis would then propel RIG-I forward by one base-pair toward the 5′ end of the tracking strand. Indeed, such directional bias along the tracking strand had been observed in a single-molecule setting. In a biological context, it is tempting to speculate that directional translocation along the tracking strand may serve to dislodge the viral replication machinery that would be actively synthesizing new viral RNA (i.e., the activating strand). However, whether RIG-I can remodel protein–RNA complexes, as had been shown for other helicases previously, remains the subject of future inquiry.

Despite having the numerous structural views of RIG-I now available, several outstanding questions remain. For example, given that the CTD has a tight affinity for the triphosphorylated duplex substrates, how does RIG-I disengage from that position to translocate along dsRNA? What are the roles of phosphorylation, ubiquitination, or tetraubiquitin binding in altering the biochemical activities of RIG-I? How do the CARDs interact with other proteins that have CARD domains, and what are the functional outcomes? Is there a role for ATP hydrolysis during the presentation of the CARD domains and in signaling?

3. Concluding Remarks

We now have complete structural views of three RNA translocases that move in different directions and function on different types of RNA. Despite the diversity in the directionalities and substrate specificities of these molecular motors, these enzymes share a common theme: they bind and hydrolyze ATP in the presence of a stimulating nucleic acid substrate. While a main research question is to decipher what ATP binding and hydrolysis

accomplish, an emerging theme appears to be that this orchestrated motion advances the protein forward by one nucleotide. It is thus tempting to speculate that perhaps any helicase containing these two RecA-like modules may be able to move along nucleic acid strands, regardless of the number of additional flanking domains. Nevertheless, there is still much to uncover what is underneath the hoods of Nature's molecular engines, and future research will undoubtedly shed light on their diverse mechanical properties.

ACKNOWLEDGMENTS

This work was funded by the Howard Hughes Medical Institute and NIH grant AI089826. A. M. P. is an Investigator with the Howard Hughes Medical Institute.

REFERENCES

Appleby, T. C., Anderson, R., Fedorova, O., Pyle, A. M., Wang, R., Liu, X., Brendza, K. M., and Somoza, J. R. (2011). Visualizing ATP-dependent RNA translocation by the NS3 helicase from HCV. *J. Mol. Biol.* **405,** 1139–1153.

Beran, R. K., and Pyle, A. M. (2008). Hepatitis C viral NS3-4A protease activity is enhanced by the NS3 helicase. *J. Biol. Chem.* **283,** 29929–29937.

Beran, R. K., Bruno, M. M., Bowers, H. A., Jankowsky, E., and Pyle, A. M. (2006). Robust translocation along a molecular monorail: the NS3 helicase from hepatitis C virus traverses unusually large disruptions in its track. *J. Mol. Biol.* **358,** 974–982.

Beran, R. K., Serebrov, V., and Pyle, A. M. (2007). The serine protease domain of hepatitis C viral NS3 activates RNA helicase activity by promoting the binding of RNA substrate. *J. Biol. Chem.* **282,** 34913–34920.

Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409,** 363–366.

Buttner, K., Nehring, S., and Hopfner, K. P. (2007). Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nat. Struct. Mol. Biol.* **14,** 647–652.

Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., and Conti, E. (2011). Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Mol. Cell* **41,** 693–703.

Chamieh, H., Ballut, L., Bonneau, F., and Le Hir, H. (2008). NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. *Nat. Struct. Mol. Biol.* **15,** 85–93.

Cheng, Z., Muhlrad, D., Lim, M. K., Parker, R., and Song, H. (2007). Structural and functional insights into the human Upf1 helicase core. *EMBO J.* **26,** 253–264.

Civril, F., Bennett, M., Moldt, M., Deimling, T., Witte, G., Schiesser, S., Carell, T., and Hopfner, K. P. (2011). The RIG-I ATPase domain structure reveals insights into ATP-dependent antiviral signalling. *EMBO Rep.* **12,** 1127–1134.

Clerici, M., Mourao, A., Gutsche, I., Gehring, N. H., Hentze, M. W., Kulozik, A., Kadlec, J., Sattler, M., and Cusack, S. (2009). Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2. *EMBO J.* **28,** 2293–2306.

Conti, E., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Curr. Opin. Cell Biol.* **17,** 316–325.

Cui, S., Eisenacher, K., Kirchhofer, A., Brzozka, K., Lammens, A., Lammens, K., Fujita, T., Conzelmann, K. K., Krug, A., and Hopfner, K. P. (2008). The C-terminal regulatory domain is the RNA 5′-triphosphate sensor of RIG-I. *Mol. Cell* **29,** 169–179.

Dumont, S., Cheng, W., Serebrov, V., Beran, R. K., Tinoco, I., Jr., Pyle, A. M., and Bustamante, C. (2006). RNA translocation and unwinding mechanism of HCV NS3 helicase and its coordination by ATP. *Nature* **439,** 105–108.

Enemark, E. J., and Joshua-Tor, L. (2006). Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* **442,** 270–275.

Fairman-Williams, M. E., Guenther, U. P., and Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.* **20,** 313–324.

Gack, M. U., Kirchhofer, A., Shin, Y. C., Inn, K. S., Liang, C., Cui, S., Myong, S., Ha, T., Hopfner, K. P., and Jung, J. U. (2008). Roles of RIG-I N-terminal tandem CARD and splice variant in TRIM25-mediated antiviral signal transduction. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 16743–16748.

Gee, P., Chua, P. K., Gevorkyan, J., Klumpp, K., Najera, I., Swinney, D. C., and Deval, J. (2008). Essential role of the N-terminal domain in the regulation of RIG-I ATPase activity. *J. Biol. Chem.* **283,** 9488–9496.

Gross, C. H., and Shuman, S. (1996). Vaccinia virions lacking the RNA helicase nucleoside triphosphate phosphohydrolase II are defective in early transcription. *J. Virol.* **70,** 8549–8557.

Hofmann, K., Bucher, P., and Tschopp, J. (1997). The CARD domain: a new apoptotic signalling motif. *Trends Biochem. Sci.* **22,** 155–156.

Hornung, V., Ellegast, J., Kim, S., Brzozka, K., Jung, A., Kato, H., Poeck, H., Akira, S., Conzelmann, K. K., Schlee, M., *et al.* (2006). 5′-Triphosphate RNA is the ligand for RIG-I. *Science* **314,** 994–997.

Ivanov, K. A., and Ziebuhr, J. (2004). Human coronavirus 229E nonstructural protein 13: Characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5′-triphosphatase activities. *J. Virol.* **78,** 7833–7838.

Jankowsky, E., Gross, C. H., Shuman, S., and Pyle, A. M. (2000). The DExH protein NPH-II is a processive and directional motor for unwinding RNA. *Nature* **403,** 447–451.

Jankowsky, E., Gross, C. H., Shuman, S., and Pyle, A. M. (2001). Active disruption of an RNA-protein interaction by a DExH/D RNA helicase. *Science* **291,** 121–125.

Jiang, F., Ramanathan, A., Miller, M. T., Tang, G. Q., Gale, M., Jr., Patel, S. S., and Marcotrigiano, J. (2011). Structural basis of RNA recognition and activation by innate immune receptor RIG-I. *Nature* **479,** 423–427.

Jones, D. M., Atoom, A. M., Zhang, X., Kottilil, S., and Russell, R. S. (2011). A genetic interaction between the core and NS3 proteins of hepatitis C virus is essential for production of infectious virus. *J. Virol.* **85,** 12351–12361.

Kang, D. C., Gopalkrishnan, R. V., Wu, Q., Jankowsky, E., Pyle, A. M., and Fisher, P. B. (2002). mda-5: An interferon-inducible putative RNA helicase with double-stranded RNA-dependent ATPase activity and melanoma growth-suppressive properties. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 637–642.

Kawaoka, J., Jankowsky, E., and Pyle, A. M. (2004). Backbone tracking by the SF2 helicase NPH-II. *Nat. Struct. Mol. Biol.* **11,** 526–530.

Kowalinski, E., Lunardi, T., McCarthy, A. A., Louber, J., Brunel, J., Grigorov, B., Gerlier, D., and Cusack, S. (2011). Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA. *Cell* **147,** 423–435.

Linder, P., and Jankowsky, E. (2011). From unwinding to clamping—The DEAD box RNA helicase family. *Nat. Rev. Mol. Cell Biol.* **12,** 505–516.

Luo, D., Xu, T., Hunke, C., Gruber, G., Vasudevan, S. G., and Lescar, J. (2008a). Crystal structure of the NS3 protease-helicase from dengue virus. *J. Virol.* **82,** 173–183.

Luo, D., Xu, T., Watson, R. P., Scherer-Becker, D., Sampath, A., Jahnke, W., Yeong, S. S., Wang, C. H., Lim, S. P., Strongin, A., *et al.* (2008b). Insights into RNA unwinding and ATP hydrolysis by the flavivirus NS3 protein. *EMBO J.* **27,** 3209–3219.

Luo, D., Ding, S. C., Vela, A., Kohlway, A., Lindenbach, B. D., and Pyle, A. M. (2011). Structural insights into RNA recognition by RIG-I. *Cell* **147,** 409–422.

Ma, Y., Yates, J., Liang, Y., Lemon, S. M., and Yi, M. (2008). NS3 helicase domains involved in infectious intracellular hepatitis C virus particle assembly. *J. Virol.* **82,** 7624–7639.

Montpetit, B., Thomsen, N. D., Helmke, K. J., Seeliger, M. A., Berger, J. M., and Weis, K. (2011). A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. *Nature* **472,** 238–242.

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.

Myong, S., Cui, S., Cornish, P. V., Kirchhofer, A., Gack, M. U., Jung, J. U., Hopfner, K. P., and Ha, T. (2009). Cytosolic viral sensor RIG-I is a 5′-triphosphate-dependent translocase on double-stranded RNA. *Science* **323,** 1070–1074.

Pichlmair, A., Schulz, O., Tan, C. P., Naslund, T. I., Liljestrom, P., Weber, F., and Reis e Sousa, C. (2006). RIG-I-mediated antiviral responses to single-stranded RNA bearing 5′-phosphates. *Science* **314,** 997–1001.

Pyle, A. M. (2009). How to drive your helicase in a straight line. *Cell* **139,** 458–459.

Pyle, A. M. (2011). RNA helicases and remodeling proteins. *Curr. Opin. Chem. Biol.* **15,** 636–642.

Raney, K. D., Sharma, S. D., Moustafa, I. M., and Cameron, C. E. (2010). Hepatitis C virus non-structural protein 3 (HCV NS3): A multifunctional antiviral target. *J. Biol. Chem.* **285,** 22725–22731.

Saikrishnan, K., Powell, B., Cook, N. J., Webb, M. R., and Wigley, D. B. (2009). Mechanistic basis of 5′–3′ translocation in SF1B helicases. *Cell* **137,** 849–859.

Singleton, M. R., Dillingham, M. S., and Wigley, D. B. (2007). Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* **76,** 23–50.

Tai, C. L., Chi, W. K., Chen, D. S., and Hwang, L. H. (1996). The helicase activity associated with hepatitis C virus nonstructural protein 3 (NS3). *J. Virol.* **70,** 8477–8484.

Thomsen, N. D., and Berger, J. M. (2009). Running in reverse: the structural basis for translocation polarity in hexameric helicases. *Cell* **139,** 523–534.

Tsu, C. A., Kossen, K., and Uhlenbeck, O. C. (2001). The Escherichia coli DEAD protein DbpA recognizes a small RNA hairpin in 23S rRNA. *RNA* **7,** 702–709.

Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S., and Wigley, D. B. (1999). Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* **97,** 75–84.

Wang, Y., Ludwig, J., Schuberth, C., Goldeck, M., Schlee, M., Li, H., Juranek, S., Sheng, G., Micura, R., Tuschl, T., *et al.* (2010). Structural and functional insights into 5′-ppp RNA pattern recognition by the innate immune receptor RIG-I. *Nat. Struct. Mol. Biol.* **17,** 781–787.

Yao, N., Hesson, T., Cable, M., Hong, Z., Kwong, A. D., Le, H. V., and Weber, P. C. (1997). Structure of the hepatitis C virus RNA helicase domain. *Nat. Struct. Biol.* **4,** 463–467.

Yoneyama, M., Kikuchi, M., Natsukawa, T., Shinobu, N., Imaizumi, T., Miyagishi, M., Taira, K., Akira, S., and Fujita, T. (2004). The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat. Immunol.* **5,** 730–737.

Yoneyama, M., Kikuchi, M., Matsumoto, K., Imaizumi, T., Miyagishi, M., Taira, K., Foy, E., Loo, Y. M., Gale, M., Jr., Akira, S., *et al.* (2005). Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity. *J. Immunol.* **175,** 2851–2858.