Published in final edited form as:

J Mol Biol. 2011 February 4; 405(5): 1139–1153. doi:10.1016/j.jmb.2010.11.034.

Visualizing ATP-dependent RNA Translocation by the NS3 Helicase from HCV

Todd C. Appleby^{1,*}, Robert Anderson¹, Olga Fedorova², Anna M. Pyle², Ruth Wang¹, Xiaohong Liu¹, Katherine M. Brendza¹, and John R. Somoza¹

¹Department of Structural Chemistry, Gilead Sciences, Inc., 333 Lakeside Drive, Foster City, California 94404, USA

²Department of Molecular Biophysics and Biochemistry, Howard Hughes Medical Institute and Yale University, Bass Center, Rm. 334A, 266 Whitney Ave., New Haven, Connecticut 06520, USA

Abstract

The structural mechanism by which non-structural protein 3 (NS3) from the hepatitis C virus (HCV) translocates along RNA is currently unknown. HCV NS3 is an ATP-dependent motor protein essential for viral replication and a member of the superfamily 2 (SF2) helicases. Crystallographic analysis using a labeled RNA oligonucleotide allowed us to unambiguously track the positional changes of RNA bound to full-length HCV NS3 during two discrete steps of the ATP hydrolytic cycle. The crystal structures of HCV NS3, NS3 bound to bromine-labeled RNA, and a tertiary complex of NS3 bound to labeled RNA and a non-hydrolyzable ATP analog provide a direct view of how large domain movements resulting from ATP binding and hydrolysis allow the enzyme to translocate along the phosphodiester backbone. While directional translocation of HCV NS3 by a single base pair per ATP hydrolyzed is observed, the 3'-end of the RNA does not shift register with respect to a conserved tryptophan residue, supporting a “spring-loading” mechanism that leads to larger steps by the enzyme as it moves along a nucleic acid substrate.

INTRODUCTION

HCV NS3 is a DExH/D-box motor protein and a member of the SF2 helicases, based on the sequence classification scheme of Gorbalenya and Koonin¹. SF2 includes a number of nucleic acid remodeling enzymes that hydrolyze ATP and move in a unidirectional fashion along a single-stranded nucleic acid referred to as the tracking strand^{2; 3; 4}. Given the range of functions for this class of enzymes, movement of HCV NS3 along a tracking strand may result not only in the unwinding of a duplex substrate, but also in the displacement of bound proteins or additional remodeling of the viral RNA. Although the exact role of NS3 helicase and translocase activity in the HCV life cycle is currently not known, fully functional NS3 is required for replication of the RNA genome^{5; 6}. Targeting HCV NS3 helicase activity may prove to be an effective avenue for the treatment of HCV infection, a disease affecting

© 2010 Elsevier Ltd. All rights reserved.

*To whom correspondence should be addressed. Address: 333 Lakeside Dr., Foster City, CA, 94404, USA; Phone: 650-522-1722; Fax: 650-522-5266; todd.appleby@gilead.com.

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.

Olga Fedorova and Anna M. Pyle have no financial interest pertaining to this work.

approximately 200 million people worldwide including up to 4.9 million people in the United States (about 1.9% of the population)^{7; 8; 9}. In addition to being an attractive anti-viral target, NS3 serves as a model SF2 helicase, and numerous studies have been performed to further characterize its ATPase, nucleic acid binding, and strand translocation/duplex unwinding activities^{10; 11; 12; 13; 14; 15; 16; 17}.

HCV NS3 is a large (~67 kDa), multifunctional enzyme. The N-terminal 181 residues of NS3 (NS3pro) together with a span of residues from its NS4A protein co-factor are responsible for serine protease activity^{18; 19}. The remaining C-terminal 450 residues of HCV NS3 (NS3hel) comprise a platform for the ATPase^{20; 21}, nucleic acid binding^{22; 23}, and duplex unwinding activities of NS3^{20; 22; 24}. Several crystal structures of full-length NS3 and isolated NS3hel have been determined^{25; 26; 27; 28; 29; 30} and confirm that NS3hel is comprised of two RecA-like α/β domains (domains 1 and 2) and a third alpha-helical domain (domain 3). Unlike other members of SF2, the full-length NS3 structure reveals that the three domains comprising NS3hel are positioned on top of NS3pro with the C-terminus bound in the protease active site²⁶. Crystal structures of superfamily 1 (SF1) helicases in complex with various substrate and substrate analogs have been reported^{31; 32; 33} and provide insight into the mechanisms of translocation and unwinding by members of this family of enzymes. Similar studies on dengue virus NS3 helicase have greatly improved our understanding of the structure-function relationships for SF2 helicases³⁴. Recently, conformational snapshots of isolated HCV NS3hel have revealed new details about nucleotide-dependent structural changes that shed light on how this enzyme translocates along single-stranded DNA (ssDNA)³⁰.

The current body of structural information and results from numerous biochemical studies provide a link between the various activities of HCV NS3 and the structural elements that physically connect them. In depth reviews of key findings and proposed mechanisms of action have been reported previously^{4; 35}. Despite a wealth of mechanistic studies and the recent structural insights into how the enzyme may translocate along ssDNA, a detailed view of intact NS3 interacting with its physiologically relevant substrate, RNA, remains lacking.

To understand how ATP binding and hydrolysis leads to translocation of NS3 along RNA, we determined the high resolution crystal structures of NS3 alone, NS3 in complex with a bromine-labeled, single-stranded RNA oligonucleotide (ssRNA), and a ternary complex of NS3 bound to ssRNA and a non-hydrolyzable ATP analog. Due to the observed enhancement in direct and functional binding of RNA to intact NS3 compared to isolated NS3hel¹⁰, we chose to conduct these studies on full-length NS3 protein. Analysis of the structures reveals details about the conformational changes in NS3 that occur in the presence and absence of nucleotide. By utilizing a labeled oligonucleotide, we are able to unambiguously track the positional changes of the RNA in the nucleotide-bound and nucleotide-free forms of the enzyme. Based on comparisons to results from the previous structural studies performed using the isolated NS3hel and ssDNA³⁰, the structures presented here reveal key differences in how the full-length enzyme interacts with an RNA substrate. By tracking the precise register of the RNA using a bromine label, we propose a molecular mechanism by which HCV NS3 couples ATP hydrolysis to movement along a ssRNA tracking strand.

RESULTS and DISCUSSION

NS3 bound to RNA

To understand how HCV NS3 interacts with the RNA tracking strand in the absence of ATP, we determined the binary complex structure of the full-length enzyme bound to a short

ssRNA oligonucleotide. In order to establish and track the register of the ssRNA in the structure, we used a poly-uridine 8-mer containing a brominated uracil base on the nucleic acid residue at position 4 (5'-U1-U2-U3-U4[Br]-U5-U6-U7-U8-3'). This bromine-labeled ssRNA will be referred to as rU8Br4 throughout this report. Although there are two molecules of NS3 in the crystallographic asymmetric unit, only one of the molecules contained strong density indicating high occupancy for the rU8Br4. It is likely that this result is an artifact of crystallization and not of functional significance. The structure reveals a sulfate ion bound to residues of the conserved Walker A motif (motif I) in the nucleotide binding site of domain 1. Lithium sulfate is a major component of the crystallization mixture. Evidence for this anion binding site was first reported by Yao and colleagues²⁶ who noted the presence of a bound phosphate ion. The resulting 2.0 Å resolution structure of full-length NS3 bound to rU8Br4 (NS3-rU8Br4) reveals a binding mode similar to that seen in structures of isolated NS3hel in complex with ssDNA (NS3hel-ssDNA)^{25, 30} and also provides new insights into key protein-RNA interactions and their implications for ATP-dependent translocation along RNA.

As seen previously in the NS3hel-ssDNA structures, the oligonucleotide in our structure binds in a large channel that separates RecA-like domains 1 (D1) and 2 (D2) from domain 3 (D3) (Fig. 1). Inspection of the maps reveals interpretable electron density for five out of eight nucleotide residues of the ssRNA (Fig. 2a). The absence of complete electron density for nucleotide residues U1 and U2 at the 5'-end of the RNA and nucleotide U8 at the 3'-terminus suggests that they do not form strong interactions with NS3 and are disordered in the structure. This result is consistent with the NS3hel-ssDNA structures, which reveal five nucleotide residues bound between the conserved "bookend" residues (Val 432 and Trp 501) of the enzyme in the nucleotide-free state. Strong residual density adjacent to a single uracil base in the difference map calculated before including the electron-dense bromine atom in the model, allowed us to identify the position of bromine-labeled nucleotide residue, uridine 4 (U4), and confirm that the ssRNA is bound to NS3 in a consistent register throughout the crystal. Based on the location of U4 and the presence of strong phosphate density at both the 5' and 3' ends of the visible strand, it is clear that the enzyme is not selectively binding to the 3' or 5' terminal end of the ssRNA.

The complex of NS3 and rU8Br4 reveals that the ssRNA binds in an extended conformation and interacts with all three helicase domains (Fig. 3a). NS3 residues in D2 interact with the 5'-end of the ssRNA, while the 3'-end forms contacts with protein residues in D1 and D3. The uracil base of U7, the last visible nucleotide at the 3'-end of the RNA, is anchored by a stacking interaction with the side chain of Trp 501 from D3. Interestingly, U7 is in the *syn* conformation while the other four visible nucleosides adopt an *anti* conformation. This result is also observed for the corresponding deoxynucleotides in the NS3hel-ssDNA structures. The vast majority of specific NS3-RNA interactions involve polar protein atoms donating hydrogen bonds to the phosphoryl oxygens of the ssRNA backbone. Protein atoms involved in hydrogen bonding include the backbone NH groups belonging to Lys 371 (motif I), Arg 393, Val 232 (motif Ia), and Gly 255, as well as the polar side chains of Thr 411 (motif V), Thr 416 (motif V), and Thr 269 (motif Ib). The side chain amino group of Lys 272 forms what appears to be the only charge-charge interaction with the ssRNA, forming contacts with the phosphate linking U7 to the disordered U8 nucleotide residue at the 3'-end of the strand.

Other than the ring stacking interaction between Trp 501 and U7, the only other contact between NS3 and an RNA base involves the side chain of Gln 556 which forms a solvent exposed hydrogen bond to the 4-carbonyl oxygen of U3. The observed lack of base recognition by NS3 is expected and supports the functional requirement for the enzyme to translocate efficiently along the HCV RNA genome in a sequence independent fashion.

Unlike the dengue virus NS3 helicase³⁴, HCV NS3 does not form direct interactions with the 2'-hydroxyl moieties of RNA. The lack of sugar specificity is consistent with the fact that despite RNA being the physiological substrate for HCV NS3, the enzyme has been shown to efficiently utilize both RNA and DNA substrates³⁶. The specific contacts observed between NS3 D2 with the 5'-end and D1 with the 3'-end of the ssRNA phosphodiester backbone in conjunction with the base stacking interaction involving Trp 501 of D3, establishes an asymmetric arrangement of protein structural elements with respect to the substrate and help promote the observed 3' to 5' polarity of the enzyme¹⁴.

Structural flexibility allows NS3 to bind both DNA and RNA substrates

Although SF2 helicases tend to be specific for either RNA or DNA, HCV NS3 can bind to and unwind both nucleic acid substrates³⁶. With ssDNA and ssRNA-bound forms of NS3 now available, we wanted to compare the structures and determine what features of NS3 account for its dual substrate specificity. After roughly aligning an ensemble of structures, we noted that there was excellent agreement of D3 across all structures and decided to perform our structural alignments using residues 487–623 of D3 as the frame of reference. Comparisons between the isolated NS3hel domain bound to ssDNA (**PDB ID: 3KQH**)³⁰ and our NS3-rU8Br4 structure reveal a number of similarities and highlight a few key differences in how NS3 interacts with DNA and RNA substrates.

In both structures, adjacent nucleotide residues base stack against one another, and in fact, the overall position of the bases in the binding cleft overlap quite well (Fig. 3b). Variation is observed along the phosphodiester backbone, however, with the divergence between overlapping phosphates increasing in the 3' to 5' direction of the two strands. Using our nucleotide residue numbering convention established by the bromine-labeled ssRNA, we observe that the deoxyribose rings of nucleotides 3 and 4 at the 5'-end of the ssDNA adopt a 2'-endo pucker (B-form-like), while the ssRNA in our structure maintains a uniform 3'-endo conformation (A-form-like) along the entire length of visible nucleotide residues. The shorter distance between consecutive phosphate groups in the ssRNA, is a consequence of the C3'-endo ring puckering. Also, unlike the ssDNA, the ssRNA in the NS3-rU8Br4 structure is able to form intramolecular hydrogen bonds between the 2'-hydroxyl groups and 4'-oxygen atoms of several adjacent ribose moieties. The hydrogen bonding interactions are observed between U3 and U4 at the 5'-end of the RNA, as well as between U5–U6 and U6–U7 at the 3'-end of the strand. Together, the intramolecular hydrogen bonding and 3'-endo conformation of the ribose rings result in a more compact structure for the ssRNA backbone compared to that observed for the ssDNA.

Because NS3 interacts mainly with the phosphodiester backbone of the nucleic acid tracking strand, we next wanted to determine how the enzyme adjusts to accommodate both ssDNA and ssRNA. The overlay of the two protein structures reveals that D2 in NS3hel-ssDNA pivots away from D1 in order to accommodate the extended phosphodiester backbone of the ssDNA. This results in an approximate 2 Å shift in the D2 residues lining the nucleic acid binding cleft relative to the corresponding residues in NS3-rU8Br4 that bind to the more compact ssRNA backbone. A significant change in the local secondary structure of the protein is also observed in this region (Fig. 3b). In NS3hel-ssDNA, residues in motif V of D2 form an extended loop segment, whereas in our NS3-rU8Br4 complex, those residues maintain a compact helical arrangement observed in the full-length apo structure. In NS3-rU8Br4, the short helix positions Thr 416 in the nucleic acid binding cleft where the side chain hydroxyl group forms a hydrogen bond with the ssRNA backbone linking nucleotide residues U4 and U5. The corresponding phosphate linkage in NS3hel-ssDNA is shifted approximately 2 Å (in the 5' direction) where the backbone NH of Leu 414 in motif V acts as a hydrogen bond donor rather than the side chain of Thr 416 which is sequestered from the nucleic acid binding cleft (over 6 Å away from the ssDNA backbone). Inter-domain

flexibility, local structural rearrangements, and a distinct lack of nucleic acid sugar and base specificity all contribute to the ability of HCV NS3 to utilize both RNA and DNA substrates.

NS3 bound to RNA and a non-hydrolyzable nucleotide analog

In order to trap HCV NS3 associated with RNA in a nucleotide-bound state, we soaked co-crystals of NS3-rU8Br4 and NS3-rU8 (NS3 bound to an unlabeled poly-uridine 8-mer) in a mixture of ATP, MgCl₂, BeCl₂, and NaF to generate the non-hydrolyzable ATP analog, ADP•BeF₃. A similar approach was used to determine structures of the trapped complexes of the myosin II and F₁-ATPase motor proteins^{37; 38} as well as the isolated HCV NS3hel domain bound to ssDNA³⁰. In all cases, the investigators concluded that the ADP•BeF₃-bound structures represented the ground state, ATP-bound conformation of the enzyme preceding hydrolysis. Enzyme studies have also shown that biochemically, ADP•BeF₃ can act as a pre-hydrolytic ATP mimic for helicase SF2 DEAD-box proteins³⁹. Analysis of our structure confirms the formation of a ternary complex of enzyme, ssRNA, magnesium, and the non-hydrolyzable ATP analog (NS3-rU8Br4/ADP•BeF₃).

The protein structures of NS3-rU8/ADP•BeF₃ (2.05 Å resolution) and NS3-rU8Br4/ADP•BeF₃ (2.3 Å resolution) are essentially identical. The higher resolution structure, NS3-rU8/ADP•BeF₃, is useful for detailed analysis of the nucleotide binding site, while the NS3-rU8Br4/ADP•BeF₃ structure allows for precise tracking of the ssRNA due to the presence of the bromine-labeled nucleotide residue. Working with full-length NS3 structures allows us to perform alignments using NS3pro as a frame of reference to monitor movements of individual helicase domains and RNA substrates relative to a fixed point. Based on an alignment to the nucleotide-free NS3-rU8Br4 structure, the ternary complex structures reveal that D1 and D2 undergo a significant conformational change upon ADP•BeF₃ binding (Fig. 4a and 5a). D2 experiences the greatest shift with an overall root mean square deviation (RMSD) of 1.3 Å² for 138 residues (based on α-carbons only) as it pivots towards D1 by approximately 10°. D1 displays a more subtle rearrangement upon binding the nucleotide analogue with an overall RMSD of 0.5 Å² for 137 residues (based on α-carbons only). The most notable shift in D1 results from the tilting of a long α-helix (residues 232–246), termed the “spring helix”³⁰, towards domain 2. This helix contains Val 232 (motif Ia) at its N-terminal end and Tyr 241 (Y motif) near its carboxy terminal end. As noted above in the analysis of the NS3-rU8Br4 structure, the backbone NH group of Val 232 forms a direct hydrogen bond with a phosphoryl oxygen from U6 of the ssRNA. In the ternary complex, the Val 232-RNA interaction persists as the spring helix tilts towards D2 allowing the side chain phenol of Tyr 241 to stack against the adenine base of ADP•BeF₃.

NS3-nucleotide interactions

The NS3-rU8/ADP•BeF₃ structure confirms that full-length NS3 interacts with ADP•BeF₃ and a coordinated magnesium ion (Mg²⁺) in a similar fashion to isolated NS3hel in complex with ADP•BeF₃ and manganese³⁰. Residues in D1 either coordinate the Mg²⁺ or contact the γ-beryllium trifluoride (γ-BeF₃) and phosphate groups of the ADP directly (Fig. 4b). Asp 290 and Glu 291 (D and E of the DEx/H sequence) in motif II, together with serine 211 of motif I (Walker A) and two well ordered water molecules, form a hydrogen bond network that results in the chelation of a single Mg²⁺ which in turn forms a bridge between the β-phosphate and γ-BeF₃ groups of the nucleotide analog, thus completing the hexacoordination of the metal ion. The side chain amino group of Lys 210 (motif I) also forms a salt-bridge between the β-phosphate and γ-BeF₃ groups. The polar hydroxyl side chain of Thr 212 and the backbone NH groups from several residues in motif I (residues 207–212) provide additional hydrogen bonds to both the α and β-phosphate groups of ADP•BeF₃.

While residues in D1 form extensive interactions with the nucleotide analog, considerably fewer specific contacts are observed between D2 and ADP•BeF₃ (Fig. 4b). The nucleotide-induced collapse of D2 towards D1 repositions Thr 419 (motif V), which now packs against the adenine base opposite of the Tyr 241 ring stacking interaction. Structural rearrangements in motif VI also result in the movement of two conserved arginine residues (Arg 464 and Arg 467) towards D1 (approximately 6 Å and 4.8 Å for Arg 464 and Arg 467, respectively), placing the basic guanidinium group of both arginines in close proximity to the negatively charged γ -BeF₃ moiety of ADP•BeF₃. This result suggests that Arg 464 and Arg 467 are key residues that can contribute to ATP hydrolysis by stabilizing the transition state and also trigger subsequent conformational changes in D2 following hydrolysis. As ATP is hydrolyzed to ADP, the bond between the β and γ -phosphates is broken, forming ADP and inorganic phosphate. If Arg 464 and Arg 467 maintain contact primarily with the γ -phosphate group (now phosphate ion), they would presumably be free to pull away from the ATP binding site, allowing D2 to relax back into the open state. ADP and phosphate would then be exposed to solvent and released from the active site, allowing the enzyme to prepare for another cycle of nucleotide binding and hydrolysis. The involvement of Arg 464 and Arg 467 in nucleotide binding has been previously proposed³⁵, and site directed mutagenesis studies have supported the important role that these two arginine residues play in the ATPase and unwinding activities of NS3^{40; 41; 42}. Gln 460 in D2 positions a water molecule near the γ -BeF₃. If activated, possibly by nearby Glu 291³⁰, this water molecule could perform an in-line nucleophilic attack on the γ -phosphate of ATP during hydrolysis. Mutations of either Gln 460 or Glu 291 completely abolish ATPase activity^{16; 41}, suggesting a critical role for this residue during catalysis.

Nucleotide induced changes in ssRNA binding

Structural rearrangements that occur in NS3 when it is bound to ADP•BeF₃ lead to changes in how the enzyme interacts with ssRNA. Overall, the total number of contacts between NS3 and ssRNA are reduced in the nucleotide-bound state which is consistent with biochemical results demonstrating that NS3 has a reduced affinity for RNA and DNA in the presence of nucleotides^{10; 14}. As discussed above, binding of ADP•BeF₃ to NS3 causes D1 and D2 to pivot together and close around the nucleotide analog. These domain movements are accompanied by a significant conformational change in motif V (Fig. 5a and b), moving Thr 416 away from the RNA binding cleft and disrupting a hydrogen bond to the phosphodiester backbone of the ssRNA. Due to the rearrangement of motif V and the loss of the Thr 416-RNA interaction in the nucleotide-bound state, the interaction between NS3 and the 5'-end of the RNA is weakened. Only four bases (U4–U7) are stacked in the nucleic acid binding cleft of NS3-rU8/ADP•BeF₃ compared to the five bases (U3–U7) that are stacked in the nucleotide-free structure (Fig. 5c).

ATP hydrolysis and 3'→5' translocation of NS3 along ssRNA

Although much is known about the interaction of NS3^{hel} with ssDNA, structures allowing us to visualize how full-length NS3 coordinates ATPase activity and movement along RNA are lacking. The current study offers views of NS3 bound to an RNA tracking strand at two key stages during the ATPase cycle. The NS3-rU8/ADP•BeF₃ structure mimics the nucleotide-bound state of NS3 immediately preceding ATP hydrolysis (the substrate complex). The NS3-rU8Br4 structure represents the step following ATP hydrolysis after the reaction products, ADP and phosphate, have been released by the enzyme (the product complex). By overlaying and comparing these two structures, we are able to identify specific ATP-dependent conformational changes that, in turn, lead to changes in ssRNA binding and reveal and overall movement of the protein along the RNA in a 3'–5' direction.

To illustrate the ATP-dependent translocation of HCV NS3 along RNA using our two structures (Fig. 6), we first consider the substrate complex of NS3 bound to ssRNA and ADPBe•F3. In the nucleotide-bound state, motif V rearranges allowing D1 and D2 to collapse together, sequestering Thr 416 from the RNA binding cleft where it is prevented from interacting with the ssRNA backbone. In this closed conformation, NS3 forms fewer interactions with the ssRNA resulting in lower affinity for the substrate. Near the 5'-end of the ssRNA, the side chain hydroxyl group of Thr 411 and the backbone NH of Arg 393 form hydrogen bonds to the phosphodiester linkage between U4 and U5. U3 is outside of the RNA binding cleft and does not form strong interactions with the enzyme. Our nucleotide-free structure reveals that after hydrolysis of ATP and release of ADP and phosphate, D1 and D2 relax back into an open conformation. The domain movement is accompanied by a rearrangement of motif V, exposing Thr 416 to the RNA binding cleft where it now interacts with the phosphodiester linkage between U4 and U5. As D2 relaxes into the open conformation, it shifts register along the ssRNA backbone by one nucleotide, allowing Thr 411 and the backbone NH of Arg 393 to form a new contact with the phosphate linkage between U4 and U3. In this open state, NS3 binds the RNA with higher affinity and effectively pulls an additional nucleotide residue (U3) into the RNA binding cleft (corresponding to a 3'→5' movement of D2 along the ssRNA). Together, these two structures suggest that the conformational changes that occur in NS3 going from the nucleotide-bound state to the nucleotide-free state result directly in movement of the enzyme along the phosphodiester backbone of the RNA tracking strand by a single nucleotide residue in the 3'→5' direction, confirming a translocation step-size of one nucleic acid residue per cycle of ATP hydrolysis.

Our structural model of NS3 translocation is consistent with the Brownian motor mechanism first suggested by Levin et al.¹⁴. The Brownian motor model proposes that the ATPase activity of NS3 allows it to cycle between high-affinity (nucleotide-free) and low-affinity (nucleotide-bound) conformational states allowing the enzyme to grab and release the RNA tracking strand. In the low affinity conformation, NS3 could potentially slide either forward or backwards along the tracking strand, but because NS3 binds with highest affinity to single-stranded (ss)/double-stranded (ds) junctions¹⁴, it will have a propensity to move in the forward (3'→5') direction. Our structures suggest that even in the absence of a ss/ds junction NS3 will move along ssRNA backbone in the 3'→5' as it transitions from the nucleotide-bound to the nucleotide-free conformational states.

NS3 translocation visualized by our structures is also in excellent agreement with a second, complementary translocation model which describes SF2 helicases as backbone stepping motors^{15; 43}. Central to this model are a pair of residues, one residing in D1 and a second in D2⁴ which act as pincers to grip the nucleic acid phosphodiester backbone. During cycles of ATP hydrolysis, conformational changes in the enzyme cause D1 and D2 to close (nucleotide-bound) and open (nucleotide-free). In the nucleotide-bound state, the pincer residues span two nucleotide residues, while in the nucleotide-free form they span three nucleotide residues⁴³, suggesting a step-size along the nucleic acid backbone by a single-nucleotide increment per molecule of ATP hydrolyzed (Fig. 7). Consistent with this model, our structures identify two conserved threonine residues (Thr 269 in D1 and Thr 411 in D2) that interact with the phosphodiester backbone of RNA in both the nucleotide-bound and nucleotide-free states. In the nucleotide-bound state, Thr 269 and Thr 411 span two bases (U5–U6). In the nucleotide-free state, the pincers open up allowing Thr 411 to step along the ssRNA backbone and span an additional base (U4) towards the 5'-end of the RNA. Site directed mutational studies reveal that replacing either Thr 269 or Thr 411 in HCV NS3 with an alanine residue results not only reduced binding to ssRNA, but also in a complete loss of helicase activity⁴⁴.

Implications for helicase activity

In the absence of a structure containing NS3 and an RNA duplex substrate, it is more difficult to speculate about the precise mechanism of strand unwinding and the effect that the single-stranded/double-stranded (ss/ds) junction might have on the translocation mechanism that we have observed. If the enzyme utilizes a passive mechanism^{45; 46}, processive unwinding would essentially proceed as a consequence of the strand translocation mechanism that we have presented above in which NS3 hydrolyzes ATP to fuel the switch between tight and weak RNA binding modes near the ss/ds junction, opportunistically capturing ssRNA that results from spontaneous melting of the duplex at the junction.

NS3 may also unwind duplex RNA using a more active strand displacement mechanism. Although we observe translocation of D2 by one nucleotide along the 5'-end of the ssRNA, D1 and D3 maintain a relatively fixed grip on the 3'-end of the substrate (no change of register along the strand). This may be evidence of a mechanism critical to NS3 helicase activity. While single-molecule studies on HCV NS3 confirm a translocation step-size of one base per ATP hydrolyzed, they also reveal a larger burst which may have functional significance for strand unwinding¹⁵. A key feature of the model that emerged from these studies is that D3 utilizes a mechanism allowing it to lag behind D1 and D2 for up to three cycles of ATP hydrolysis, effectively building up tension in the system. After three rounds of D1 and D2 opening and closing, D3 springs forward by three bases, providing NS3 with the necessary energy to unwind three additional base pairs at the RNA duplex junction. Our structures offer support to a similar "spring-loaded" mechanism of unwinding. The movement of RNA observed going from the nucleotide-bound to nucleotide-free structures of NS3 presented here, reveals that D2 can step by one nucleotide residue at the 5'-end of the ssRNA while D1 and D3 maintain a fixed grip on the 3'-end of the ssRNA. The fact that U7 remains stacked against Trp501 in both the nucleotide-bound and nucleotide-free structures suggests that Trp 501 can act as an anchor point during individual translocation steps. According to the proposed model, the tension that accumulates in the system during two additional translocation steps might then trigger a conformational change in NS3 allowing bases at the 3'-end of the strand to slip past Trp 501 as the enzyme springs forward along the tracking strand in the 5' direction. Mutational analysis has confirmed a critical role for Trp 501 not only in ssRNA binding, but also for duplex unwinding⁴⁴. Additional structural studies utilizing RNA substrates containing ss/ds junctions will be required to fully characterize NS3 unwinding activity.

The structures reported here provide details about the conformational changes that occur in HCV NS3 during cycles of nucleotide binding and product release. More importantly, we link these structural rearrangements directly to changes in ssRNA binding. The observation that NS3 moves by a single nucleotide residue in the 3'→5' direction along ssRNA as a result of transitioning between closed (nucleotide-bound) and open (nucleotide-free) states confirms a step size of one base per ATP hydrolyzed. Together, these structures provide tremendous insight into the molecular mechanism by which HCV NS3 couples ATP hydrolysis to translocation along ssRNA.

EXPERIMENTAL PROCEDURES

Construct design and protein expression

The expression construct, pET15b-4A-NS3FL, is similar to that used previously to determine the structure of full-length NS3²⁶. Briefly, the construct was engineered as a single-chain polypeptide with the core peptide (GSVVIVGRIILS) of HCV NS4A tethered to the N-terminus of full-length NS3 (Con-1 sequence) using a short linker (Gly-Ser-Gly-Ser). The NS3 Con-1 DNA was amplified from the HCV 1b I389luc-ubi-neo/NS3-3'/ET replicon

plasmid provided by R. Bartenschlager. The coding sequence of this fused 4A-NS3 peptide was subsequently cloned into a pET15b expression vector at the Nde I and Xho I sites, generating an N-terminal, Thrombin-cleavable His₆-tag (MGSSHHHHHSSGLVPRGS).

After confirming the sequence, the expression vector was transformed into BL21DE3 CodonPlus *E. coli* cells. A 1 L overnight culture was prepared from a single colony and used to inoculate 18 L of 2YT medium containing 100 µg/ml ampicillin and 34 mg/ml chloramphenicol in a C20 fermentor (Sartorius BBI System Inc., Bethlehem, PA). When the cell density reached an OD₆₀₀ of 1.0, the temperature of the culture was reduced from 37°C to 22°C. 30 mM ZnSO₄ and 0.5 mM IPTG were added to induce recombinant protein expression. Cells were harvested after overnight induction and frozen at -80 °C prior to purification of the protein.

Protein purification

Cell pellets were thawed and resuspended (10 mL/g cells) in lysis buffer (50 mM Tris, pH 7.6, 300 mM NaCl, 4 mM MgCl₂, 0.1% CHAPS, 5% glycerol, and 2 mM mercaptoethanol). The cell suspension was then passed three times through a microfluidizer at 18,000 psi. The homogenate was cleared by centrifugation at 15,500 rpm for 45 minutes, and the resulting supernatant was loaded onto a Ni-NTA column which was pre-equilibrated with buffer A (lysis buffer supplemented with 50 mM imidazole). Proteins were eluted with a gradient of 500 mM imidazole and selected fractions were pooled. This Ni-NTA pool was diluted three-fold with buffer B (25 mM HEPES pH 7.5, 10% glycerol, 50 mM NaCl, 0.05% n-octyl-β-octylglucoside, 5 mM DTT, and 4 mM MgCl₂) before it was loaded onto a pre-equilibrated Heparin HP column. The recombinant protein was eluted using a 1 M NaCl gradient. Fractions containing the His₆-4A-NS3 protein were pooled and TCEP was added to a final concentration of 10 mM. The Heparin-HP pool was concentrated and loaded onto a pre-equilibrated Sephacryl S-100 16/60 column. The protein was eluted in buffer C (25 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol and 2 mM TCEP). Fractions containing purified His₆-4A-NS3 were pooled and TCEP was added to a final concentration of 10 mM. The protein was concentrated to 7–10 mg/ml and stored at 4°C prior to crystallization.

Oligonucleotide synthesis and purification

A poly-uridine RNA oligonucleotide with a single 5-bromo-uridine substitution at position four of the 8-mer (5'-U1-U2-U3-U4[Br]-U5-U6-U7-U8-3') was synthesized on an automated DNA-RNA synthesizer MerMade 6 (BioAutomation) using phosphoramidite chemistry. The labeled oligonucleotide (rU8Br4) was removed from the polymer support by incubation with 28–30% ammonium hydroxide:ethanol mix (3:1) at room temperature for 22–24 hours. 2'-hydroxyl groups were deprotected as previously described⁴⁷. Excess of the fluoride reagent was removed by incubation with isopropoxytrimethylsilane at room temperature for 20 minutes⁴⁸. rU8Br4 was subsequently precipitated by diethyl ether, purified on a denaturing 20% polyacrylamide gel, and eluted from the gel as previously reported⁴⁹.

Crystallization and structure determination

Preliminary crystallization attempts using purified protein with an intact N-terminal His₆-tag proved successful, therefore Thrombin cleavage of the fusion protein was never performed. Although disordered in the structure, the His₆-tag is located on the opposite side of the helicase domain with respect to the RNA binding site (~25 Å away) and should not interfere directly with RNA binding. Diffraction-quality crystals of His₆-4A-NS3 grow at room temperature using the hanging-drop vapor diffusion technique. Drops (6 µl) contain a 2:1 mixture of protein (diluted to 5.0 mg/mL) and reservoir solution. The protein solution was prepared as described above. The reservoir solution contained 20% (w/v) PEG 3350, 160

mM LiSO₄, 80 mM Bis-Tris (pH 6.5) and 10% (v/v) glycerol. Drops are equilibrated against 1.0 ml of reservoir solution following microseeding. Crystals continue to grow for approximately 7 to 10 days.

Data on all crystals utilized in this study were collected on the Sector 5 beamlines (5.0.1 or 5.0.2) at the Advanced Light Source (Berkeley, CA). Prior to data collection, all crystals were soaked briefly in reservoir solution supplemented with 15% glycerol and flash frozen in liquid nitrogen. The data were collected at cryogenic temperatures on either a Quantum-210 CCD detector (beamline 5.0.1) or a Quantum-315 CCD detector (beamline 5.0.2) using 0.5° oscillations until greater than 90% overall completeness was observed. All data sets were processed with the programs DENZO and SCALEPACK⁵⁰. Data collection statistics are summarized in Table 1. All crystals used in this study are orthorhombic (space group P2₁2₁2₁) with similar unit cell dimensions (a= 91 Å, b= 111Å, c= 142 Å) and contain two molecules in the asymmetric unit.

Full-length NS3

The structure of apo His₆-4A-NS3 (NS3) was determined by molecular replacement using data collected to 1.95 Å resolution, the previously determined full-length NS3 structure (**PDB ID: 1CU1**)²⁶ as the search model, and the program EPMR⁵¹. The new model was subjected to rigid-body refinement followed by several iterative cycles of torsion angle annealing in CNX (Accelrys, San Diego, CA) and manual refitting in COOT⁵². Although the Thrombin-cleavable, N-terminal His₆-tag was left intact, it is disordered in the structure. The first residues visible at the N-terminal region of the molecule belong to the fused NS4A peptide (GSVVIVGRIILS). The C-terminus of NS3 is visible in the structure and is bound in the active site of the protease domain. Water molecules and four sulfate ions were eventually added to the final model after an inspection of difference density maps and based on potential hydrogen bonding to nearby protein atoms. Lithium sulfate is a major component of the crystallization mixture.

Full-length NS3 bound to RNA

Bromine labeled single-stranded RNA used for crystal studies was synthesized as described above. Unlabeled poly-uridine ssRNA of eight nucleotides in length (rU8) was purchased from TriLink Biotechnologies (San Diego, CA). His₆-4A-NS3 (diluted to 5.0 mg/mL) was mixed with a 3-fold molar excess of either rU8 or the rU8Br4 ssRNA and allowed to incubate on ice for at least one hour prior to crystallization using the procedure described above to generate apo NS3 crystals. 24 hours prior to data collection, the crystals were harvested and soaked in reservoir solution supplemented with either 1mM rU8 or rU8Br4. Crystals of full-length NS3 bound to rU8 (NS3-rU8) and full-length NS3 bound to the labeled rU8Br4 (NS3-rU8Br4) were then flash-cooled as described above.

Data were collected on a frozen NS3-rU8Br4 crystal to 2.0 Å resolution. Using the apo NS3 structure determined above as a starting model, the processed NS3-rU8Br4 data were used to perform rigid-body refinement and several rounds of torsion angle annealing. After making necessary adjustments to the new model, inspection of difference electron density maps revealed strong density in the nucleic acid binding cleft for one molecule of the asymmetric unit (molecule A). The second NS3 molecule (molecule B) contained weak, non-contiguous density for the ssRNA indicating very low occupancy at this site. Based on these results, the final model contains only one ssRNA molecule associated with NS3 molecule A (Fig. 2a). The position of the bromine-labeled uridine residue was determined by omitting the bromine atom from the initial ssRNA model and calculating an Fo-Fc difference electron density map. This map revealed a unique, strong peak adjacent to the C5

atom of a single uridine base in the oligonucleotide (U4). During later stages of refinement water molecules and two sulfate ions were also added to the model.

Full-length NS3 bound to RNA and ADP•BeF₃

Crystals of NS3-rU8 (prepared as above) were soaked in reservoir solution supplemented with 1 mM ATP, 1 mM MgCl₂, 1mM BeCl₂, and 5 mM NaF for ten to fifteen hours using a procedure similar to that reported by Alani and colleagues⁵³. Crystals displayed varying amounts fraying and cracking within four to six hours of soaking. The most visibly intact crystals were harvested and flash-cooled prior to data collection. Diffraction data were collected on a single NS3-rU8/ADP•BeF₃ crystal and used to perform rigid-body refinement with apo NS3 as the starting model. The crystallographic R and free R values remained higher than expected compared to the earlier results from both the apo NS3 and NS3-rU8Br4 refinements. During subsequent rounds of torsion angle annealing and manual rebuilding of the model, it was apparent that there were significant conformational changes in D1 and D2 of NS3. The nucleotide analog, ADP•BeF₃, and a magnesium ion were subsequently modeled into regions of strong difference density (Fig. 2c) located at the interface between D1 and D2 in both molecules A and B. Density for the ssRNA in the RNA binding cleft of both NS3 molecules is extremely weak. It is possible that since additional rU8 was not included in the overnight soaking mixture, the RNA diffused out of the crystal resulting in poor occupancy. Although weak density is visible for a portion of the phosphodiester backbone of the RNA associated with molecule A, no RNA was included in the final model. Several water molecules were added to the model during later rounds of refinement. The resulting 2.05 Å resolution structure provides a detailed view of the NS3 bound to ADP•BeF₃.

A similar soaking procedure was carried out using NS3 crystallized in the presence of the labeled RNA, rU8Br4. For this experiment, the soaking solution contained 1 mM ATP, 1 mM MgCl₂, 1mM BeCl₂, and 5 mM NaF supplemented with 1 mM rU8Br4 RNA. The crystals were removed from the soaking solution after 6 hours due to visible deterioration. Data was subsequently collected on a frozen crystal of NS3-rU8Br4/ADP•BeF₃ to 2.3 Å resolution. The structure was determined using the protein coordinates from the refined NS3-rU8/ADP•BeF₃ structure (above) as the starting model. One magnesium ion and a single molecule of the nucleotide analog, ADP•BeF₃, were subsequently modeled into a region of strong difference density located at the interface between D1 and D2 in molecule A. Overall, the density for the ssRNA was weaker than that observed in the NS3-rU8Br4 maps. There was, however, contiguous density for the phosphodiester backbone in the RNA binding cleft of molecule A, and four consecutive uridine nucleotide residues were unambiguously fit into the remaining density (Fig. 2b). The position of the bromine-labeled uridine was determined in the same fashion as the NS3-rU8Br4 structure described above. Density for NS3 molecule B suggests that the nucleotide binding site is only partially occupied in the crystal. This observation may be a result of the reduced soaking time used in this experiment. Longer soaks were performed, but this resulted in poorly diffracting crystals (<3.3 Å resolution). Based on these results, the final model includes molecule A bound to ADP•BeF₃, a magnesium ion, and rU8Br4 while molecule B is modeled with D1 and D2 in the open state without nucleotide and RNA bound.

Structural alignments were performed and figures were generated using the molecular graphics program, PyMol (Schrödinger, LLC). Analysis of NS3 motions in the presence of ssRNA and nucleotide was aided by comparing the NS3-rU8Br4 and NS3-rU8Br4/ADP•BeF₃ structures using the DynDom Protein Domain Motion Analysis server (<http://fizz.cmp.uea.ac.uk/dyndom/>)⁵⁴. Structures have been deposited in the Protein Data Bank (<http://www.rcsb.org>) with accession numbers 3O8B (Apo NS3), 3O8C (NS3-rU8Br4), 3O8D (NS3-rU8/ADP•BeF₃), and 3O8R (NS3-rU8Br4/ADP•BeF₃).

Acknowledgments

The authors would like to thank the staff of The Advanced Light Source at the Lawrence Berkeley National Laboratory. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under contract No. DE-AC02-05CH11231. Todd C. Appleby, Robert L. Anderson, Ruth Wang, Xiaohong Liu, Katherine M. Brendza, and John R. Somoza are all employees and shareholders of Gilead Sciences, Inc. The majority of this work was funded by Gilead Sciences, Inc.

REFERENCES

1. Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Res.* 1989; 17:4713–4730. [PubMed: 2546125]
2. Bork P, Koonin EV. An expanding family of helicases within the 'DEAD/H' superfamily. *Nucleic Acids Res.* 1993; 21:751–752. [PubMed: 8382805]
3. Rocak S, Linder P. DEAD-box proteins: the driving forces behind RNA metabolism. *Nat Rev Mol Cell Biol.* 2004; 5:232–241. [PubMed: 14991003]
4. Pyle AM. Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu Rev Biophys.* 2008; 37:317–336. [PubMed: 18573084]
5. Beran RK, Lindenbach BD, Pyle AM. The NS4A protein of hepatitis C virus promotes RNA-coupled ATP hydrolysis by the NS3 helicase. *J Virol.* 2009; 83:3268–3275. [PubMed: 19153239]
6. Lam AM, Frick DN. Hepatitis C virus subgenomic replicon requires an active NS3 RNA helicase. *J Virol.* 2006; 80:404–411. [PubMed: 16352565]
7. Brown RS Jr, Gaglio PJ. Scope of worldwide hepatitis C problem. *Liver Transpl.* 2003; 9:S10–S13. [PubMed: 14586889]
8. Armstrong GL, Wasley A, Simard EP, McQuillan GM, Kuhnert WL, Alter MJ. The prevalence of hepatitis C virus infection in the United States, 1999 through 2002. *Ann Intern Med.* 2006; 144:705–714. [PubMed: 16702586]
9. Purcell R. The hepatitis C virus: overview. *Hepatology.* 1997; 26:11S–14S. [PubMed: 9305657]
10. Beran RK, Serebrov V, Pyle AM. The serine protease domain of hepatitis C viral NS3 activates RNA helicase activity by promoting the binding of RNA substrate. *J Biol Chem.* 2007; 282:34913–34920. [PubMed: 17921146]
11. Frick DN, Banik S, Rypma RS. Role of divalent metal cations in ATP hydrolysis catalyzed by the hepatitis C virus NS3 helicase: magnesium provides a bridge for ATP to fuel unwinding. *J Mol Biol.* 2007; 365:1017–1032. [PubMed: 17084859]
12. Frick DN, Rypma RS, Lam AM, Gu B. The nonstructural protein 3 protease/helicase requires an intact protease domain to unwind duplex RNA efficiently. *J Biol Chem.* 2004; 279:1269–1280. [PubMed: 14585830]
13. Lam AM, Keeney D, Frick DN. Two novel conserved motifs in the hepatitis C virus NS3 protein critical for helicase action. *J Biol Chem.* 2003; 278:44514–44524. [PubMed: 12944414]
14. Levin MK, Gurjar M, Patel SS. A Brownian motor mechanism of translocation and strand separation by hepatitis C virus helicase. *Nat Struct Mol Biol.* 2005; 12:429–435. [PubMed: 15806107]
15. Myong S, Bruno MM, Pyle AM, Ha T. Spring-loaded mechanism of DNA unwinding by hepatitis C virus NS3 helicase. *Science.* 2007; 317:513–516. [PubMed: 17656723]
16. Tai CL, Pan WC, Liaw SH, Yang UC, Hwang LH, Chen DS. Structure-based mutational analysis of the hepatitis C virus NS3 helicase. *J Virol.* 2001; 75:8289–8297. [PubMed: 11483774]
17. Zhang C, Cai Z, Kim YC, Kumar R, Yuan F, Shi PY, Kao C, Luo G. Stimulation of hepatitis C virus (HCV) nonstructural protein 3 (NS3) helicase activity by the NS3 protease domain and by HCV RNA-dependent RNA polymerase. *J Virol.* 2005; 79:8687–8697. [PubMed: 15994762]
18. Failla C, Tomei L, De Francesco R. Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus nonstructural proteins. *J Virol.* 1994; 68:3753–3760. [PubMed: 8189513]

19. Bartenschlager R, Ahlborn-Laake L, Mous J, Jacobsen H. Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions. *J Virol.* 1993; 67:3835–3844. [PubMed: 8389908]
20. Jin L, Peterson DL. Expression, isolation, and characterization of the hepatitis C virus ATPase/RNA helicase. *Arch Biochem Biophys.* 1995; 323:47–53. [PubMed: 7487072]
21. Suzich JA, Tamura JK, Palmer-Hill F, Warren P, Grakoui A, Rice CM, Feinstone SM, Collett MS. Hepatitis C virus NS3 protein polynucleotide-stimulated nucleoside triphosphatase and comparison with the related pestivirus and flavivirus enzymes. *J Virol.* 1993; 67:6152–6158. [PubMed: 8396675]
22. Gwack Y, Kim DW, Han JH, Choe J. Characterization of RNA binding activity and RNA helicase activity of the hepatitis C virus NS3 protein. *Biochem Biophys Res Commun.* 1996; 225:654–659. [PubMed: 8753814]
23. Kanai A, Tanabe K, Kohara M. Poly(U) binding activity of hepatitis C virus NS3 protein, a putative RNA helicase. *FEBS Lett.* 1995; 376:221–224. [PubMed: 7498546]
24. Kim DW, Gwack Y, Han JH, Choe J. C-terminal domain of the hepatitis C virus NS3 protein contains an RNA helicase activity. *Biochem Biophys Res Commun.* 1995; 215:160–166. [PubMed: 7575585]
25. Kim JL, Morgenstern KA, Griffith JP, Dwyer MD, Thomson JA, Murcko MA, Lin C, Caron PR. Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure.* 1998; 6:89–100. [PubMed: 9493270]
26. Yao N, Reichert P, Taremi SS, Prosser WW, Weber PC. Molecular views of viral polypeptide processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase. *Structure.* 1999; 7:1353–1363. [PubMed: 10574797]
27. Yao N, Hesson T, Cable M, Hong Z, Kwong AD, Le HV, Weber PC. Structure of the hepatitis C virus RNA helicase domain. *Nat Struct Biol.* 1997; 4:463–467. [PubMed: 9187654]
28. Cho HS, Ha NC, Kang LW, Chung KM, Back SH, Jang SK, Oh BH. Crystal structure of RNA helicase from genotype 1b hepatitis C virus. A feasible mechanism of unwinding duplex RNA. *J Biol Chem.* 1998; 273:15045–15052. [PubMed: 9614113]
29. Mackintosh SG, Lu JZ, Jordan JB, Harrison MK, Sikora B, Sharma SD, Cameron CE, Raney KD, Sakon J. Structural and biological identification of residues on the surface of NS3 helicase required for optimal replication of the hepatitis C virus. *J Biol Chem.* 2006; 281:3528–3535. [PubMed: 16306038]
30. Gu M, Rice CM. Three conformational snapshots of the hepatitis C virus NS3 helicase reveal a ratchet translocation mechanism. *Proc Natl Acad Sci U S A.* 2010; 107:521–528. [PubMed: 20080715]
31. Saikrishnan K, Powell B, Cook NJ, Webb MR, Wigley DB. Mechanistic basis of 5'-3' translocation in SF1B helicases. *Cell.* 2009; 137:849–859. [PubMed: 19490894]
32. Lee JY, Yang W. UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. *Cell.* 2006; 127:1349–1360. [PubMed: 17190599]
33. Velankar SS, Soultanas P, Dillingham MS, Subramanya HS, Wigley DB. Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell.* 1999; 97:75–84. [PubMed: 10199404]
34. Luo D, Xu T, Watson RP, Scherer-Becker D, Sampath A, Jahnke W, Yeong SS, Wang CH, Lim SP, Strongin A, Vasudevan SG, Lescar J. Insights into RNA unwinding and ATP hydrolysis by the flavivirus NS3 protein. *EMBO J.* 2008; 27:3209–3219. [PubMed: 19008861]
35. Frick DN. The hepatitis C virus NS3 protein: a model RNA helicase and potential drug target. *Curr Issues Mol Biol.* 2007; 9:1–20. [PubMed: 17263143]
36. Pang PS, Jankowsky E, Planet PJ, Pyle AM. The hepatitis C viral NS3 protein is a processive DNA helicase with cofactor enhanced RNA unwinding. *EMBO J.* 2002; 21:1168–1176. [PubMed: 11867545]
37. Fisher AJ, Smith CA, Thoden JB, Smith R, Sutoh K, Holden HM, Rayment I. X-ray structures of the myosin motor domain of Dictyostelium discoideum complexed with MgADP.BeFx and MgADP.AIF4. *Biochemistry.* 1995; 34:8960–8972. [PubMed: 7619795]

38. Kagawa R, Montgomery MG, Braig K, Leslie AG, Walker JE. The structure of bovine F1-ATPase inhibited by ADP and beryllium fluoride. *EMBO J.* 2004; 23:2734–2744. [PubMed: 15229653]
39. Liu F, Putnam A, Jankowsky E. ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. *Proc Natl Acad Sci U S A.* 2008; 105:20209–20214. [PubMed: 19088201]
40. Kwong AD, Kim JL, Lin C. Structure and function of hepatitis C virus NS3 helicase. *Curr Top Microbiol Immunol.* 2000; 242:171–196. [PubMed: 10592661]
41. Kim DW, Kim J, Gwack Y, Han JH, Choe J. Mutational analysis of the hepatitis C virus RNA helicase. *J Virol.* 1997; 71:9400–9409. [PubMed: 9371600]
42. Wardell AD, Errington W, Ciaramella G, Merson J, McGarvey MJ. Characterization and mutational analysis of the helicase and NTPase activities of hepatitis C virus full-length NS3 protein. *J Gen Virol.* 1999; 80(Pt 3):701–709. [PubMed: 10092010]
43. Hopfner KP, Michaelis J. Mechanisms of nucleic acid translocases: lessons from structural biology and single-molecule biophysics. *Curr Opin Struct Biol.* 2007; 17:87–95. [PubMed: 17157498]
44. Lin C, Kim JL. Structure-based mutagenesis study of hepatitis C virus NS3 helicase. *J Virol.* 1999; 73:8798–8807. [PubMed: 10482634]
45. Amaratunga M, Lohman TM. Escherichia coli rep helicase unwinds DNA by an active mechanism. *Biochemistry.* 1993; 32:6815–6820. [PubMed: 8392863]
46. Geiselmann J, Wang Y, Seifried SE, von Hippel PH. A physical model for the translocation and helicase activities of Escherichia coli transcription termination protein Rho. *Proc Natl Acad Sci U S A.* 1993; 90:7754–7758. [PubMed: 7689228]
47. Wincott F, DiRenzo A, Shaffer C, Grimm S, Tracz D, Workman C, Sweedler D, Gonzalez C, Scaringe S, Usman N. Synthesis, deprotection, analysis and purification of RNA and ribozymes. *Nucleic Acids Res.* 1995; 23:2677–2684. [PubMed: 7544462]
48. Song Q, Jones RA. Use of silyl ethers as fluoride scavengers in RNA synthesis. *Tetrahedron Lett.* 1999; 40:4653–4654.
49. Pyle AM, Green JB. Building a kinetic framework for group II intron ribozyme activity: quantitation of interdomain binding and reaction rate. *Biochemistry.* 1994; 33:2716–2725. [PubMed: 8117737]
50. Otwinowski Z, Minor W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology.* 1997; 276:307–326.
51. Kissinger CR, Gehlhaar DK, Smith BA, Bouzida D. Molecular replacement by evolutionary search. *Acta Crystallogr D Biol Crystallogr.* 2001; 57:1474–1479. [PubMed: 11567162]
52. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 2004; 60:2126–2132. [PubMed: 15572765]
53. Alani E, Lee JY, Schofield MJ, Kijas AW, Hsieh P, Yang W. Crystal structure and biochemical analysis of the MutS.ADP.beryllium fluoride complex suggests a conserved mechanism for ATP interactions in mismatch repair. *J Biol Chem.* 2003; 278:16088–16094. [PubMed: 12582174]
54. Hayward S, Berendsen HJ. Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins.* 1998; 30:144–154. [PubMed: 9489922]

Fig. 1. Structure of Full-length HCV NS3 bound to RNA. **(a)** Cartoon representation of NS3 viewed from the side. NS3hel domains 1 (D1), domain 2 (D2), and domain 3 (D3) are colored yellow, green, and orange, respectively. NS3pro (gray) with N-terminally-fused NS4A peptide (red) lies beneath NS3hel in this orientation. N and C termini and ssRNA are labeled accordingly. **(b)** A top view of NS3 (colored as above). NS3pro is removed for clarity. The bound sulfate ion is shown as a space-filling model, while the 5' and 3'-ends of the ssRNA (stick model) are labeled accordingly. **(c)** Conserved SF2 helicase motifs mapped onto NS3. Motif I (residues 204–211), motif Ia (residues 230–235), motif Ib (residues 268–272), motif II (residues 290–293), motif III (residues 322–324), motif IV (residues 365–

372), motif V (residues 411–419), and motif VI (460–467) are colored according to the legend. Tyr 241 (motif Y) is shown as a stick model.

Fig. 2. Electron density maps for selected ligands. Regions of the 2Fo-Fc maps (light blue) are shown for bound RNA in (a) the NS3-rU8Br4 structure (2.0Å) and (b) the NS3rU8Br4/ADP•BeF₃ structure (2.3Å). Fo-Fc density calculated prior to adding bromine to U4 nucleotide residues in both models is contoured at 4σ and shown in red. (c) A 2Fo-Fc density map (light blue) for ADP•BeF₃ bound to NS3 (2.05Å) is superimposed on the final model of the nucleotide, magnesium ion, and two coordinating water molecules.

Fig. 3. HCV NS3 binding to single-stranded nucleic acid. **(a)** Close-up view of ssRNA bound to NS3. RNA is shown as sticks colored by atom type (carbon, cyan; nitrogen, dark blue; oxygen, red; phosphate, orange; bromine, burgundy). NS3 is shown as a transparent surface and colored by domain (as in Fig. 1). The side chain or backbone atoms of NS3 residues interacting with the RNA are represented by sticks and are labeled accordingly. The side chains of Lys 371 and Arg 393 are disordered in the structure and omitted here for clarity. Nucleotide residues (U3–U7) and the 5' and 3'-ends of the strand are labeled. **(b)** Overlay of the NS3-rU8Br4 and NS3hel-ssDNA (**PDB ID: 3KQH**) structures. RNA is shown as above, while the DNA is shown with carbons colored yellow. Protein residues in motif V of NS3-

rU8Br4 (cyan) and NS3hel-ssDNA (yellow) are shown as cartoon coils. Leu 414 of NS3hel-ssDNA and the side chain atoms for Thr 416 and Trp 501 in both structures are represented by sticks. Hydrogen bonds between protein atoms and the phosphodiester backbone linking nucleotides 4 and 5 are represented by red dashed lines. Distances between corresponding phosphoryl oxygens of DNA and RNA are indicated by dashed lines (black).

Fig. 4.

(a) Nucleotide induced structural rearrangement of NS3. The structure of nucleotide-free NS3-rU8Br4 (left) compared to NS3-rU8Br4/ADP•BeF₃ (right). NS3 represented by cartoon ribbons with transparent surfaces while RNA is shown as sticks. Upon binding to ADP•BeF₃ (shown as spheres), D1 (yellow) and D2 (green) pivot towards each other, collapsing around the nucleotide. Arrows indicate the distance between the C α carbons of residues Lys 244 (D1) and Gly 468 (D2) before and after ADP•BeF₃ binding. (b) Close-up of the NS3 ATPase site. Specific interactions between ADP•BeF₃ and the two NS3 domains, D1 (right panel) and D2 (left panel), are diagramed separately for clarity. D1 residues (yellow carbons) in direct contact with either ADP•BeF₃ or the Mg²⁺ ion are shown as sticks and labeled accordingly. D2 residues (green carbons) interacting with the nucleotide analogue are also labeled. Gln 460 from D2, which coordinates a water molecule near the γ -position of the nucleotide, is also shown.

Fig. 5. Nucleotide induced conformational changes in NS3 lead to differences in RNA binding. **(a)** Structural overlay of nucleotide-free NS3-rU8Br4 (colored according to Fig. 1) and NS3-rU8Br4/ADP•BeF₃ (gray). **(b)** Local conformational change in motif V of D2. Motif V (residues 411–419) in the nucleotide-free structure (green helical coil and sticks) rearranges upon binding to ADP•BeF₃ (spheres), resulting in the repositioning of Thr 416. For clarity, only U5 of the RNA is shown (cyan sticks, nucleotide-free structure; gray sticks, nucleotide-bound structure). **(c)** Comparison of RNA binding to NS3 in the nucleotide-bound and nucleotide-free states. The nucleotide-bound structure (+Nuc, gray coil and sticks) and the

nucleotide-free structure (–Nuc, green coils and cyan sticks) are overlaid. Thr 416, Trp 501 and RNA nucleotides are labeled accordingly.

Fig. 6. Schematic representation of nucleotide induced changes in RNA binding during a single round of ATP hydrolysis. The colored boxes represent D1 (yellow), D2 (green), and D3 (orange). Specific NS3 residues and atoms involved in binding RNA are labeled and shown to interact with either the uridine bases (light orange hexagons) or the phosphodiester backbone (red circles) of the RNA strand. Ribose sugars (gray pentagons) are numbered from 5'→3' based on the position of the bromine-labeled U4 in both structures (labeled with small purple hexagons). The straight arrow indicates the relative movement of NS3 with respect to RNA.

Fig. 7. Schematic representation of NS3 stepping along an RNA tracking strand during multiple rounds of the ATPase cycle. Domains 1, 2, and 3 are colored yellow, green and orange respectively. The phosphodiester backbone of the tracking strand is shown as small circles labeled by nucleotide residue number (P7-P2, 5'→3'). The small, upside down triangles represent conserved Thr 269 of D1 (yellow) and Thr 411 of D2 (green). Three rounds of ATP binding and release are depicted (labeled 1–3).

Table 1

Crystallographic Data and Refinement Statistics

	Apo NS3	NS3-rU8Br4	NS3-rU8/ ADP•BeF ₃	NS3-rU8Br4/ ADP•BeF ₃
Data				
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell parameters (Å)	91.3, 112.0, 139.9	91.7, 110.2, 142.9	89.8, 110.4, 144.4	91.0, 110.8, 142.0
Resolution range (Å)	50.0 – 1.95	50.0 – 2.0	50.0 – 2.05	50.0 – 2.3
Total reflections	615,872	549,461	458,798	304,871
Unique Reflections	102,574	97,089	90,501	64,105
Completeness ^a (%)	98 (95)	99 (94)	99 (99)	99 (99)
R _{sym} ^b (%)	6.7 (48.6)	4.8 (46.4)	6.5 (42.7)	4.8 (39.2)
<I>/<s(I)>	21.8 (2.6)	29.3 (3.1)	22.6 (3.0)	29.6 (4.3)
Model				
Molecules in asymmetric unit	2	2	2	2
Number of water molecules	991	824	846	305
Other ligands	4 SO ₄ ²⁻	2 SO ₄ ²⁻ , 1 RNA	2 Mg ²⁺ 2 ADP•BeF ₃	1 RNA, 1 Mg ²⁺ 1 ADP•BeF ₃
Resolution used for refinement (Å)	50.0 – 1.95	50.0 – 2.0	50.0 – 2.05	50.0 – 2.3
Sigma cutoff (F/s(F))	0.0	0.0	0.0	0.0
R-factor	18.4%	20.0%	19.1%	21.6%
R-free	22.5%	23.2%	23.3%	26.3%
RMS deviations from ideal geometry:				
Bonds (Å)	0.007	0.008	0.007	0.008
Angles (°)	1.04	1.08	1.14	1.15

^aThe numbers in parentheses are values for the highest resolution shell.

^b $R_{\text{sym}} = \frac{\sum_{\text{hkl}} \sum_i |I_{\text{hkl},i} - \langle I_{\text{hkl}} \rangle|}{\sum_{\text{hkl}} \langle I_{\text{hkl}} \rangle}$