Two Novel Conserved Motifs in the Hepatitis C Virus NS3 Protein Critical for Helicase Action*

Angela M. I. Lam, David Keeney, and David N. Frick‡

From the Department of Biochemistry and Molecular Biology, New York Medical College, Valhalla, New York 10595

The hepatitis C virus (HCV) NS3 helicase shares several conserved motifs with other superfamily 2 (SF2) helicases. Besides these sequences, several additional conserved motifs are conserved among the various HCV genotypes and quasispecies. The roles of two such motifs are examined here. The first motif (YRGXDV) forms a loop that connects SF2 helicase motifs 4 and 5, at the tip of which is Arg<sup>393</sup>. When Arg<sup>393</sup> is changed to Ala, the resulting protein (R393A) retains a nucleic acid-stimulated ATPase but cannot unwind RNA. R393A also unwinds DNA more slowly than wild type and translocates poorly on single-stranded DNA (ssDNA). DNA and RNA stimulate ATP hydrolysis catalyzed by R393A like the wild type, but the mutant protein binds ssDNA more weakly both in the presence and absence of the non-hydrolyzable ATP analog ADP(BeF<sub>3</sub>). The second motif (DFSLDPFII) forms a loop that connects two anti-parallel sheets between SF2 motifs 5 and 6. When Phe<sup>444</sup> in this Phe-loop is changed to Ala, the resulting protein (F444A) is devoid of all activities. When Phe<sup>438</sup> is changed to Ala, the protein (F438A) retains nucleic acid-stimulated ATPase, but does not unwind RNA. F438A unwinds DNA and translocates on ssDNA at about half the rate of the wild type. Equilibrium binding data reveal that this uncoupling of ATP hydrolysis and unwinding is due to the fact that the F438A mutant does not release ssDNA upon ATP binding like the wild type. A model is presented explaining the roles of the Arg-clamp and the Phe-loop in the unwinding reaction.

The hepatitis C virus (HCV)<sup>1</sup> has caused a global crisis by infecting more than 2% of the world’s population (1). Combination therapies using interferon and the antiviral drug ribavirin eliminate HCV in a majority of patients but such treatments are costly and cause debilitating side effects. Vaccine and drug development has been difficult because the only natural hosts are humans and chimpanzees, and the virus cannot be conventionally cultivated in vitro. As a result, the enzymes encoded by the virus have been intensely studied for rational drug design (2). Several groups have reported potent compounds that inhibit the HCV RNA-dependent RNA polymerase (3) and its serine protease (4). However, few inhibitors of the HCV helicase have been explored as antivirals because less is known about the mechanism of action of the HCV helicase and if or how it differs from similar cellular enzymes.

HCV helicase is part of the viral non-structural protein 3 (NS3). The N-terminal one-third of NS3 is a chymotrypsin-like serine protease, and the C-terminal NS3 portion is a helicase that is fueled by the hydrolysis of ATP. Although the HCV helicase only unwinds duplex RNA during viral replication, the purified protein unwinds DNA dramatically more efficiently than RNA (5). In these reactions, HCV helicase requires a 3′-single-stranded (ss) DNA or RNA overhang to initiate unwinding and is therefore classified as a 3′-5′ helicase. The structure of the HCV helicase has been determined four times using x-ray crystallography: twice as an apoenzyme (6, 7), once with a bound ssDNA (8) and once as part of the full-length NS3 protein (9). NMR structures of parts of the HCV helicase have also been reported (10, 11). HCV helicase consists of 3 domains, with ATP likely binding between the structurally similar domains 1 and 2. One strand of the unwound duplex binds in the cleft that separates domain 3 from domains 1 and 2 (8). Gorbalenya and Koonin (12) have classified the HCV helicase as a member of helicase superfamily 2, which contains several large protein families that all share several signature sequences. The SF2 helicase signature sequence motifs line the “ATP-binding” cleft between domains 1 and 2. Numerous studies have elucidated the roles of the SF2 motifs in HCV helicase function, and structure-based mutagenesis has revealed a few other critical conserved residues (reviewed in Refs. 2 and 13). Little is known about the function of residues outside these conserved motifs.

In the present study, sequence alignments and structure-based site-directed mutagenesis were used to identify two additional conserved motifs in domain 2 of the HCV helicase that are critical for helicase action. After introducing individual point mutations into these regions of the enzyme, ATPase, unwinding, translocation rates, and DNA binding were analyzed. The first conserved motif examined is part of a loop connecting SF2 helicase motifs 4 and 5. At the tip of this loop is Arg<sup>393</sup>, which contacts the backbone of DNA (8). To test the importance of Arg<sup>393</sup>, the residue was changed to Ala, and the R393A mutant was purified and characterized. Data show how this “Arg-clamp” motif aids the unwinding of the double helix. The second conserved motif spans two conserved phenylalanines and connects a pair of β-sheets extending from domain 2. This Phe-loop motif is far from the known ssDNA binding site and the putative ATP binding site. To examine the role of the Phe-loop, two conserved Phe in this motif were changed to Ala. Analyses of the purified mutants reveal that the Phe-loop

---

* This work was supported by the AASLD Liver Scholar Award from the American Liver Foundation and National Institutes of Health Grant AI052395. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be identified in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Biochemistry and Molecular Biology, New York Medical College, Valhalla, NY 10595. Tel.: 914-594-4190; Fax: 914-594-4058; E-mail: David_Frick@NYMC.edu.

1 The abbreviations used are: HCV, hepatitis C virus; NS3, nonstructural protein 3; NS5B, nonstructural protein 5B; eIF-4a, eukaryotic translation initiation factor 4a; ss, single-stranded; VAST, Vector Alignment Search Tool; NCBI, National Center for Biotechnology Information; PDB, Protein Data Bank; MOPS, 4-morpholinepropanesulfonic acid.
might help stabilize the rotation of domain 2 by anchoring it to domain 3. Such a conformational change can be visualized by comparing currently available structures of HCV helicase and could occur upon ATP binding or hydrolysis. A model is presented to explain the role of the Arg-clamp and the Phe-loop in the helicase mechanism that combines the domain rotation model proposed by Yao et al. (6) and the ratcheting inchworm model proposed by Kim et al. (8).

EXPERIMENTAL PROCEDURES

Materials—RNase-free reagents were purchased from Ambion Inc. (Austin, TX), and RNA oligonucleotides were from Dharmacon (Lafayette, CO). DNA oligonucleotides were obtained from either Integrated DNA Technologies (Corvalle, IA) or Genelinck (Hawthorne, NY). Plasmid pET24a and BL21(DE3) cells were obtained from Novagen (Madison, WI). Beryllium fluoride was from Alfa Aesar (Ward Hill, MA).

Mutagenesis—The QuikChange site-directed mutagenesis kit (Strategene, La Jolla, CA) was used to construct each mutant from the plasmid p24Hel-1a (14), which contains the helicase region of NS3 isolated from a infectious clone of HCV genotype 1a H77c (15). The following oligonucleotides were utilized in this PCR-based mutagenesis: 5'-CATGCGGCCCGCTTACGGAAC-GCTGCTGTCTG-3'; R933A(-), 5'-GACGACAGCTGTCGACGGCCGGAGCAC-GTAGGACCCGAGTCTGGG-T3'; F438A(+), 5'-GTCGTCACTGACGTCGATGCGACGGCCTGATCTACCTACCGCTACCATTGA-3'; F444A(-), 5'-CTCATGGTTAAGAGTGGTACGGCTGACTGGCCGAGTG-3'; and F444A(+), 5'-GACGCTCTGACCTACCCGCTACCTAGG-GAGCAGC(CA3); and F444A(-), 5'-CTCTGGTTGCTAAATGTACGCGTGCTGAGGTGGTCAAGGCTG-3'. Mutations were confirmed by sequencing the DNA plasmid. All constructs place the HCV helicase under control of a T7(lac) promoter, which contains a lac operator between the T7 promoter and the transcription start site (16). Use of this hybrid promoter was necessary to control basal expression of the HCV helicase, which appears to be toxic to Escherichia coli. Each of the four recombinant proteins possesses an additional MAS sequence at their N termini and a C-terminal His tag (PNSSSVDKLAAALEHHHHHHHH). BL21(DE3) cells harboring plasmids were grown in LB broth, induced with 1,000 nM isopropyl-1-thio-1-p-galactopyranoside, and grown for 3 h at 23°C. Protein Purification—All proteins were purified as previously described (14). Briefly, cells were harvested and lysed by two passes through a French Press, and the clarified soluble extract was fractionated on three columns: a 10-ml Ni-NTA column, a 100-ml Sephacryl S-200HR gel filtration column, and finally a 0.5 ml DEAE column. All constructs yielded 1-2 mg of pure helicase/liter culture except for those containing the F444A mutant. Only a small amount of the F444A mutant was present in the soluble fraction of a cellular crude extract. In the overwhelming majority of F444A was found in inclusion bodies, which were discarded. About 50 μg of F444A was purified from a 2-l culture.

ATPase Assays—Initial rates of ATP hydrolysis were measured as described previously (14). Reactions were run at 37°C in 25 mM MOPS (pH 6.5), 5 mM MgCl2, 0.1% Tween 20, 100 nM helicase. Data were fit to the Michaelis-Menten equation by non-linear regression described previously (14). Reactions were run at 37°C, but DNA unwinding assays were performed at 37°C because rates of DNA unwinding at 37°C were too rapid to be accurately measured.

Translocation Assays—Translocation on ssDNA was assessed using an assay, developed by Raney and co-workers (17, 18), that measures the helicase catalyzed displacement of streptavidin from a biotinylated DNA substrate (5'-[32P]gtccagccacggccagcagtgtcggccag-CTGCGCTGTCTG-3'). The streptavidin-bound biotinylated DNA substrate 5'-[32P]GTbiovGGCAGGGGACGGCGAGGCCTTTTTTTTTTTTTTTTTTTTTTTTTT-3' (5 nt) and various concentrations of HCV helicase were incubated in reaction buffer (25 mM MOPS, pH 6.5, 3 mM MgCl2, 0.1% Tween 20, 37°C) for 30 min. Reactions were initiated by the addition of 5 mM ATP and 6 μM biotin (to trap released streptavidin) and terminated at various times with the stopping buffer used in the unwinding assays. After electrophoresis on a 15% non-denaturing polyacrylamide gel, reactions were analyzed with a Molecular Dynamics phosphorimager using ImageQuant software (Amersham Biosciences, Piscataway, NJ).

Fluorescence-based Nucleic Acid Binding Assay—The quenching of intrinsic protein fluorescence upon formation of an enzyme-nucleic acid complex was monitored as described previously (14). Aliquots of DNA were added to the helicase in 2 ml of 25 mM MOPS, pH 6.5, 3 mM MgCl2, and 0.1% Tween 20 at 23°C. Fluorescence was measured by exciting the sample at 280 nm and reading the emission at 340 nm with a Varian Cary Eclipse fluorescence spectrophotometer. After correcting data for intrinsic protein filter effects, fractional fluorescence FN0 was calculated by dividing the fluorescence values by the initial fluorescence of the protein in the absence of DNA. FN0 was fit to the total concentration of added 18-nt long poly(U) DNA oligonucleotide ([NA]T) using Equation 1, which accounts for ligand depletion.

\[
F_{N0} = \frac{\left(\frac{K_D + [\text{NA}]_T + [\text{E}]_T}{[\text{E}]_T}\right)}{\left(\frac{K_D + [\text{NA}]_T + [\text{E}]_T + 4[\text{NA}]_T[\text{E}]_T}{[\text{E}]_T}\right)}
\]

In Equation 1, \(\Delta F_{\text{MAX}}\) is the maximum change in fractional fluorescence resulting when all enzyme molecules are bound to nucleic acid. \(K_D\) is the dissociation constant, and \(\text{E}_T\) is total enzyme concentration. The procedure was repeated at three different enzyme concentrations for each enzyme/nucleotide pair. \(K_D\) was calculated by globally fitting 3 data sets to Equation 1 using Graphpad Prism Version 4.0 (San Diego, CA).

Gel Shift Assays—The same [32P]DNA duplex used in the unwinding assays (1 nM) was incubated with increasing concentrations of helicase (1–1,000 nM) in 25 mM MOPS, pH 6.5, 3 mM MgCl2, 0.1% Tween 20. After electrophoresis in a 15% Tris borate/EDTA polyacrylamide gel, the amounts of bound and free DNA were determined using a phosphorimager.

RESULTS

The helicase encoded by Hepatitis C virus has been extensively studied both as a model helicase and as a potential drug target. One problem with designing drugs that function by inhibiting viral enzymes is that the virus often rapidly evolves resistance by selecting mutations in inhibitor binding sites. Ideal inhibitors should therefore bind to critical invariant residues.

Sequence Alignments—To locate critical invariant residues in the HCV helicase, 163 NS3 sequences were obtained from the Hepatitis C Virus data base (s2as02.genes.nig.ac.jp) and aligned using the program ClustalX (version 1.83) (19). The results are shown in Fig. 1A. The consensus generated by ClustalX is shown below the sequence of the helicase from genotype 1a (isolate H77c). The NS3 protease occupies the first 180 amino acids that are not shown in Fig 1A. Such an analysis reveals that only about one-third of the residues in the helicase

Tri-ChI, pH 7.5, 20 mM EDTA, 0.5% SDS, 0.1% Nonidet P-40, 0.1% bromphenol blue, 0.1% xylene cyanole FF, 50% glycerol. 250 nM trap DNA was also added before electrophoresis to prevent re-annellation. Products were analyzed by non-denaturing polyacrylamide gel electrophoresis as described previously (14). In some assays, excess trap DNA, which consisted of the unlabeled shorter oligonucleotide, was added when the reactions were initiated. RNA unwinding assays were performed at 37°C, but DNA unwinding assays were performed at 23°C because rates of DNA unwinding at 37°C were too rapid to be accurately measured.

The Arg-Clamp and Phe-Loop in the HCV Helicase

44515
Identification of the Arg-clamp and Phe-loop in the HCV helicase.

A, consensus generated from an alignment of HCV NS3 sequences obtained from 163 different isolates is shown below the sequence of the helicase from HCV genotype 1a used in this study (isolate H77c, NCBI accession AF011751). Asterisks designate amino acids that are identical in all sequences, colons mark very conserved residues, and periods indicate conserved residues as defined by the program ClustalX (19). Locations of conserved motifs are designated above the sequence.

B, domain 2 of the HCV helicase was aligned with the analogous domain of other SF2 helicases using the Vector Alignment Search Tool (41) available at the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/Structure/VAST/vast.shtml). These structures include eIF4A from Saccharomyces cerevisiae (NCBI accession 1FUU), the DNA repair protein UvrB from Bacillus caldotenax (NCBI accession 1D9Z), and a DEAD-box.

Fig. 1. Identification of the Arg-clamp and Phe-loop in the HCV helicase. A, consensus generated from an alignment of HCV NS3 sequences obtained from 163 different isolates is shown below the sequence of the helicase from HCV genotype 1a used in this study (isolate H77c, NCBI accession AF011751). Asterisks designate amino acids that are identical in all sequences, colons mark very conserved residues, and periods indicate conserved residues as defined by the program ClustalX (19). Locations of conserved motifs are designated above the sequence. B, domain 2 of the HCV helicase was aligned with the analogous domain of other SF2 helicases using the Vector Alignment Search Tool (41) available at the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/Structure/VAST/vast.shtml). These structures include eIF4A from Saccharomyces cerevisiae (NCBI accession 1FUU), the DNA repair protein UvrB from Bacillus caldotenax (NCBI accession 1D9Z), and a DEAD-box.
are invariant in all HCV sequences reported. About one-third are conserved, and the rest of the residues are highly variable among the known HCV genotypes and quasispecies.

Many of the conserved amino acids cluster around the helicase "signature" motifs described by Gorbalenya and Koonin (12), but it should be noted that not all residues are strictly conserved even within the signature motifs. For example, Lys452 in Motif 1, which helps form the Walker-type phosphate binding "P-loop," is sometimes replaced by an Arg. Likewise, His293 in the signature "DECH-box" of motif 2 is a Tyr in HCV genotype 4a (20). Also shown on Fig. 1A are a TXGX motif that was shown to interact with bound nucleic acid (21), and Trp501, which intercalates between the nucleic acid bases acting like a bookend to aid unwinding (22–25).

Fig. 1A highlights additional conserved motifs outside the SF2 motifs. This study focuses on two of these motifs, which we have named the Arg-clamp and the Phe-loop. The Arg-clamp spans amino acids 392–397, where 5 of 6 conserved amino acids are invariant. An examination of the structure of HCV helicase bound to a oligonucleotide reveals that the side chain of Arg293 is only 2.3 Å from a bridging phosphate in the DNA backbone (8). To test if this is a functionally important contact, Arg293 was altered to an Ala using site-directed mutagenesis. The R393A mutant protein was purified and compared with the wild type (see below). The second conserved feature examined here is the motif spanning the absolutely conserved amino acids 437–444. In NS3, this region is the tip of an extended β-loop. This Phe-loop links two phenylalanines, one on each end of each β-strand that forms the β-loop. Kim et al. (8) proposed that this β-loop might function like the L45 loop that binds ssDNA in ssDNA-binding proteins. In support of this role, Lam et al. (14) has shown that a nearby residue at position 450 is involved in DNA binding and unwinding. Alternatively, this Phe-loop could play a more structural role. As noted by Yao et al. (6) the two conserved Phe in the Phe-loop pack into a pocket containing Phe438 and Trp432 of domain 3. Thus, the Phe-loop could either be involved in nucleic acid interactions or as an important contact to link domains 2 and 3. To test these ideas, Phe438 and Phe444 in the Phe-loop were changed to Ala, and the recombinant purified proteins were characterized.

Although the Arg-clamp and Phe-loop are conserved in HCV and related viruses, structural alignments of the HCV helicase and other SF2 helicases reveal that these motifs are unique to the HCV-like family of enzymes (Fig. 1B). This alignment reveals that neither the Phe-loop nor the Arg-clamp is present in representatives of other families in SF2. Alignments of the HCV helicase with 21 other human DEAD-box proteins also failed to uncover a human equivalent of these motifs (data not shown). However, as pointed out by Caruthers et al. (26), other DEAD-box proteins contain a nearby conserved Arg (Arg298 in eIF4A) that could likewise interact with bound nucleic acids.

Three site-directed mutants were constructed to examine the role of the Arg-clamp and Phe-loop: R393A, F438A, and F444A. Each mutant protein retained the same C-terminal His tag as the wild type (Hel-1a), was expressed in E. coli, and was purified to apparent homogeneity using the same protocol (Fig. 2). Here, the wild-type enzyme is the NS3 protein isolated from infectious clone H77c of HCV genotype 1a (15). The proteins lack the first 166 amino acids of NS3, which comprise the NS3 protease. Both R393A and F438A behaved like the wild type throughout expression and purification. However, F444A was far less soluble in crude extracts than the other proteins. Only a small amount of F444A (50 µg) was obtained from a 1-liter culture, whereas over 2 mg of the other proteins were purified from 1-liter cultures.

**ATPase Assays—**ATP hydrolysis fuels the action of the HCV helicase. Based on structural alignments with similar helicases that have been crystallized in the presence of ATP analogs (27), ATP likely binds and is hydrolyzed in the cleft between domains 1 and 2 of HCV helicase, far from the Arg-clamp and Phe-loop motifs. Thus, it is not surprising that the R393A and F438A substitutions have little effect on ATP hydrolysis (Fig. 3).

On the other hand, the F444A protein retains no detectable ability to hydrolyze ATP under any condition (Fig. 3A). Furthermore, F444A did not exhibit any detectable helicase activity and did not bind DNA or RNA in any assays described below. Since insufficient amounts of F444A were available for biophysical analysis, the inability that the F444A protein was not folded properly could not be ruled out. Therefore, because no general conclusions regarding the essential role of F444A can be firmly established, the negative data obtained with this protein are not shown in the assays below.

HCV helicase hydrolyzes ATP fairly rapidly even in the absence of DNA or RNA. In the absence of nucleic acids, Hel-1a hydrolyzes ATP with a turnover rate (kcat) of 0.82 ± 0.36 s⁻¹, assuming the protein functions as a monomer and is 100% active. R393A hydrolyzes ATP faster at 1.1 ± 0.2 s⁻¹, and F438A hydrolyzes somewhat more slowly at 0.54 ± 0.1 s⁻¹. The confidence intervals suggest that these differences are not significant. In the presence of a saturating amount of poly(U) RNA, ATP hydrolysis by Hel-1a is stimulated 160 fold to 132 ± 10 s⁻¹. R393A-catalyzed ATP hydrolysis is stimulated by RNA 115-fold to 127 ± 7 s⁻¹, and the F438A-catalyzed reaction is stimulated 222-fold to 120 ± 11 s⁻¹. Again, these rate constants are not significantly different (Fig. 3B).

The Kₐₐ values for ATP hydrolysis are likewise similar for reactions catalyzed by Hel-1a, R393A, and F438A (Fig. 3C).
The only minor differences between the mutants and the wild type occur in the absence of RNA, where the _K_m_ with Hel-1a is 21 ± 4 μM, and the _K_m_ with R393A is 44 ± 9 μM and the _K_m_ with F438A is 43 ± 9 μM. In the presence of RNA, there are no _K_m_ differences: Hel-1a, 920 ± 200 μM; R393A, 860 ± 150 μM; F438A, 1000 ± 270 μM.

Although the Phe-loop and the Arg-clamp are far from the putative ATP binding site, interactions with nucleic acids could conceivably influence the activation of ATP hydrolysis by various nucleic acids. For this reason, rates of ATP hydrolysis were measured in the presence of various concentrations of different RNA homopolymers (average lengths 2500 nt), and the parameter _K_RNA was determined. _K_RNA is analogous to a _K_m_ and is defined as the concentration of RNA that leads to 50% maximum ATP hydrolysis. Fig. 3D shows that the nucleic acid stimulation profiles of the mutants are unchanged. Poly(U) stimulates each protein better (lower _K_RNA) than poly(C). Poly(A) stimulates each enzyme poorly, whereas poly(G) does not stimulate any of the helicases (not shown). Poly(U) RNA stimulates ATP hydrolysis by R393A with a _K_RNA (67 ± 8 μM) about 2-fold higher than Hel-1a (35 ± 3 μM). The poly(U)-stimulated _K_RNA seen with F438A, 48 ± 3 μM, is more similar to that seen with the wild type. Similar differences between Hel-1a, R393A, and F438A are seen when the enzymes are stimulated by poly(C) or poly(A) RNA (Fig. 3D). The experiments summarized in Fig. 3D were also performed using all the DNA and RNA oligonucleotides used below in the helicase and binding assays. No differences in the stimulation of ATP hydrolysis catalyzed by the wild type or mutant proteins were noted (data not shown).

In Fig. 3, ATP hydrolysis was measured under optimal conditions for RNA unwinding (MOPS buffer, pH 6.5). It should be noted that assays performed with the wild type enzyme at physiological pH yield somewhat different results (14). At pH 7.5 (Tris-Cl buffer), all values for _K_RNA are about 100-fold higher, although the enzyme retains the same nucleic acid stimulation profile (poly(U) > poly(C) > poly(A)). No additional differences between the mutants and wild type were noted when rates of ATP hydrolysis were measured at pH 7.5 or in other buffer systems (data not shown).

**Unwinding Assays**—Considering that both R393A and F438A interact with RNA (Fig. 3D) and hydrolyze ATP as rapidly as wild type (Fig. 3B), one might reasonably expect that the mutant proteins would also unwind DNA and RNA like the wild type. However, although many point mutants in HCV helicase affect both ATPase and helicase, some disrupt the helicase function but leave the ATPase intact (21–24). Rigorous analyses of mutations that uncouple ATPase from helicase have yielded valuable insights into the mechanism of helicase action (28).

Unwinding assays were based on those described by Pang et al. (5). RNA and DNA duplexes were made from two complementary oligonucleotides. The shorter oligonucleotide is labeled with 32P, and the longer strand forms a 3′-single stranded tail necessary for helicase action. The helicase and substrate were combined, and the unwinding reaction was
initiated with ATP. After termination, unwound products were detected using non-denaturing polyacrylamide gel electrophoresis. No unwinding was observed in the absence of ATP, although such a reaction has been previously reported for HCV helicase (29). Also, when a boiled substrate was incubated under these conditions in the absence of protein, no re-annealing was observed.

When RNA is used as a helicase substrate, a minimum of 50 nM of Hel-1a is required to detect unwinding at 37 °C (Fig. 4A). The total amount of RNA unwind reaches a maximum between 100 and 250 nM enzyme. 500 nM Hel-1a unwinds RNA initially more rapidly but unwinds less total RNA possibly because of ATP depletion. Even under optimal conditions (250 nM enzyme) no RNA unwinding by either R393A or F438A can be detected (Fig. 4A). Likewise, neither mutant unwinds RNA when present in higher or lower concentrations (Fig. 4C).

As has been previously established, the HCV helicase unwinds DNA better than RNA (5, 30). DNA unwinding assays were therefore used to detect and analyze residual helicase activity in these mutants that lack RNA helicase activity. DNA unwinding assays were performed under the same conditions as the RNA unwinding assays except that the reactions were incubated at 23 °C. HCV helicase unwinds DNA so rapidly at 37 °C that rates could not be accurately measured. Under these conditions at 23 °C, both R393A and F438A retain some ability to unwind DNA (Fig. 5). Only 25 nM of the wild-type enzyme is necessary to detect DNA unwinding, which reaches a maximum at 100 nM enzyme (Fig. 5A). More than 50 nM of R393A is necessary to detect unwinding, and even in the presence of 200 nM R393A, more than half of the DNA remains annealed. A more linear response is seen in the presence of various amounts of F438A, but that protein also unwinds DNA more slowly than wild type (Fig. 5C). The initial rates (0–120 s) of DNA unwinding can be fit to protein concentration using Equation 2,

\[ V = V_{max}[E]/K_{cat} + [E] \]  

(Eq. 2)

for both Hel-1a and F438A (the data for R393A do not describe a hyperbola). Since protein exceeds substrate, the value, \( K_{cat} \), represents an apparent dissociation constant for the enzyme and DNA substrate. The \( V_{max} \) obtained for F438A (19 ± 4) is about one-third that with Hel-1a (56 ± 20), and the \( K_{cat} \) for F438A (151 ± 20) is 3-fold higher than Hel-1a (48 ± 20).

The above analysis suggests that the mutant proteins have a lower affinity for the DNA duplex than the wild type. Although all three proteins are fully associated with the DNA before the reaction is initiated, as shown below in direct binding studies, the mutants might fall from the substrate before unwinding can be detected. To test this hypothesis, a DNA trap was added to the unwinding reactions when they were initiated. The trap should act to sequester helicase that falls from the substrate as the reaction proceeds. The trap used here consists of the shorter of the two oligonucleotides that form the duplex substrate. The reactions were initiated with ATP and various trap concentrations and contained 100 nM of either Hel-1a (Fig. 6A), R393A (Fig. 6B), or F438A (Fig. 6C). Although even the highest concentrations of trap only slightly inhibited the reaction catalyzed by the wild type, as little as 50 nM trap (5× the enzyme concentration), significantly affected the unwinding rates of the two mutants (Fig. 6D).

Translocation Assays—The mechanism by which a helicase unwinds a double helix is complex and still not completely understood. Most current models suggest that a helicase moves along one strand while actively displacing the complementary strand. These models are primarily based on the fact that most helicases require a single-stranded region in order to unwind a double helix, and it is widely assumed that this single-stranded region is required for helicase loading. The polarity of the required ssDNA tail is usually indicative of the directionality of helicase movement. The HCV helicase requires a 3′ ssDNA tail and such a tail is provided on the longer strand of the helicase substrate utilized above. The results of the unwinding assays suggest that the mutant helicases might be defective in their ability to move along DNA (or RNA) in a 3′-5′ direction. To more directly analyze the movement of the helicases on ssDNA, the translocation of each protein on ssDNA was measured using an assay developed by Raney and co-workers (17, 18).

This assay measures helicase catalyzed displacement of streptavidin from a biotinylated oligonucleotide. An oligonucleotide with the same sequence as the longer strand used in the unwinding assay, which contains a biotin at its 5′-end annealed to streptavidin, is used here to investigate apparent movements of the HCV helicase on ssDNA.

Fig. 7 shows that both F438A and R393A translocate poorly along ssDNA. In these assays, biotinylated [32P]oligonucleotide bound to streptavidin is incubated with the HCV helicase. Reactions are initiated with ATP, terminated at various times, and analyzed on native polyacrylamide gels. In such gels, the streptavidin annealed substrates migrate more slowly than the displaced products. The wild type (Fig. 7A) more rapidly displaces streptavidin from the [32P]oligonucleotide than either R393A (Fig. 7B) or F438A (Fig. 7C). None of the proteins displace streptavidin bound to the same oligonucleotide with a biotin-labeled 3′-end, confirming the 3′-5′ directionality for the helicase.

When these translocation assays are performed at different helicase concentrations, it becomes apparent that the streptavidin displacement rates are dependent on protein concentration (Fig. 7D). R393A displaces streptavidin only at high enzyme concentrations (>500 nM). However, even at the highest concentration tested, R393A functions at less than one-tenth the rate of Hel-1a. For Hel-1a and F438A, initial rates of

![Image](image-url)
streptavidin displacement can be fit to the general rate Equation 3.

\[ V_0 = \frac{V_{\text{max}} [E]}{K_{\text{app}} + [E]} \]  

(Eq. 3)

In these assays, enzyme is in excess of substrate, and thus total enzyme approximately equals free enzyme. Such an analysis suggests that F438A interacts with the ssDNA with a \( K_{\text{app}} \) (530 ± 120 nM) that is 6-fold higher than that seen with Hel-1a (85 ± 17 nM). It should be noted, however, that these apparent dissociation constants, like the \( K_{\text{RNA}} \) and \( K_{\text{Hel}} \) values determined above, are not direct measurements of the affinity of the proteins for nucleic acids. They are instead complex functions dependent on translocation rates, dissociation rates, association rates, and processivity. The \( V_{\text{max}} \) for the F438A catalyzed reaction (3 ± 1.3) is less than half of Hel-1a (7.2 ± 2).

**Binding Assays**—Because DNA binds HCV helicase near Trp\(^{301}\) (8), nucleic acid binding significantly affects intrinsic protein fluorescence. Preugschat et al. (31) first used intrinsic protein fluorescence quenching to measure the affinity of HCV helicase for nucleic acids (31). Their data were later confirmed using the same techniques by Levin and Patel (32), who subsequently showed that the affinity of HCV helicase for DNA decreases substantially in the presence of the non-hydrolysable ATP analog ADP[BeF\(_3\)] (33). Thus, by monitoring changes in protein fluorescence upon titration with an oligonucleotide in the presence and absence of ADP[BeF\(_3\)], one can determine the thermodynamic parameters describing the interaction of HCV helicase at different points in the reaction cycle. This was done with Hel-1a, R393A, and F438A.

Before examining DNA binding, an interaction between ADP[BeF\(_3\)] and the HCV helicase mutants was confirmed by examining the ability of ADP[BeF\(_3\)] to inhibit ATP hydrolysis catalyzed by the mutants (Fig. 8A). BeF\(_3\) has been proposed to coordinate ADP in a position analogous to the \( \gamma \)-phosphate of ATP (34). The formation of this complex requires excess F\(^-\), which is supplied in the form of NaF. Although NaF and ADP alone inhibit ATP hydrolysis (35), the addition of beryllium fluoride leads to very potent inhibition (33). This effect is seen not only with the wild type, but also with R393A and F438A (Fig. 8A).

Fig. 8B shows a typical set of titrations that were performed using three amounts of Hel-1a. Because the enzyme concentration was comparable to the DNA concentration, the data were fit to an equation accounting for ligand depletion (Equation 1). This process was repeated for each enzyme in the presence and absence of ADP[BeF\(_3\)]. The resulting \( K_{\text{D}} \) values are compared in Fig. 8C. As shown in Fig. 8C, the apparent \( K_{\text{D}} \) for the wild-type Hel-1a is 2.4 ± 0.6 nM. This value is similar to that previously reported using HCV helicase derived from different HCV isolates (31, 32). In the absence of ADP[BeF\(_3\)], both mutants bind DNA somewhat weaker. The \( K_{\text{D}} \) for F438A is 4.1 ± 1.1 and that for R393A is 5.3 ± 0.9 nM. In the presence of
ADP[BeF₃], HCV helicase releases ssDNA, and the K_D obtained with the wild type increases 7-fold to 17 ± 2 nM. R393A also releases some of its grip on DNA and its K_D increases 5-fold to 27 ± 3 nM. On the other hand the K_D obtained with the Phe-loop mutant F438A does not increase significantly. The K_D describing the interaction of F438A with ssDNA is 5.7 ± 1.3 nM. This value is similar to that obtained with F438 in the absence of ADP[BeF₃] (4.1 ± 1.1 nM), although both values represent a significantly lower affinity for ssDNA than that seen with the wild type in the absence of the ATP analogue. Similar effects were seen when an RNA oligonucleotide was used instead of a DNA oligonucleotide (data not shown). However, as seen by others (32), RNA bound the enzyme about 5-fold weaker than DNA.

The dissociation constants determined for the HCV helicase and DNA in the presence of ADP(BeF₃) in this study are lower than those reported by Levin et al. (33). Whereas in some cases Levin et al. found a 100-fold reduction in affinity, only about a 5-fold reduction was seen here. The simplest explanation for this difference is that Levin et al. used a recombinant HCV helicase with an N-terminal fusion peptide (derived from the T7 gene 10 protein) in addition to a C-terminal His tag (see Refs. 36 and 37). The proteins here contained only a C-terminal His tag. It is possible that N-terminal fusion peptides, and likewise the N-terminal NS3 protease domain, influence helicase activity in a manner that affects these dissociation constants.

The above dissociation constants indicate that under the conditions which the unwinding assays were performed, almost all protein is associated with the DNA substrates. Thus, both mutant proteins bind nucleic acids and hydrolyze ATP similar to the wild type. These observations suggest that the mutant proteins fall from the duplex DNA substrate before unwinding occurs or can be detected. In other words, the mutants appear less stably bound to nucleic acids. Electrophoretic mobility shift analyses (gel-shift assays) were used to examine the stability of the helicase-DNA complexes. In these assays, helicase and the substrate used to measure DNA unwinding were incubated in the absence of ATP, and the complex was separated using non-denaturing polyacrylamide gel electrophoresis. The protein remaining bound during electrophoresis caused the DNA to be shifted. As seen in Fig. 9, Hel-1a, R393A, and F438A all caused a gel-shift, but dramatically different amounts of the three proteins were necessary to shift the DNA. Whereas 60 nM of Hel-1a was required to shift 50% of the DNA, 200 nM of F438A and 800 nM of R393A were necessary to achieve the same effect.

**DISCUSSION**

We have demonstrated that two sequence motifs in domain 2 of the HCV helicase are critical for DNA/RNA binding and strand separation. These regions, designated here as the Arg-clamp and Phe-loop, are conserved among all known HCV genotypes and quasispecies (Fig. 1A). Unlike other SF2 motifs, the Arg-clamp and Phe-loop are not shared with other protein families in SF2 (Fig. 1B) and thus may represent better targets
for rational drug design. The data support a mechanism for the HCV helicase outlined in Fig. 10 that is based on the rigid domain rotation model proposed by Yao et al. (6), the ratcheting inchworm model proposed by Kim et al. (8) and the Brownian motor mechanism proposed by Levin et al. (33).

Fig. 10A shows the locations of the Arg-clamp and the Phe-loop in the structure of HCV helicase with a bound DNA oligonucleotide (determined by Kim et al., Ref. 8). Also shown are the products of ATP hydrolysis catalyzed by Hel-1a, R393A, or F438A with 18 nucleotide long polyuridylate DNA oligonucleotide. Fractional fluorescence was fit to Equation 1 to obtain a $K_d$. C, titrations in B were repeated using Hel-1a, R393A, and F438A in the presence and absence of 0.1 mM ADP, 5 mM NaF, and 0.5 mM BeF$_2$.  

It was previously unclear whether the DNA bound in the structure shown in Fig. 10A represents the strand on which the helicase is translocating in a 3′ to 5′ direction or the complementary strand. The data obtained here demonstrate that the interaction between Arg$_{393}$ and the strand, which is seen in the structure, is critical for binding (Figs. 8 and 9) unwinding (Figs. 4–6) and translocation on ssDNA (Fig. 7). Combined, these findings imply that the DNA (Fig. 10A) is the strand on which
the helicase is translocating (the longer strand in the helicase substrates). Such an arrangement is depicted in Fig. 10C.

The structure in Fig. 10A provides a snapshot along the helicase reaction pathway. It likely represents a ground state where DNA is tightly bound to the enzyme. To function as a molecular motor, HCV helicase must exit this ground state and temporarily release DNA. The binding and hydrolysis of ATP provides this energy. Levin et al. (33) recently showed using the non-hydrolyzable ATP analog ADP(βF₃)₇ that when HCV helicase binds ATP, its affinity for ssDNA decreases. While ATP is bound, the protein can slide along DNA until hydrolysis occurs and the protein again clamps more tightly in a new position. Kim et al. (8) proposed that a unidirectional movement along ssDNA (or RNA) results from the closing of the cleft between domains 1 and 2 upon DNA binding. Unidirectional translocation occurs because the DNA slides through one end of the protein when the cleft opens and the other end when the cleft closes. When the cleft is open, Trp501 acts like a bookend and Trp-helix is connected to the C terminus of domain 2, a domain rotation would lead to movements of the Trp-helix. A slight movement of Trp501 would remove the bookend and allow DNA to slip past Trp501, as suggested by Kim et al. (8).

A scheme depicting how conformational changes could lead to duplex unwinding is presented in Fig. 10C. ATP binding leads to domain closure and a rotation of domain 2. The rotation is accomplished by pivoting around the linkage between domains 1 and 2 and simultaneously altering the twist of the anti-parallel β-sheets connecting the Phe-loop to the rest of domain 2, as proposed by Yao et al. (6). Because one end of the Trp-helix is connected to the C terminus of domain 2, a domain 2 rotation would lead to movements of the Trp-helix. A slight movement of Trp501 would remove the bookend and allow DNA to slip past Trp501, as suggested by Kim et al. (8).

The properties of R393A demonstrate that Arg393 clamps ssDNA (or RNA) into the known ssDNA binding site both in the presence and absence of ATP. The model in Fig. 10C proposes that these contacts allow the enzyme to actively disrupt the double helix. Contacts between Arg393 and the backbone of DNA occur both when ATP is absent (open conformation) and present (closed conformation). In the absence of ATP, Hel-1a tightly binds leading to a change in free energy of ~49 kJ/mol. R393A binds more weakly with a smaller ΔG of ~47 kJ/mol. In the presence of a non-hydrolyzable ATP analog, DNA binds more weakly to both enzymes. The ΔG for binding of the wild type to DNA in the presence of ADP[BeF₃]₇ is ~44 kJ/mol, whereas that for R393A is ~42 kJ/mol. Under both conditions R393 contributes equally to binding (about 2 kJ/mol). The gel-shift experiments (Fig. 9) reveal that these subtle changes in binding energy cause R393A to bind DNA more than 10-fold less stably. Thus, although DNA and RNA still bind R393A, and ATP hydrolysis is stimulated, the mutant enzyme probably fails from the nucleic acid during translocation before unwinding can be detected.

The model in Fig. 10C implies that interactions between Arg393 and DNA in the presence and absence of ATP lead to the active disruption of the double helix. If ATP binding leads to a domain closure, and rotation of domain 2, then one would expect that Arg393 would move away from the DNA backbone. If such a rotation removes Arg393 from contacts with DNA, then there would be no difference between the affinities of Hel-1a and R393A for DNA in the presence of ADP[BeF₃]₇. Because there is such a difference, Arg393 likely interacts with DNA when the protein is both in the open and in the closed conformation. If the DNA remains associated with Arg393, as suggested by the equilibrium binding data, then this would lead to a twisting and disruption of the double helix. Such a disruption would occur much less frequently in the R393A mutant that lacks the appropriate contacts between domain 2 and the nucleic acid backbone.

Fig. 10. Structure and function of the Arg-clamp and the Phe-loop in the HCV helicase. A, structure of HCV helicase bound to DNA as determined by Kim et al. (PDB 1AIV, Ref. 8). Domain 1 is colored red, domain 2 is blue, and domain 3 is gray. The bound DNA oligonucleotide is shown as light blue sticks and the orange/red spheres (sulfate ion) marks the putative ATP binding site. Key amino acid residues are also shown as sticks: Arg393 (orange), Phe398 (purple), Phe444 (brown), and Trp501 (blue). B, structural alignment of three HCV helicase structures from PDB 1HEI (6) (chain A (blue) and chain B (purple)) and PDB 1AIV (8) (brown). The Arg-clamp, Trp-helix, and Phe-loop are highlighted as tubes. C, cartoon depicting DNA unwinding by HCV helicase. In the absence of ATP, the helicase is tightly bound to DNA in an open conformation. Trp501 is intercalated between the bases like a bookend. In the absence of ATP, the helicase is tightly bound to DNA in an open conformation. Trp501 is intercalated between the bases like a bookend. Movements of the Trp-helix provide a pivot point for the rotation of the Trp-helix. Movements of the Trp-helix allow the protein to slide on DNA until ATP is hydrolyzed, and the protein returns to the closed conformation.
The F438A substitution also uncouples ATP hydrolysis from unwinding activity but its role in the reaction cycle is more difficult to interpret. The F438A protein retains an ATPase activity indistinguishable from the wild-type protein (Fig. 3), but translocates poorly (Fig. 7) and this leads to reduced DNA unwinding (Figs. 5 and 6) and no RNA unwinding (Fig. 4). This poor translocation correlates with weak DNA binding (Fig. 9) and the curious fact that F438A binds DNA more tightly than wild type in the presence of ADP(BeF3) (Fig. 8). We propose that the characteristics of the F438A mutant result from the fact that the hydrophobic pocket that includes Phe438 provides a pivot point for domain 2 rotations. Phe438 contacts Phe444 and residues that could act in a manner similar to the Arg-clamp and the Phe-loop, account for these differences is unclear, and this question warrants further investigation.

Acknowledgments—We thank Fred Jaffe, Ruth Gallagher, and Ryan Rypma for valuable discussions and technical assistance.

REFERENCES

1. Alter, M. J., Kruzson-Moran, D., Naiman, O. V., McQuillan, G. M., Gao, F., Moyer, L. A., Kaslow, R. A., and Margolis, H. S. (1999) *N. Engl. J. Med.* 341, 556–562.
2. Frick, D. N. (2003) *Curr. Org. Chem.* 7, in press.
3. Lueveque, V. J., and Wang, Q. M. (2002) *Cell. Mol. Life Sci.* 59, 909–919.
4. Narjes, F., Koch, U., and Stemmer, C. (2003) *Exp. Opin. Investig. Drugs* 12, 153–163.
5. Pang, P. S., Jankowsky, E., Planet, P. J., and Pyle, A. M. (2002) *EMBO J.* 21, 1168–1176.
6. Yao, N., Hesson, T., Cable, M., Hong, Z., Kwong, A. D., Le, H. V., and Weber, P. C. (1997) *Nat. Struct. Biol.* 4, 463–467.
7. Cho, H. S., Ha, N. C., Kang, L. W., Chung, K. M., Back, S. H., Jang, S. K., and Oh, B. H. (1998) *J. Biol. Chem.* 273, 15045–15052.
8. Kim, J. L., Morgenstern, K. A., Griffith, J. P., Dwyer, M. D., Thomson, J. A., Murcko, M. A., Lin, C., and Caron, P. R. (1998) *Structure* 6, 89–100.
9. Yao, N., Rechert, P., Taremi, S. S., Prusoff, W. W., and Weber, P. C. (1999) *Structure Fold Des.* 7, 1353–1363.
10. Liu, D., Wang, Y. S., Gesell, J. J., and Wyss, D. F. (2001) *J. Mol. Biol.* 314, 543–561.
11. Gesell, J. J., Liu, D., Madison, V. S., Hesson, T., Wang, Y. S., Weber, P. C., and Wyss, D. F. (2001) *Protein Eng.* 14, 573–582.
12. Gerbalenae, A. E., and Koonin, E. V. (1993) *Curr. Opin. Struct. Biol.* 3, 419–429.
13. Kwong, A. D., Kim, J. L., and Lin, C. (2000) *Curr. Top Microbiol. Immunol.* 242, 171–196.
14. Lam, A. M. I., Keeney, D., Eckert, P. Q., and Frick, D. N. (2003) *J. Virol.* 77, 3950–3961.
15. Yanagi, M., Purrell, R. H., Emerson, S. U., and Bukh, J. (1997) *Proc. Natl. Acad. Sci. U. S. A.* 84, 8738–8743.
16. Dubendorff, J. W., and Studier, F. W. (1991) *J. Mol. Biol.* 219, 45–59.
17. Morris, P. D., and Raney, K. D. (1999) *Biochemistry* 38, 5164–5171.
18. Morris, P. D., Byrd, A. K., Tackett, A. J., Cameron, C. E., Tanega, P., Ott, R., Fanning, E., and Raney, K. D. (2002) *Biochemistry* 41, 2372–2378.
19. Thompson, J. D., Gibson, T. J., Pfeiffer, F., Jeanninouf, F., and Higgins, D. G. (1997) *Nucleic Acids Res.* 25, 4876–4882.
20. Chamberlain, R. W., Adams, N., Saeed, K., Simon, P. P., and Elliott, R. M. (1997) *J. Gen. Virol.* 78, 1341–1347.
21. Lin, C., and Kim, J. L. (1999) *J. Virol.* 73, 8798–8807.
22. Preuschat, F., Danger, D. P., Carter, L. H., and Davis, R. G., and Porter, D. J. (2000) *Biochemistry* 39, 5174–5181.
23. Basilio, C., Lahm, A., De Francesco, R., and Gallinari, P. (2000) *J. Gen. Virol.* 81, 1649–1658.
24. Tai, C. H., Pan, W. C., Liaw, S. H., Yang, U. C., Hwang, H. L., and Chen, D. S. (2001) *J. Virol.* 75, 8289–8297.
25. Kim, J. W., Seo, M. Y., Shelat, A., Kim, C. S., Kwon, T. W., Lo, H. H., Moustakas, D. T., Sun, J., and Han, J. H. (2003) *J. Virol.* 77, 571–582.
26. Caruthers, J. M., Johnson, E. R., and McKay, D. B. (2000) *Proc. Natl. Acad. Sci. U. S. A.* 97, 13080–13085.
27. Korelev, S., Yao, N., Lehman, T. M., Lehman, W. P., and Waksman, G. (1998) *Protein Sci.* 7, 605–610.
28. Soutulanas, P., Dillingham, M. S., Wiley, P., Webb, M. R., and Wignall, D. B. (2000) *EMBO J.* 19, 3799–3810.
29. Porter, D. J., and Preuschat, F. (2000) *Biochemistry* 39, 5166–5173.
30. Gwack, Y., Kim, D. W., Han, J. H., and Choe, J. (1997) *Eur. J. Biochem.* 250, 47–54.
31. Preuschat, F., Averett, D. R., Clarke, B. E., and Porter, D. J. T. (1996) *J. Biol. Chem.* 271, 24449–24557.
32. Levin, M., and Patel, S. S. (2002) *J. Biol. Chem.* 277, 29377–29385.
33. Levin, M. K., Gurus, M. M., and Patel, S. S. (2003) *J. Biol. Chem.* 278, 23311–23316.
34. Penomarev, M. A., Timofeev, V. P., and Levittsky, D. I. (1995) *FEBS Lett.* 371, 261–263.
35. Porter, D. J. (1998) *J. Biol. Chem.* 273, 7390–7396.
36. Kim, D. W., Gwack, Y., Han, J. H., and Choe, J. (1995) *Biochem. Biophys. Res. Commun.* 215, 169–166.
37. Levin, M., and Patel, S. S. (1999) *J. Biol. Chem.* 274, 31839–31846.
38. Velankar, S. S., Soutulanas, P., Dillingham, M. S., Subramanaya, H. S., and Wignall, D. B. (1999) *Cell* 97, 75–84.
39. Korelev, S., Haeh, J., Gaus, G. H., Lehman, T. M., and Waksman, G. (1997) *Cell* 90, 635–647.
40. Du, M. X., Johnson, R. B., Sun, X. L., Staicu, K. A., Colaccino, J., and Wang, Q. M. (2002) *Biochem. J.* 363, 147–155.
41. Gibiat, J. F., Madej, T., and Bryant, S. H. (1996) *Curr. Opin. Struct. Biol.* 6, 377–385.