4604–4614 *Nucleic Acids Research, 2012, Vol. 40, No. 10* *Published online 28 January 2012*
doi:10.1093/nar/gks046

Mycobacterium smegmatis RqlH defines a novel clade of bacterial RecQ-like DNA helicases with ATP-dependent 3′–5′ translocase and duplex unwinding activities

Heather Ordonez, Mihaela Unciuleac and Stewart Shuman*

Molecular Biology Program, Sloan-Kettering Institute, NY 10065, USA

Received December 14, 2011; Revised January 6, 2012; Accepted January 10, 2012

ABSTRACT

The *Escherichia coli* RecQ DNA helicase participates in a pathway of DNA repair that operates in parallel to the recombination pathway driven by the multisubunit helicase–nuclease machine RecBCD. The model mycobacterium *Mycobacterium smegmatis* executes homologous recombination in the absence of its helicase–nuclease machine AdnAB, though it lacks a homolog of *E. coli* RecQ. Here, we identify and characterize *M. smegmatis* RqlH, a RecQ-like helicase with a distinctive domain structure. The 691-amino acid RqlH polypeptide consists of a RecQ-like ATPase domain (amino acids 1–346) and tetracysteine zinc-binding domain (amino acids 435–499), separated by an RqlH-specific linker. RqlH lacks the C-terminal HRDC domain found in *E. coli* RecQ. Rather, the RqlH C-domain resembles bacterial ComF proteins and includes a phosphoribosyltransferase-like module. We show that RqlH is a DNA-dependent ATPase/dATPase that translocates 3′–5′ on single-stranded DNA and has 3′–5′ helicase activity. These functions inhere to RqlH-(1–505), a monomeric motor unit comprising the ATPase, linker and zinc-binding domains. RqlH homologs are distributed widely among bacterial taxa. The mycobacteria that encode RqlH lack a classical RecQ, though many other *Actinobacteria* have both RqlH and RecQ. Whereas *E. coli* K12 encodes RecQ but lacks a homolog of RqlH, other strains of *E. coli* have both RqlH and RecQ.

INTRODUCTION

Bacterial DNA helicases are nucleic acid-dependent NTPases that play important roles in DNA replication, recombination and repair. We are interested in the recombination/repair functions of DNA helicases in *Mycobacteria* (1–8), a genus of the phylum *Actinobacteria* that includes the slow-growing human pathogen *M. tuberculosis* and its more rapidly growing avirulent relative *M. smegmatis*. Mycobacteria have three distinct pathway options for the repair of DNA double-strand breaks (DSBs): (i) RecA-dependent homologous recombination (HR); (ii) RecA-independent homology-directed single-strand annealing (SSA); and (iii) non-homologous end-joining (NHEJ) driven by Ku and DNA ligase D (9,10). The HR and SSA mechanisms involve the resection of DSB ends by the coordinated actions of a helicase and a nuclease. Mycobacteria have evolved a distinctive division of helicase labor between the two homology-dependent pathways, whereby the mycobacterial helicase–nuclease machine AdnAB is dedicated to the HR pathway while the RecBCD machine is responsible for SSA (10). AdnAB is a heterodimer of two subunits, both consisting of an N-terminal UvrD-like motor domain and a C-terminal RecB-like nuclease domain (4). The 3′–5′ translocase/helicase activity of the AdnB subunit initiates and propagates duplex unwinding from a DSB end, during which the AdnA and AdnB nucleases incise the displaced 5′- and 3′-strands, respectively (4–6). Whereas ablation of *recA* abolishes mycobacterial HR *in vivo*, *adnAB* ablation merely reduces the efficiency of HR, implying the existence of a parallel mycobacterial subpathway (not involving RecBCD) for DNA unwinding and resection (10).

To date, two other mycobacterial helicases have been identified and characterized: UvrD1 and UvrD2 (1–3,7,8).

*To whom correspondence should be addressed. Tel: +1 212 639 7145; Fax: +1 212 717 3623; Email: shuman@ski.mskcc.org

As the name implies, mycobacterial UvrD1 is homologous to *E. coli* UvrD, a prototypal superfamily I DNA helicase. Yet, despite having a vigorous DNA-dependent ATPase activity, UvrD1 is a feeble helicase *per se* (1,7). The distinctive property of UvrD1 is that it requires Ku in order to catalyze efficient ATP-dependent unwinding of 3′-tailed duplex DNAs (1). UvrD1, Ku and DNA form a stable ternary complex, suggesting that Ku (or a different protein partner) might serve as a processivity factor for unwinding by UvrD1. Ablation of *uvrD1* sensitizes *Mycobacterium smegmatis* to killing by ultraviolet and ionizing radiation and *tert*-butylhydroperoxide, thereby attesting to its function in DNA repair (1,11).

Mycobacterial UvrD2 is a DNA-dependent ATPase with a vigorous 3′–5′ duplex unwinding activity (1). UvrD2 is an atypical helicase with respect to its domain organization, insofar as its N-terminal ATPase domain resembles the superfamily I helicase UvrD, yet it has a C-terminal HRDC domain, which is a feature of RecQ-type superfamily II DNA helicases. The ATPase and HRDC domains are connected by a CxxC-(14)-CxxC tetracysteine module that defines a new clade of UvrD2-like bacterial helicases found only in *Actinomycetales* (2). Attempts to disrupt the *M. smegmatis* or *M. tuberculosis uvrD2* gene were unsuccessful unless a second copy of *uvrD2* was present elsewhere in the chromosome (2,8), indicating that UvrD2 is essential for growth of mycobacteria. The UvrD2 HRDC domain is not required for ATPase or helicase activities *in vitro*, and is dispensable for survival of *M. tuberculosis* (2,8). Deletion of the tetracysteine module of UvrD2 abolishes duplex unwinding while preserving ATP hydrolysis and mycobacterial viability (2,8). Such findings suggest that UvrD2 performs an essential motor function *in vivo* that does not require duplex DNA unwinding.

Studies of other model bacteria implicate the RecQ DNA helicase and the separate RecJ nuclease in a RecA- and RecFOR-dependent HR pathway that operates in parallel to the pathway driven by multisubunit helicase–nuclease machines. For example, *Bacillus subtilis* DSB repair relies on RecA, but loss of AddAB, the *Bacillus* motor-nuclease machine homologous to mycobacterial AdnAB (5,12) elicits only moderate sensitivity to killing by DNA damaging agents (13). The AddAB-independent DSB repair pathway in *B. subtilis* requires a RecQ-type DNA helicase and RecJ (13). In *E. coli*, the RecQ helicase and RecJ nuclease drive end resection in RecFOR-dependent DSB repair when RecBCD is absent (14).

An earlier phylogenetic survey of then available bacterial proteomes noted mycobacterial homologs of RecFOR, but failed to detect mycobacterial homologs of RecQ or RecJ (15). Here, we identify, purify, and biochemically characterize the first example of a mycobacterial **R**ec**Q**-**l**ike **h**elicase (named RqlH): the 691-amino acid protein encoded by *M. smegmatis gene* MSMEG_5935. We report that RqlH resembles *E. coli* RecQ in possessing DNA-dependent ATPase, and ATP-dependent 3′–5′ helicase activities, but is distinguished from 'classical' RecQ proteins by its unique domain composition (Figure 1). RqlH defines a new clade of superfamily II helicases distributed widely among bacterial taxa.

MATERIALS AND METHODS

RqlH proteins

The ORFs encoding full-length RqlH (MSMEG_5935) and a C-terminal truncation mutant RqlH-(1–505) were PCR-amplified from *M. smegmatis* genomic DNA with primers that introduced a BglII site immediately flanking the start codon and a HindIII site downstream of the stop codon. The PCR products were digested with BglII and HindIII and inserted between the BamHI and HindIII sites of pET28b-His_{10}Smt3 to generate expression plasmids encoding the RqlH or RqlH-(1–505) polypeptides fused to an N-terminal His_{10}Smt3 tag. Mutations K49A and D148A were introduced into the expression plasmids by PCR with mutagenic primers. All of the plasmid inserts were sequenced to verify that no unintended coding changes were acquired during amplification and cloning.

pET28b-His_{10}Smt3-RqlH and pET28b-His_{10}Smt3-RqlH-(1–505) plasmids were transformed into *E. coli* BL21(DE3) cells. Cultures (1 l) amplified from single kanamycin-resistant transformants were grown at 37°C in LB broth containing 60 μg/ml kanamycin until the A_{600} reached 0.6. The cultures were chilled on ice for 45 min, then adjusted to 2% (v/v) ethanol and 0.5 mM isopropyl-β-D-thiogalactopyranoside and incubated for 16 h at 17°C with constant shaking. All subsequent steps were performed at 4°C. Cells were harvested by centrifugation and resuspended in 25 ml of buffer A (50 mM Tris–HCl, pH 8.0, 250 mM NaCl, 1 mM DTT, 0.1% Triton X-100, 10% sucrose) containing one protease inhibitor cocktail tablet (Roche). Lysozyme was added to a concentration of 1 mg/ml. After incubation for 30 min, the lysate was sonicated to reduce viscosity and the insoluble material was removed by centrifugation at 38 000*g* for 1 h. The supernatant was mixed for 1 h with 3 ml of Ni–NTA agarose resin (Qiagen) that had been equilibrated with buffer A. The resin was recovered by centrifugation and resuspended in 30 ml of buffer A. The washed resin was then recovered by centrifugation, resuspended in 10 ml of buffer A containing 50 mM imidazole, and then poured into a column. The column was eluted stepwise with 5-ml aliquots of buffer A containing 100, 200 and 500 mM imidazole. The polypeptide compositions of the eluate fractions were monitored by SDS–PAGE. The recombinant RqlH or RqlH-(1–505) proteins eluted in the 200 and 500 mM imidazole fractions, which were pooled, supplemented with Smt3-specific protease Ulp1 (at a RqlH:Ulp1 ratio of 1000:1) and then dialyzed against 4 l of buffer B (50 mM Tris–HCl, pH 8.0, 250 mM NaCl, 1 mM DTT, 0.05% Triton X-100, 10% glycerol) for 3 h to remove the His_{10}Smt3 tag. The tag-free RqlH or RqlH-(1–505) proteins were separated from His_{10}Smt3 by applying the digests to a 3-ml Ni–NTA agarose column that had been equilibrated with buffer B and then collecting the flow-through fraction.

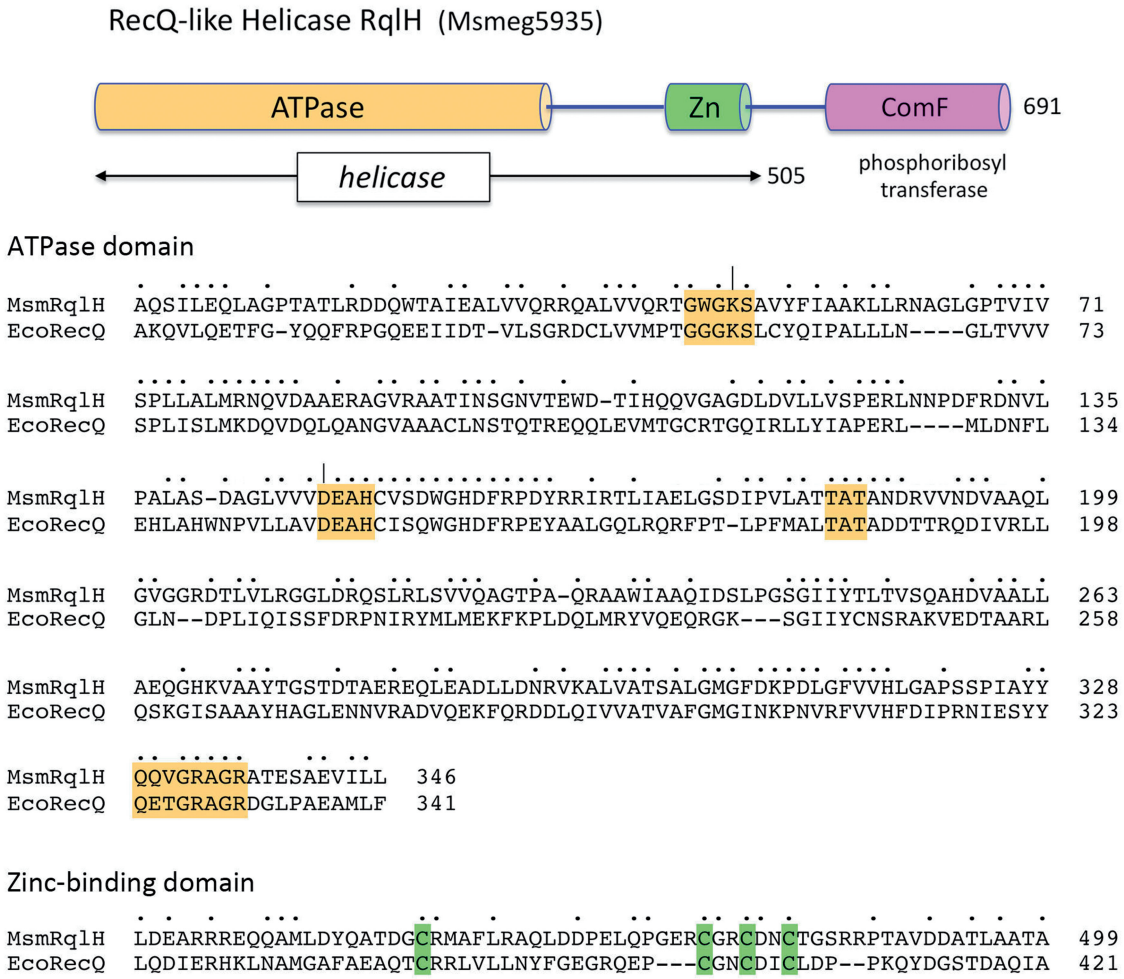

Figure 1. Domain organization of *M. smegmatis* RqlH. The 691-amino acid RqlH polypeptide is depicted in the top panel as a linear array (with the N-terminus at left and the C-terminus at right) and the known or imputed domains drawn as cylinders spanning their segments of the primary structure. RqlH is composed of an N-terminal ATPase domain (beige) flanked by a downstream tetracysteine zinc-binding domain (green) and a C-terminal ComF domain (magenta) that includes motifs found in amidophosphoribosyltransferase enzymes. The N-terminal 505-amino acid segment of RqlH that includes the ATPase and zinc-binding domains and suffices for helicase activity is indicated below the domain cartoon. The amino acid sequences of the RqlH ATPase and zinc-binding domains are aligned to the homologous domains of *E. coli* RecQ in the middle and bottom panels, respectively. Positions of side chain identity/similarity are indicated by dots above the sequences. Gaps in the alignments are denoted by dashes. The superfamily II ATPase motifs I, II, III and VI are highlighted in beige in the middle panel; the Lys49 and Asp148 residues that were mutated to alanine are denoted by |. The four putative zinc-binding cysteines are highlighted in green in the bottom panel.

In order to deplete residual nucleic acids, the tag-free RqlH and RqlH-(1–505) preparations were applied to 2-ml DEAE-Sephacel columns that had been equilibrated with buffer B. RqlH and RqlH-(1–505) were recovered in the flow through fraction, which was then adjusted to 1 mM EDTA. Protein concentrations were determined with the BioRad dye reagent using BSA as the standard. The yields of RqlH and RqlH-(1–505) were ~10 mg/l of culture.

Velocity sedimentation

Aliquots (150 µg in 0.2 ml) of RqlH or RqlH-(1–505), either alone or mixed with catalase (50 µg), BSA (50 µg), and cytochrome c (100 µg), were applied to 4.8-ml 15–30% glycerol gradients containing 50 mM Tris–HCl, pH 8.0, 250 mM NaCl, 1 mM DTT, 1 mM EDTA, 0.05% Triton X-100. The gradients were centrifuged at 50 000 rpm for 17 h at 4°C in a Beckman SW55Ti rotor. Fractions (0.2 ml) were collected from the bottoms of the tubes.

Nucleoside triphosphatase assay

Reaction mixtures containing (per 10 µl) 20 mM Tris–HCl, pH 8.0, 1 mM DTT, 5 mM $MgCl_2$, 1 mM [α^{32}P]ATP (Perkin-Elmer Life Sciences) and DNA and RqlH or RqlH-(1–505) as specified were incubated for 30 min at 37°C. The reactions were quenched by adding 2 µl of 5 M formic acid. An aliquot (2 µl) of the mixture was applied to a polyethyleneimine-cellulose TLC plate, which was developed with 0.45 M ammonium sulfate. The radiolabeled material was visualized by autoradiography and ^{32}P-ADP formation was quantified by scanning the TLC plate with a Fujix BAS-2500 imager.

Streptavidin displacement assay of RqlH translocation on DNA

Synthetic 34-mer oligodeoxynucleotides of otherwise identical nucleobase sequence containing a Biotin-ON internucleotide spacer either at the fourth position from the 5′-terminus or the second position from the 3′-terminus were purchased from Eurofins MWG Operon. These DNAs were 5′-end-labeled with [γ^{32}P]ATP by using T4 polynucleotide kinase and then purified by electrophoresis through a native 18% polyacrylamide gel. Steptavidin–DNA (SA–DNA) complexes were formed by pre-incubating 50 nM biotinylated ^{32}P-DNA with 4 μM streptavidin (SA) (Sigma) in 20 mM Tris–HCl, pH 8.0, 1 mM DTT, 5 mM $MgCl_2$ and 1 mM ATP for 10 min at room temperature. The mixtures were supplemented with 20 μM free biotin (Fisher) and the displacement reactions (10 μl, containing 0.5 pmol biotinylated ^{32}P-DNA) were initiated by adding 1 pmol (100 nM) RqlH-(1–505). After incubation for 15 min at 37°C, the reactions were quenched by adding 3 μl of a solution containing 200 mM EDTA, 0.6% SDS, 25% glycerol and 15 μM of a unlabeled single-stranded DNA (the 24-mer oligodeoxynucleotide shown in Figure 4B) to mask any binding of RqlH-(1–505) to ^{32}P-DNA released from the SA–DNA complex. The reaction products were analyzed by electrophoresis through a 15-cm native 18% polyacrylamide gel containing 89 mM Tris–borate, 2.5 mM EDTA. The free ^{32}P-labeled 34-mer DNA and the SA–DNA complexes were visualized by scanning the gel with a Fujix BAS-2500 imaging apparatus.

Helicase assay

The 5′ ^{32}P-labeled strand was prepared by reaction of a synthetic oligodeoxynucleotide with T4 polynucleotide kinase and [γ^{32}P]ATP. The labeled DNA was separated from ATP by gel-filtration through a G-50 micro column (GE Healthcare) and then annealed to a 5-fold excess of a complementary DNA strand to form the various substrates shown in the figures. The annealed DNAs were purified by electrophoresis through a native 12% polyacrylamide gel, eluted from an excised gel slice by incubation for 12 h at 4°C in 200 μl TE (10 mM Tris–HCl, pH 7.5, 1 mM EDTA), recovered by ethanol precipitation and resuspended in TE. Helicase reaction mixtures (10 μl) containing 20 mM Tris–HCl, pH 8.0, 1 mM DTT, 35 mM $MgCl_2$, 1 pmol (100 nM) radiolabeled DNA and RqlH-(1–505) as specified were preincubated for 5 min at room temperature. The reactions were initiated by adding 5 mM ATP and 10 pmol of an unlabeled oligonucleotide identical to the labeled strand of the helicase substrate. Addition of excess unlabeled strand was necessary to prevent the spontaneous reannealing of the unwound ^{32}P-labeled DNA strand. The reaction mixtures were incubated for 30 min at 37°C and then quenched by adding 3 μl of a solution containing 2% SDS, 330 mM EDTA and 10 μg proteinase K. After protease digestion for 10 min at 37°C, the mixtures were supplemented with 2 μl of 50% glycerol, 0.3% bromophenol blue. The reaction products were analyzed by electrophoresis through a 15-cm 12% polyacrylamide gel in 89 mM Tris–borate, 2.5 mM EDTA. The products were visualized by autoradiography.

RESULTS AND DISCUSSION

Distinctive domain organization of *M. smegmatis* RqlH, a putative RecQ-like helicase

The genomic *M. smegmatis* open reading frame MSMEG_5935 encodes a 691-amino acid polypeptide with a 346-amino acid N-terminal domain that resembles the N-terminal ATPase domain of the 610-amino acid *E. coli* RecQ protein, to the extent of 163 positions of amino acid side chain identity/similarity (Figure 1). In particular, the mycobacterial protein and RecQ share a set of classical superfamily II helicase motifs (highlighted in Figure 1) that comprise the RecQ phosphohydrolase active site (16). Consequently, we named the mycobacterial protein RqlH (**R**ec**Q**-**l**ike **h**elicase) in anticipation of its biochemical characterization. In addition, the RqlH segment from amino acids 435–499 is homologous to the tetracysteine zinc-binding domain (amino acids 362–421) of *E. coli* RecQ (Figure 1). The RecQ zinc-binding domain, which has a distinctive fold and spacing of its zinc-binding cysteines, is a signature feature of the extended RecQ family (16,17). However, whereas the zinc-binding module of *E. coli* RecQ directly abuts the ATPase domain, the ATPase and putative zinc-binding domains of mycobacterial RqlH are separated by a 90-amino acid spacer that has no apparent counterpart in *E. coli* RecQ. Moreover, RqlH lacks the functionally important C-terminal HRDC domain found in *E. coli* RecQ and many other RecQ family members (17–22). Rather, the C-terminal 150-amino acid segment of mycobacterial RqlH displays primary structure similarity to bacterial ComF proteins, which are implicated in bacterial competence for DNA transformation (23). ComF proteins (which are typically 200–250 amino acids in length) include a C-terminal domain module homologous to the active site of the bacterial enzyme glutamine phosphoribosylpyrophosphate amidotransferase (GPATase) (24) that catalyzes the first step in *de novo* purine biosynthesis. RqlH shares this primary structure homology to GPATase (Supplementary Figure S1), which includes several active site amino acids that bind the phosphoribosylpyrophosphate•Mn^{2+} substrate complex in the GPATase crystal structure (24) (Supplementary Figure S1).

To our knowledge, there has been no prior characterization of an enzyme with the distinctive domain composition seen in *M. smegmatis* RqlH. Key questions of interest to us were: (i) is RqlH a nucleic acid-dependent phosphohydrolase and, if so, what is its substrate and cofactor specificity? (ii) can RqlH couple NTP hydrolysis to mechanical work, especially duplex unwinding? (iii) is the unique C-terminal ComF domain required for such activities? (iv) are RqlH-like helicases found in the proteomes of other taxa?

Recombinant RqlH is a DNA-dependent ATPase

To evaluate the enzymatic and physical properties of RqlH, we produced the protein in *E. coli* as a His_{10}Smt3 fusion and isolated it from a soluble extract by nickel-agarose chromatography. The His_{10}Smt3 tag was removed with the Smt3-specific protease Ulp1 and the native RqlH was separated from the tag and residual nucleic acids by Ni–agarose and DEAE-Sephacel chromatography steps. SDS–PAGE revealed a predominant ~70 kDa polypeptide corresponding to RqlH (Figure 2A). Reaction of increasing amounts of purified RqlH with 1 mM [α^{32}P]ATP in the presence of magnesium and salmon sperm DNA resulted in progressive hydrolysis of the labeled ATP to [α^{32}P]ADP, to an extent of 96% ATP hydrolysis at saturating enzyme levels (Figure 2B). A mutated version of the RqlH protein was prepared in which Lys49 in motif I (the P-loop or Walker A-box), which contacts the β and γ phosphates of ATP in the *E. coli* RecQ crystal structure (16), was replaced by alanine. A second mutated version of RqlH was prepared in which the metal-binding Asp148 in motif II was changed to alanine. The purity of the RqlH-K49A and RqlH-D148A proteins was comparable to that of wild-type RqlH (Figure 2A). The motif I and II mutations abolished the ATPase activity of RqlH (Figure 2B). These results verify that the phosphohydrolase activity is intrinsic to the recombinant RqlH protein.

We also purified a truncated version, RqlH-(1–505), that spans the RecQ-like ATPase and zinc-binding domains, but lacks the C-terminal 186-amino acid segment that includes the ComF-like domain. SDS–PAGE revealed a predominant ~55 kDa polypeptide corresponding to RqlH-(1–505) (Figure 2A). RqlH-(1–505) catalyzed ATP hydrolysis in the presence of salmon sperm DNA; the specific activity of the truncated protein was similar to that of full-length RqlH (Figure 2B). The K49A and D148A mutations of RqlH-(1–505) abolished its ATPase activity (Figure 2B). We conclude that the C-terminal domain of RqlH is dispensable for ATP hydrolysis. We attempted to purify a more extensively truncated variant, RqlH-(1–465), that lacked the zinc-binding domain, but this protein was intractably insoluble when expressed in *E. coli* and therefore not amenable to purification and characterization.

The quaternary structure of RqlH-(1–505) was gauged by zonal velocity sedimentation through a 15–30% glycerol gradient. The RqlH-(1–505) protein was analyzed alone and, in a parallel gradient, as a mixture with marker proteins catalase (248 kDa), BSA (66 kDa) and cytochrome c (13 kDa). The RqlH-(1–505) polypeptide sedimented as a major discrete peak in fractions 17–19, with a slight shoulder on the heavy side of the peak (Figure 3A). The ATPase activity profile also peaked in fractions 17–19 and tracked with the abundance of the 55-kDa RqlH-(1–505) polypeptide (Figure 3B). RqlH-(1–505) sedimented identically in the parallel glycerol gradient with internal standards (data not shown). The sedimentation peaks of the three marker proteins are denoted by arrows on the ATPase activity profile, placing the RqlH-(1–505) peak on the 'light' side of BSA (Figure 3B). We surmise that RqlH-(1–505) is a monomer in solution. By contrast, full-length RqlH sedimented diffusely across the gradient in fractions extending from the BSA peak to the very bottom of the gradient (data not shown), suggesting that it was prone to aggregation. In light of these findings, we focused our subsequent biochemical studies on the catalytic domain RqlH-(1–505).

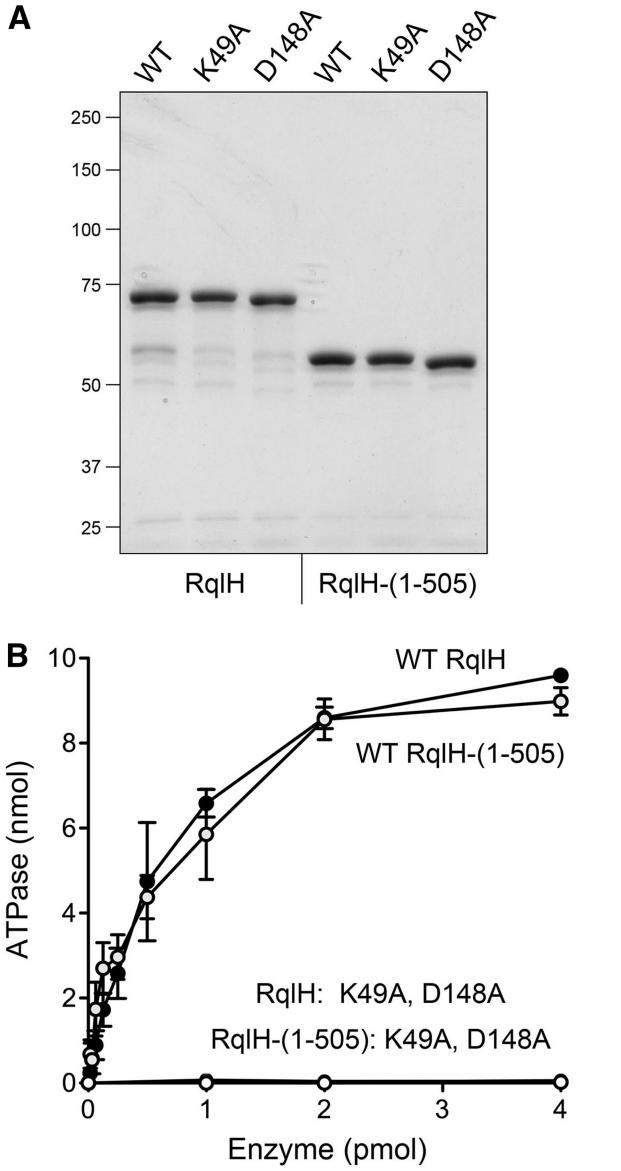

Figure 2. RqlH and RqlH-(1–505) are DNA-dependent ATPases. (**A**) Purification. Aliquots (2.5 μg) of full-length wild-type RqlH and full-length mutants RqlH-K49A and RqlH-D148A were analyzed by SDS–PAGE in parallel with aliquots (2.5 μg) of the truncated proteins RqlH-(1–505), RqlH-(1–505)-K49A and RqlH-(1–505)-D148A. The Coomassie blue-stained gel is shown. The positions and sizes (kDa) of marker are indicated on the left. (**B**) ATPase reaction mixtures containing 20 mM Tris–HCl, pH 8.0, 1 mM DTT, 5 mM $MgCl_2$, 1 mM (10 nmol) [α^{32}P]ATP, 2 μg sonicated salmon sperm DNA and RqlH or RqlH-(1–505) proteins as specified were incubated for 30 min. Reaction products were analyzed by PEI-cellulose TLC. Each datum is the average of three protein titration experiments ± SEM.

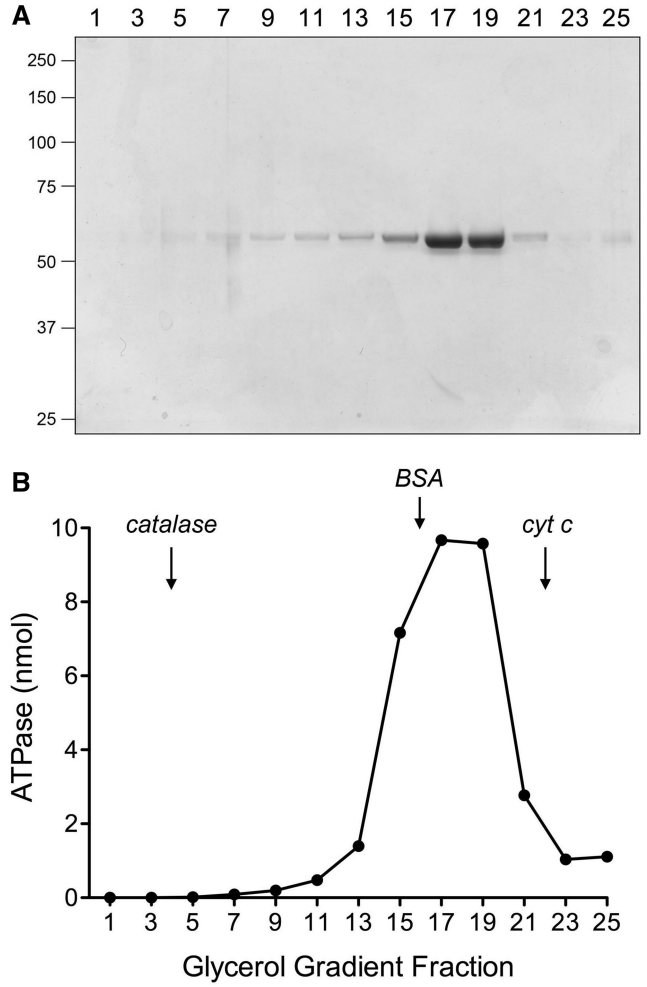

Figure 3. Glycerol gradient sedimentation of RqlH-(1–505). (**A**) RqlH-(1–505) was sedimented through a 15–30% glycerol gradient as described under Experimental Procedures. Aliquots (20 µl) of the odd numbered fractions (fraction 1 being the bottom of the gradient) were analyzed by SDS–PAGE. The Coomassie-stained gel is shown; the positions and sizes (kDa) of marker polypeptides are indicated on the left. (**B**) ATPase reaction mixtures containing 20 mM Tris–HCl, pH 8.0, 1 mM DTT, 5 mM $MgCl_2$, 10 nmol [α^{32}P]ATP, 2 µg sonicated salmon sperm DNA and 1 µl of the indicated glycerol gradient fractions were incubated at 37°C for 30 min. The activity profile is plotted. The peak positions of the internal standards catalase, BSA and cytochrome C analyzed in a parallel glycerol gradient are indicated by vertical arrows.

Metal and DNA cofactor requirements for ATPase activity

The ATPase activity of RqlH-(1–505) was optimal from pH 5.5 to 9.5 in Tris buffer; activity declined sharply at pH 5.0 and was virtually nil at pH 4.5 (Supplementary Figure S2B). No ATP hydrolysis was detected when magnesium was omitted. The divalent cation requirement was satisfied to varying degrees by 1 mM calcium, cobalt, magnesium or manganese, whereas copper, nickel and zinc were ineffective (Supplementary Figure S2A). Hydrolysis of 1 mM ATP was optimal between 2 and 7.5 mM magnesium (data not shown).

ATP hydrolysis was strictly dependent on an exogenous DNA cofactor. Whereas the nucleic acid requirement was satisfied by sonicated salmon sperm DNA, an equivalent amount of heat-denatured salmon sperm DNA resulted in a 2-fold increase in the extent of ATP hydrolysis (Figure 4A). Closed circular and linear pUC19 DNAs were relatively feeble activators of ATP hydrolysis; heat denaturation of the linear plasmid elicited an 8-fold increase in ATPase activity compared to an equivalent amount of unheated linear pUC19 (Figure 4A). These results suggest that single-stranded DNA is the preferred cofactor for the RqlH-(1–505) ATPase. This was demonstrated more directly by testing the ability of single-stranded DNAs of varying length to activate ATP hydrolysis by 0.1 µM enzyme (Fig 4B). Titration of the oligonucleotides revealed a hyperbolic dependence of ATP hydrolysis on the amount of 44-mer, 36-, 30-, 24- or 12-mer strands (Figure 4B). Non-linear regression fitting of the data to a one-site binding model in Prism yielded apparent K_d values as follows: 0.19 µM 44-mer; 0.19 µM 36-mer; 0.40 µM 30-mer; 0.37 µM 24-mer; and 0.35 µM 12-mer. In contrast, a 6-mer oligonucleotide was ineffective at up to 1.25 µM concentration (Figure 4B).

NTP substrate specificity and steady state kinetic parameters

NTP specificity was examined by colorimetric assay of the release of P_i from unlabeled ribonucleotides ATP, GTP, CTP or UTP and deoxynucleotides dATP, dGTP, dCTP and dTTP. RqlH-(1–505) displayed specificity for hydrolysis of ATP and dATP (Figure 5A). We determined steady-state kinetic parameters by measuring the velocity of ATP hydrolysis as a function of ATP concentration in the presence of 1.25 µM 44-mer single-stranded DNA cofactor (Figure 5B). From a non-linear regression curve fit of the data to the Michaelis-Menten equation, we calculated that RqlH-(1–505) has a K_m of 180 µM ATP and a k_{cat} of 17/s. The ATPase turnover number of RqlH-(1–505) is similar to values of 24/s and 35/s reported for *E. coli* RecQ (29,30).

RqlH-(1–505) translocates unidirectionally on single-stranded DNA

NTP hydrolysis by nucleic acid-dependent phosphohydrolases is often coupled to mechanical work—either duplex unwinding or displacement of protein–nucleic acid complexes—as a consequence of translocation of the phosphohydrolase enzyme along the nucleic acid. To address whether RqlH has translocase activity, we employed the streptavidin (SA) displacement assay developed by Kevin Raney and colleagues (25–27), as implemented in our previous studies of the mycobacterial AdnAB translocase (5). The 5′ ^{32}P-labeled 34-mer DNA oligonucleotides containing a single biotin moiety at the fourth internucleotide from the 5′ end or the second internucleotide from the 3′-end were pre-incubated with excess SA to form a stable SA–DNA complex (Figure 6, lanes 2 and 9) that was easily resolved from the free biotinylated 34-mer DNA (Figure 6, lanes 1 and 8) during native PAGE. The translocation assay scores the motor-dependent displacement of SA from the DNA in

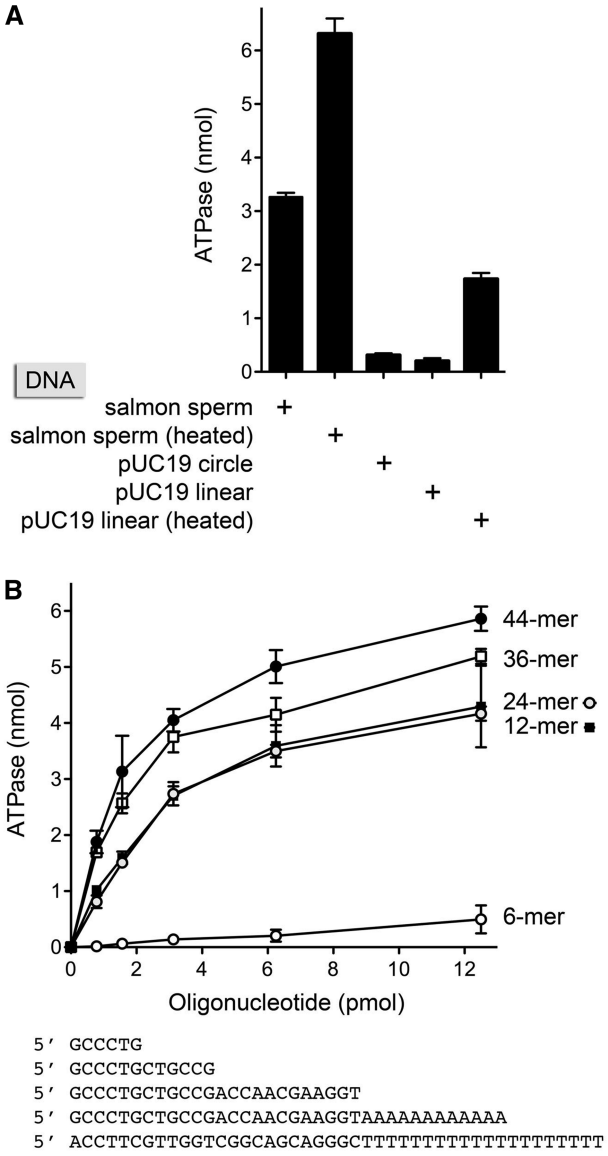

Figure 4. DNA dependence of ATP hydrolysis. (**A**) ATPase reaction mixtures containing 20 mM Tris–HCl (pH 8.0), 1 mM DTT, 5 mM $MgCl_2$, 1 mM (10 nmol) [α^{32}P]ATP, 1 pmol (100 nM) RqlH-(1–505) and 1 μg of the indicated salmon sperm or pUC19 DNAs were incubated at 37°C for 30 min. Blunt-end linear pUC19 was generated by digesting circular pUC19 with SmaI. Heat-denaturation of the DNAs was performed by incubating a 1 mg/ml DNA solution at 95°C for 10 min followed by quenching on ice. The extents of ATP hydrolysis are plotted. Each datum is the average of three separate experiments ± SEM. (**B**) ATPase reaction mixtures containing 20 mM Tris–HCl (pH 8.0), 1 mM DTT, 5 mM $MgCl_2$, 1 mM (10 nmol) [α^{32}P]ATP, 1 pmol (100 nM) RqlH-(1–505) and increasing amounts of 44-, 36-, 24-, 12- or 6-mer DNA oligonucleotides as specified were incubated at 37°C for 30 min. The nucleobase sequences of the oligonucleotides are shown at bottom. The extents of ATP hydrolysis are plotted as a function of added DNA (pmol). Each datum is the average of three separate DNA titration experiments ± SEM.

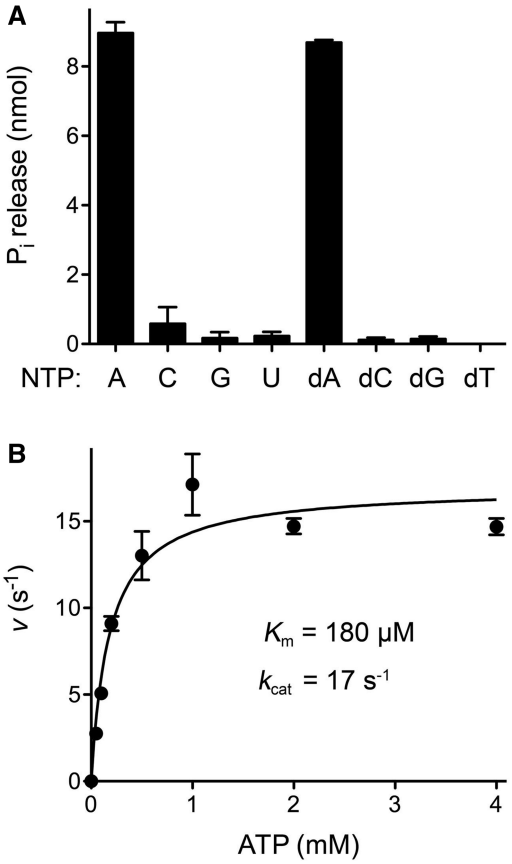

Figure 5. Nucleotide substrate specificity and kinetic parameters. (**A**) Reaction mixtures (10 μl) containing 20 mM Tris–HCl, pH 8.0, 1 mM DTT, 5 mM $MgCl_2$, 2 μg sonicated salmon sperm DNA, 0.5 pmol (50 nM) RqlH-(1–505) and 1 mM of the indicated NTP/dNTP were incubated at 37°C for 30 min. The reactions were quenched with 990 μl of malachite green reagent (Biomol Research Laboratories). Phosphate release was quantified by measuring A_{620} and interpolating the value to a phosphate standard curve. The values were corrected for the low levels of phosphate measured in control reaction mixtures containing 1 mM of the indicated NTP/dNTP but no added enzyme. Each datum is the average of three separate experiments ± SEM. (**B**) Reaction mixtures (60 μl) containing 20 mM Tris–HCl, pH 8.0, 1 mM DTT, 5 mM $MgCl_2$, 1.25 μM 44-mer oligonucleotide (shown in Figure 4B), 200 nM RqlH-(1–505) and either 0.05, 0.1, 0.2, 0.5, 1.0, 2.0 or 4.0 mM [α^{32}P]ATP were incubated at 37°C. Aliquots (5 μl) were withdrawn at 1, 2, 3 and 5 min and quenched immediately with formic acid. The extents of ATP hydrolysis were plotted as a function of time for each ATP concentration and the initial rates were derived by linear regression analysis in Prism. The initial rates (pmol/s) were divided by the molar amount of input enzyme to obtain a turnover number v (per second), which is plotted in the figure as a function of ATP concentration. Each datum is the average of three separate time course experiments ± SEM. A non-linear regression curve fit of the data to the Michaelis-Menten equation (in Prism) is shown. The K_m and k_{cat} values are indicated.

the presence of ATP and excess free biotin, which instantly binds to free SA and precludes SA rebinding to the labeled DNA. The rationale of the assay is that directional tracking of the motor along the DNA single strand will displace SA from one DNA end, but not the other. As depicted in Figure 6, the enzyme acts like a 'cow-catcher' on a locomotive engine. When moving 3′–5′, it can displace SA as it collides with the 5′-biotin–SA. In contrast, a 3′-biotin–SA is not expected to be displaced by a 3′–5′ translocase, because the motor moves away from the SA and simply falls off the free 5′ end. The converse outcomes apply to a 5′–3′ translocase; it displaces a 3′ SA, but not a 5′ SA. The instructive finding was that

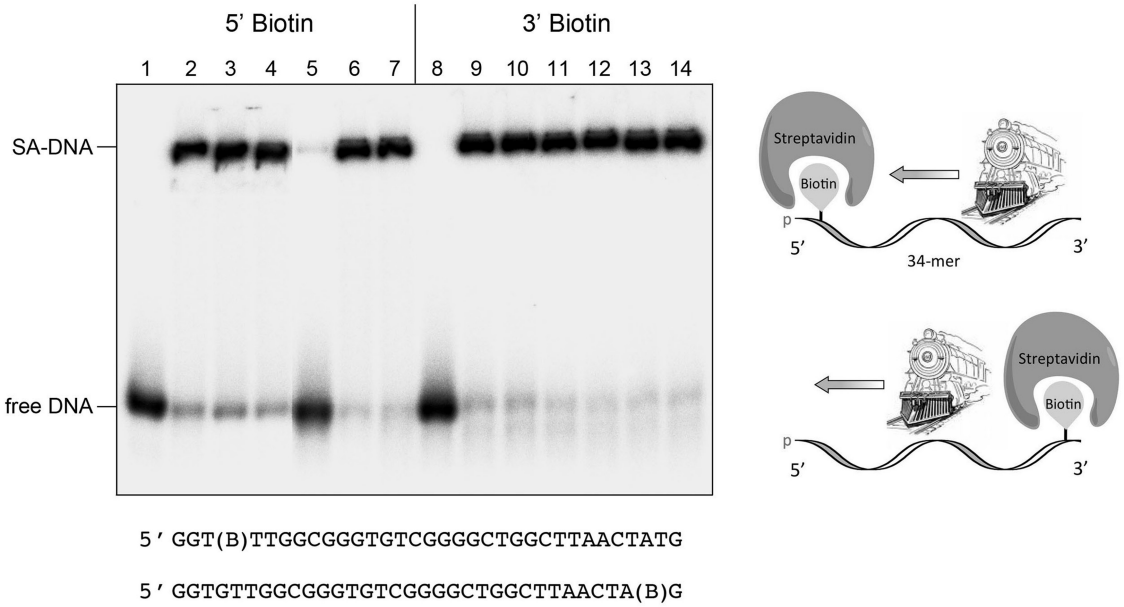

Figure 6. Directionality of RqlH-(1–505) translocation on ssDNA. Shown in the right panel is a schematic representation of directional 5′ SA displacement by the RqlH-(1–505) motor acting as a cowcatcher to pry apart the otherwise stable SA–biotin interaction. Translocase assays were performed as described under Experimental Procedures. Native PAGE analysis of the reaction products is shown in the left panel. The species corresponding to SA–DNA complex and free DNA are indicated. The nucleobase sequences of the 5′ ^{32}P-labeled 5′ or 3′ biotinylated 34-mer single-stranded DNAs are shown at bottom with (B) signifying the position of the biotin spacer. The complete translocase reaction mixtures in lanes 5 and 12 contained 1 mM ATP, 5 mM $MgCl_2$, 0.5 pmol ^{32}P-labeled 3′ or 5′ biotinylated DNA attached to SA, and 1 pmol of RqlH-(1–505). Enzyme was omitted from control reactions in lanes 2 and 9. Enzyme and SA were both omitted from control reactions in lanes 1 and 8. ATP was omitted from the reactions in lanes 3 and 10. Magnesium was omitted from the reactions in lanes 4 and 11. The complete reaction mixtures in lanes 6 and 13 contained 1 pmol of the RqlH-(1–505)-K49A mutant protein in lieu of RqlH-(1–505). The complete reaction mixtures in lanes 7 and 14 contained 1 pmol of the RqlH-(1–505)-D148A mutant protein instead of RqlH-(1–505).

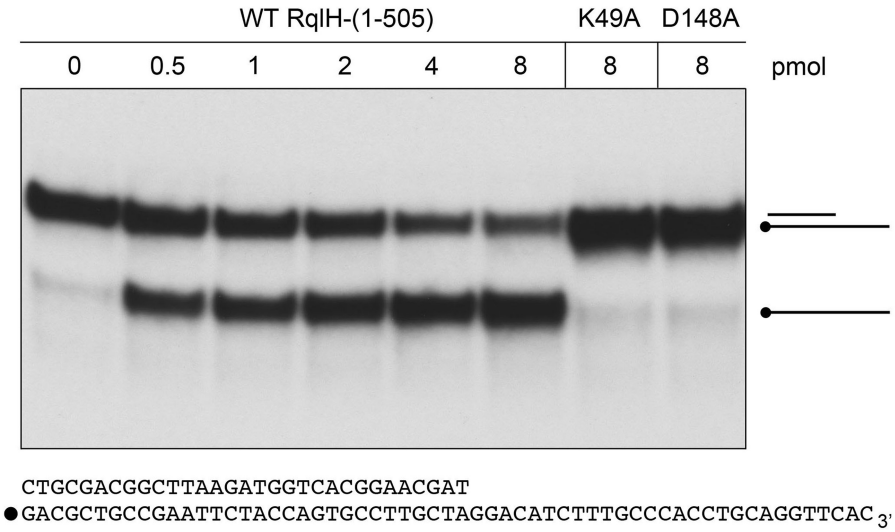

Figure 7. RqlH-(1–505) unwinds a 3′-tailed DNA duplex. Helicase reaction mixtures (10 µl) contained 1 pmol of 3′-tailed duplex substrate (depicted at bottom, with the 5′ ^{32}P-labeled denoted by filled circle) and the indicated amounts of the RqlH-(1–505), RqlH-(1–505)-K49A, or RqlH-(1–505)-D148A proteins. The reaction products were analyzed by native PAGE and visualized by autoradiography.

RqlH-(1–505) readily displaced SA from a 5′ biotin–SA complex on the 34-mer single-stranded DNA to yield the free ^{32}P-labeled 34-mer strand (Figure 6, lane 5), but was unable to displace SA from a 3′ biotin–SA complex tested in parallel (Figure 6, lane 12). Stripping of the 5′ biotin–SA complex by RqlH-(1–505) to liberate free DNA depended on ATP (Figure 6, lane 3) and magnesium (Figure 6, lane 4). The ATPase-defective RqlH-(1–505) mutants K49A and D148A displayed no detectable 3′–5′ translocase activity (Figure 6, lanes 6 and 7).

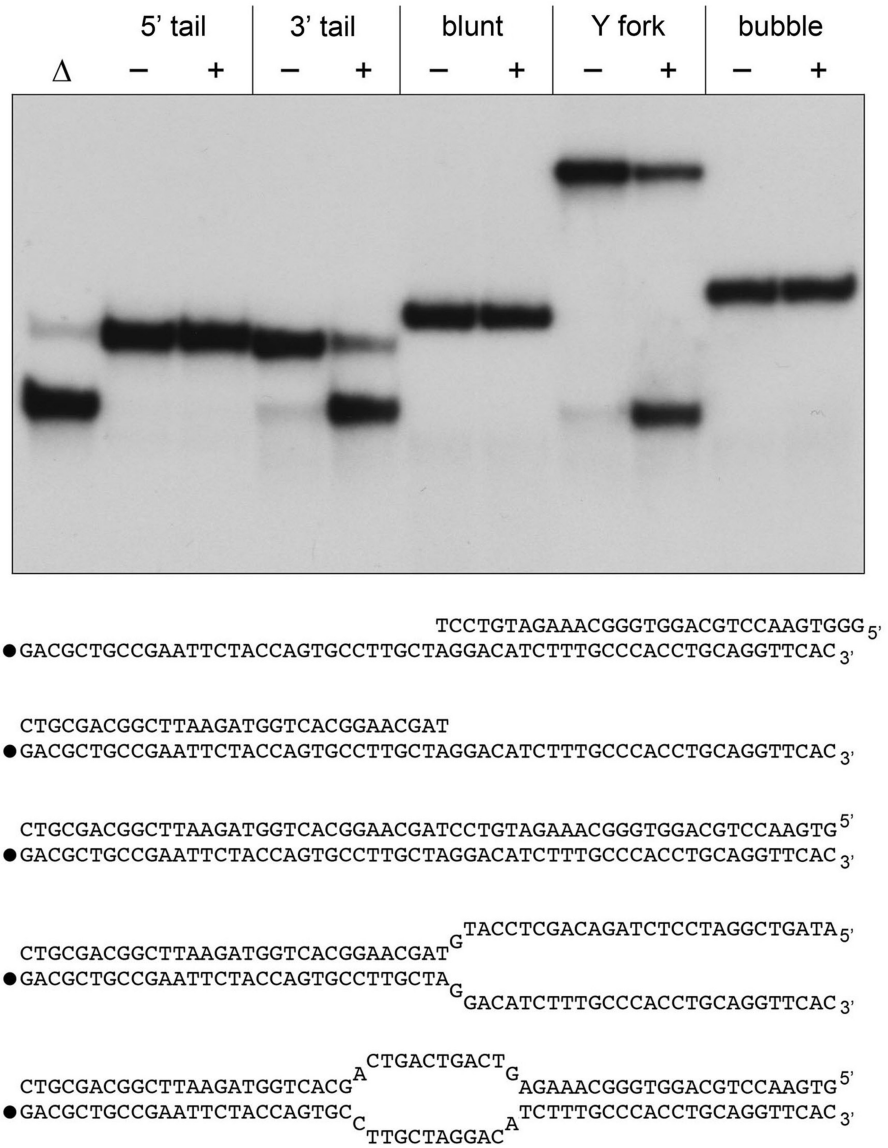

Figure 8. Directionality of duplex unwinding and requirement for a loading strand. Helicase reaction mixtures (10 μl) contained 1 pmol of the indicated ^{32}P-labeled DNA substrate and (where indicated by plus symbol) 4 pmol RqlH-(1–505). Enzyme was omitted from the reaction mixtures in lanes –. A reaction mixture lacking enzyme that was heat denatured prior to PAGE is included in lane Δ. The reaction products were analyzed by native PAGE and visualized by autoradiography (top panel). The 5′-tail, 3′-tail, blunt, Y form, and bubble substrates are shown in the bottom panel, with the with the 5′ ^{32}P-labeled denoted by filled circle.

RqlH-(1–505) unwinds a 3′-tailed DNA duplex

In light of the 3′–5′ translocase activity demonstrated above, we tested RqlH-(1–505) for helicase activity with a 3′-tailed duplex substrate consisting of a 31-bp duplex with a 28-nt 3′ single-strand tail to serve as a potential 'loading strand' (Figure 7). The helicase assay format we used entailed preincubation of RqlH-(1–505) and labeled DNA, followed by initiation of unwinding by addition of ATP, with simultaneous addition of a 'trap' of excess unlabeled 59-mer displaced strand that: (i) minimizes reannealing of any ^{32}P-labeled 59-mer that was unwound by RqlH-(1–505) and (ii) competes with the loading strand for binding to any free RqlH-(1–505) or RqlH-(1–505) that dissociated from the labeled DNA without unwinding it. Consequently, the assay predominantly gauges a single round of strand displacement by RqlH-(1–505) bound to the labeled 3′-tailed duplex prior to the onset of ATP hydrolysis. We found that RqlH-(1–505) unwound the 3′-tailed DNA substrate to yield a radiolabeled free single strand that migrated faster than the input tailed duplex during native PAGE (Figure 7); the helicase reaction product comigrated with free 59-mer generated by thermal denaturation of the substrate (Figure 8). The yield of unwound product increased with input enzyme in excess of the input substrate (Figure 7), as expected for the single-turnover assay format. The ATPase-defective RqlH-(1–505) mutants K49A and D148A were inactive in duplex unwinding (Figure 7).

RqlH-(1–505) failed to unwind a 59-bp blunt duplex DNA substrate (Figure 8), thereby attesting to a

requirement for a single strand tail to serve as a loading strand for the helicase. RqlH-(1–505) also failed to unwind a 5′-tailed duplex substrate (Figure 8). RqlH-(1–505) did unwind a 'Y fork' duplex consisting of a 31-bp duplex with one blunt end and 28-nt 5′ and 3′ single-strand tails at the other end, but it did not unwind a doubly blunt-ended duplex with a 12-nt internal bubble (Figure 8). These results establish that RqlH-(1–505) is a unidirectional motor, powered by ATP hydrolysis, that tracks 3′–5′ along the loading strand and unwinds duplex DNA.

Phylogenetic distribution of RqlH helicases

The biochemical properties of *M. smegmatis* RqlH elucidated above are consonant with those of *E. coli* RecQ, insofar as RqlH is a single-strand DNA-dependent ATPase/dATPase that translocates 3′–5′ on single-stranded DNA and has 3′–5′ helicase activity (28–31). All of these activities of RqlH are inherent to the monomeric 505-amino acid core catalytic unit, comprising the RecQ-like ATPase and zinc-binding domains, separated by an RqlH-specific linker domain. The basic motor functions of RqlH do not require the C-terminal 187-amino acid segment that includes the ComF-like domain. Nonetheless, it is conceivable that the C-terminus confers some added functional value, e.g. by mediating protein–protein interactions or allowing for unwinding of unusual nucleic structures other than those surveyed presently.

In our view, the domain organization is sufficiently diverged from that of the classical *E. coli* RecQ to warrant the designation of RqlH as the founding member of a new RecQ-like clade. A search of the NCBI data base with *M. smegmatis* RqlH recovered dozens of RqlH-like proteins, composed of the same domains, arrayed in the same order, and with similar inter-domain spacing. Among the mycobacterial taxa, RqlH is present in species *M. vanabaalenii*, *gilvum*, *KMS, JL* and *JDM601*, none of which have a homolog of classical *E. coli* RecQ with its HRDC domain (Supplemental Table S1). RqlH is not found in *M. tuberculosis*. However, RqlH is present in the proteomes of more than 50 other species of the phylum *Actinobacteria* (many of which have both RqlH and RecQ). Indeed, RqlH homologs are distributed widely in the bacterial domain of life, being present in species belonging to the phyla *Cyanobacteria*, *Chloroflexi*, *Firmicutes*, *Bacteriodetes* and *Proteobacteria*. Among the gamma-subdivision of proteobacteria, it is noteworthy that whereas the *E. coli* K12 strain (the standard model for bacterial biochemistry and genetics) encodes RecQ but lacks a homolog of RqlH, many other strains of *E. coli* have both RqlH and a classical RecQ in their proteomes (Table S1). An alignment of *M. smegmatis* and *E. coli* CFT073 RqlH proteins highlights primary structure conservation across the entire polypeptide, to the extent of 374/691 positions of side chain identity/similarity (Supplementary Figure S3).

In conclusion, we unveil RqlH as a novel clade of superfamily II helicase, with a RecQ-type motor, found in diverse bacterial taxa. Our initial biochemical analysis and phylogenetic analyses provide impetus for future studies, including: (i) characterization of the *in vivo* effects of ablating *rqlH*, singly and in combination with other potentially redundant superfamily II helicases; and (ii) biochemical and structural analyses of the distinctive C-terminal domain.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1 and Supplementary Figures 1–3.

FUNDING

U.S. National Institutes of Health [grant AI64693]. SS is an American Cancer Society Research Professor. Funding for open access charge: NIH (grant AI64693).

Conflict of interest statement. None declared.

REFERENCES

1. Sinha,K.M., Stephanou,N.C., Gao,F., Glickman,M.S. and Shuman,S. (2007) Mycobacterial UvrD1 is a Ku-dependent DNA helicase that plays a role in multiple DNA repair events, including double-strand break repair. *J. Biol. Chem.*, **282**, 15114–15125.
2. Sinha,K.M., Stephanou,N.C., Unciuleac,M.C., Glickman,M.S. and Shuman,S. (2008) Domain requirements for DNA unwinding by mycobacterial UvrD2, an essential DNA helicase. *Biochemistry*, **47**, 9355–9364.
3. Sinha,K.M., Glickman,M.S. and Shuman,S. (2009) Mutational analysis of Mycobacterium UvrD1 identifies functional groups required for ATP hydrolysis, DNA unwinding, and chemomechanical coupling. *Biochemistry*, **48**, 4019–4030.
4. Sinha,K.M., Unciuleac,M.C., Glickman,M.S. and Shuman,S. (2009) AdnAB: a new DSB-resecting motor-nuclease from Mycobacteria. *Genes Dev.*, **23**, 1423–1437.
5. Unciuleac,M.C. and Shuman,S. (2010) Characterization of the mycobacterial AdnAB DNA motor provides insights to the evolution of bacterial motor-nuclease machines. *J. Biol. Chem.*, **285**, 2632–2641.
6. Unciuleac,M.C. and Shuman,S. (2010) Double-strand break unwinding and resection by the mycobacterial helicase-nuclease AdnAB in the presence of mycobacterial SSB. *J. Biol. Chem.*, **285**, 34319–34329.
7. Curti,E., Smerdon,S.J. and Davis,E.O. (2007) Characterization of the helicase activity and substrate specificity of *Mycobacterium tuberculosis* UvrD. *J. Bacteriol.*, **189**, 1542–1555.
8. Williams,A., Güthlein,C., Beresford,N., Böttger,E.C., Springer,B. and Davis,E.O. (2011) UvrD2 is essential in Mycobacterium tuberculosis, but its helicase activity is not required. *J. Bacteriol.*, **193**, 4487–4494.
9. Aniukwu,J., Glickman,M.S. and Shuman,S. (2008) The pathways and outcomes of mycobacterial NHEJ depend on the structure of the broken DNA ends. *Genes Dev.*, **22**, 512–527.
10. Gupta,R., Barkan,D., Redelman-Sidi,G., Shuman,S. and Glickman,M.S. (2011) Mycobacteria exploit three genetically distinct DNA double-strand break repair pathways. *Mol. Microbiol.*, **79**, 316–330.
11. Güthlein,C., Wanner,R.M., Sander,P., Davis,E.O., Bosshard,M., Jiricny,J., Böttger,E.C. and Springer,B. (2009) Characterization of the mycobacterial NER system reveals novel functions of the *uvrD1* helicase. *J. Bacteriol.*, **191**, 555–562.
12. Yeeles,J.T.P. and Dillingham,M.S. (2007) A dual-nuclease mechanism for DNA break processing by AddAB-type helicase-nucleases. *J. Mol. Biol.*, **371**, 66–78.

13. Sanchez,H., Kidane,D., Cozar,M.C., Graumann,P.L. and Alonso,J.C. (2006) Recruitment of *Bacillus subtilis* RecN to DNA double-strand breaks in the absence of DNA end processing. *J. Bacteriol.*, **188**, 353–360.
14. Handa,N., Morimatsu,K., Lovett,S.T. and Kowalczykowski,S.C. (2009) Reconstitution of initial steps of dsDNA break repair by the RecF pathway of *E. coli*.. *Genes Dev.*, **23**, 1234–1245.
15. Rocha,E.P.C., Cornet,E. and Michel,B. (2005) Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genetics*, **1**, 247–259.
16. Bernstein,D.A., Zittel,M.C. and Keck,J.L. (2003) High-resolution structure of the *E. coli* RecQ helicase core. *EMBO J.*, **22**, 4910–4921.
17. Bernstein,K.A., Gangloff,S. and Rothstein,R. (2010) The RecQ DNA helicases in DNA repair. *Annu. Rev. Genet.*, **44**, 393–417.
18. Bernstein,D.A. and Keck,J.L. (2005) Conferring substrate specificity to DNA helicases: role of the RecQ HRDC domain. *Structure*, **13**, 1173–1182.
19. Kitano,K., Yoshihara,N. and Hakoshima,T. (2007) Crystal structure of the HRDC domain of human Werner syndrome protein WRN. *J. Biol. Chem.*, **282**, 2717–2728.
20. Wu,L., Chan,K.L., Ralf,C., Bernstein,D.A., Garcia,P.L., Bohr,V.A., Vindigni,A., Janscak,P., Keck,J.L. and Hickson,I.D. (2005) The HRDC domain of BLM is required for the dissolution of double Holliday junctions. *EMBO J.*, **24**, 2679–2687.
21. Bernstein,D.A. and Keck,J.L. (2003) Domain mapping of *Escherichia coli* RecQ defines the roles of conserved N- and C-terminal regions in the RecQ family. *Nucleic Acids Res.*, **31**, 2778–2785.
22. Killoran,M.P. and Keck,J.L. (2006) Three HRDC domains differentially modulate *Deinococcus radiodurans* RecQ DNA helicase biochemical activity. *J. Biol. Chem.*, **281**, 12849–12857.
23. Nakasugi,K., Svenson,C.J. and Neilan,B.A. The competence gene, *comF*, from *Synechocyst*is sp. strain PCC 3803 is involved in natural transformation, phototactic motility and piliation. *Microbiology*, **152**, 3623–3631.
24. Krahn,J.M., Kim,J.H., Burns,M.R., Parry,R.J., Zalkin,H. and Smith,J.L. (1997) Coupled formation of an amidotransferase interdomain ammonia channel and a phosphoribosyltransferase active site. *Biochemistry*, **36**, 11061–11068.
25. Morris,P.D. and Raney,K.D. (1999) DNA helicases displace streptavidin from biotin-labeled oligonucleotides. *Biochemistry*, **38**, 5164–5171.
26. Morris,P.D., Byrd,A.K., Tackett,A.J., Cameron,C.E., Tanega,P., Ott,R., Fanning,E. and Raney,K.D. (2002) Hepatitis C virus NS3 and simian virus 40 T antigen helicases displace streptavidin from 5′-biotinylated oligonucleotides but not from 3′-biotinylated oligonucleotides: evidence for directional bias in translocation on single-stranded DNA. *Biochemistry*, **41**, 2372–2378.
27. Byrd,A.K. and Raney,K.D. (2004) Protein displacement by an assembly of helicase molecules aligned along single-stranded DNA. *Nat. Struct. Mol. Biol.*, **11**, 531–538.
28. Umezu,K., Nakayama,K. and Nakayama,H. (1990) *Escherichia coli* RecQ protein is a DNA helicase. *Proc. Natl Acad. Sci. USA*, **87**, 5363–5367.
29. Harmon,F.G. and Kowalczykowski,S.C. (2001) Biochemical characterization of the DNA helicase activity of the *Escherichia coli* RecQ helicase. *J. Biol. Chem.*, **276**, 232–243.
30. Xu,H.Q., Deprez,E., Zhang,A.H., Tauc,P., Ladjimi,M.M., Brochon,J.C., Auclair,C. and Xi,X.G. (2003) The *Escherichia coli* RecQ helicase functions as a monomer. *J. Biol. Chem.*, **278**, 34925–34933.
31. Zhang,X.D., Dou,S.X., Xie,P., Hu,J.S., Wang,P.Y. and Xi,X.G. (2006) *Escherichia coli* RecQ is a rapid, efficient, and monomeric helicase. *J. Biol. Chem.*, **281**, 12655–12663.