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).
As the name implies, mycobacterial UvrD1 is homologous to *E. coli* UvrD, a prototypical superfamly 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 superfamly I helicase UvrD, yet it has a C-terminal HRDC domain, which is a feature of RecQ-type superfamly 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 RecQ-like helicase (named RqIH); the 691-amino acid protein encoded by *M. smegmatis* gene MSMEG_5935. We report that RqIH 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). RqIH defines a new clade of superfamly II helicases distributed widely among bacterial taxa.

**MATERIALS AND METHODS**

**RqIH proteins**

The ORFs encoding full-length RqIH (MSMEG_5935) and a C-terminal truncation mutant RqIH-(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-His10Smt3 to generate expression plasmids encoding the RqIH or RqIH-(1–505) polypeptides fused to an N-terminal His10Smt3 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-His10Smt3-RqIH and pET28b-His10Smt3-RqIH-(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 A660 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 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 RqIH or RqIH-(1–505) proteins eluted in the 200 and 500 mM imidazole fractions, which were pooled, supplemented with Smt3-specific protease Ulp1 (at a RqIH:Ulp1 ratio of 1000:1) and then dialyzed against 41 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 His10Smt3 tag. The tag-free RqIH or RqIH-(1–505) proteins were separated from His10Smt3 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.
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 MgCl₂, 1 mM [α³²P]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 radioactivity was visualized by autoradiography and [³²P]ADP formation was quantified by scanning the TLC plate with a Fujix BAS-2500 imager.

---

**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 four putative zinc-binding cysteines are highlighted in green in the bottom panel.
Streptavidin displacement assay of RqlH translocation on DNA

Synthetic 34-mer oligodeoxynucleotides of otherwise identical nucleobase sequence containing a Biotin-ON (1–505) as specified were preincubated for 5 min at room temperature. The mixtures were supplemented with 200 mM CaCl2, 0.6% SDS, 25% glycerol and 15 mM of a unlabeled single-stranded DNA (the 24-mer oligodeoxynucleotide shown in Figure 4B) to mask any binding of RqlH-(1–505) to 32P-DNA released from the SA–DNA complex. 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 (RecQ-like helicase) in anticipation of its biochemical characterization. In addition, the RqlH segment from amino acids 435–499 is homologous to the tetracyclisteine 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•Mn2+ 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<sub>10</sub>Smt3 fusion and isolated it from a soluble extract by nickel-agarose chromatography. The His<sub>10</sub>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<sup>32</sup>P]ATP in the presence of magnesium and salmon sperm DNA resulted in progressive hydrolysis of the labeled ATP to [α<sup>32</sup>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).
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 Kd 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 Pi 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 Km of 180 μM ATP and a km 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' 32P-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...
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...
RqlH-(1–505) readily displaced SA from a 5’ biotin–SA complex on the 34-mer single-stranded DNA to yield the free 32P-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 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’ 32P-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 MgCl2, 0.5 pmol 32P-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’ 32P-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) 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 32P-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 mono-meric 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. vanbaalenii, 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


