

# The Pif1 family in prokaryotes: what are our helicases doing in your bacteria?

Matthew L. Bochman, Colleen P. Judge, and Virginia A. Zakian

Department of Molecular Biology, Princeton University, Princeton, NJ 08544

**ABSTRACT** Pif1 family helicases, which are found in nearly all eukaryotes, have important roles in both nuclear and mitochondrial genome maintenance. Recently, the increasing availability of genome sequences has revealed that Pif1 helicases are also widely found in diverse prokaryotes, but it is currently unknown what physiological function(s) prokaryotic Pif1 helicases might perform. This Perspective aims to briefly introduce the reader to the well-studied eukaryotic Pif1 family helicases and speculate on what roles such enzymes may play in bacteria. On the basis of our hypotheses, we predict that Pif1 family helicases are important for resolving common issues that arise during DNA replication, recombination, and repair rather than functioning in a eukaryotic-specific manner.

## Monitoring Editor

Douglas Kellogg  
University of California,  
Santa Cruz

Received: Feb 22, 2011

Revised: Mar 22, 2011

Accepted: Apr 13, 2011

## INTRODUCTION TO THE EUKARYOTIC PIF1 FAMILY HELICASES

The prototypical member of the Pif1 family of helicases is the *Saccharomyces cerevisiae* Pif1 (ScPif1) protein, which was initially isolated in a genetic screen for mutations affecting mitochondrial DNA (mtDNA) recombination (reviewed in Bochman *et al.*, 2010). It was later discovered that ScPif1 is a superfamily 1b helicase with a variety of cellular roles. Aside from affecting mtDNA maintenance, ScPif1 also uses its helicase activity to perform several nuclear functions, including telomerase inhibition (Zhou *et al.*, 2000; Boule *et al.*, 2005), Okazaki fragment processing (Pike *et al.*, 2010), and resolving G-quadruplex (G4) structures (Ribeyre *et al.*, 2009; Paeschke *et al.*, 2011).

Additionally, *S. cerevisiae* encodes a second Pif1 family helicase known as *S. cerevisiae* Rrm3 (ScRrm3). Like ScPif1, ScRrm3 has roles in maintaining the stability of both the mitochondrial and nuclear genomes (Bochman *et al.*, 2010). ScRrm3 is a component of the replisome (Azvolinsky *et al.*, 2006), where it is thought to act as an accessory helicase, facilitating timely DNA replication by disrupting stable, nonnucleosomal protein–DNA complexes that would otherwise inhibit replication fork progression (Ivessa *et al.*, 2003; Azvolinsky

*et al.*, 2009). Two-dimensional gel electrophoresis of replication intermediates demonstrates that the absence of ScRrm3 results in paused and broken DNA replication forks (Ivessa *et al.*, 2000, 2003). Mutations in *RRM3* also result in local increases in recombination at regions containing sites of ScRrm3 action, such as the rDNA (Keil and McWilliams, 1993) and a tRNA-rich stretch of chromosome VII (Ivessa *et al.*, 2003).

Whereas multiple fungi also encode two Pif1 family helicases, other organisms contain variable numbers of homologues, from one in fission yeast and humans to 23 in the fungal arthropod pathogen *Metarhizium anisopliae* ARSEF 23. Most multicellular organisms, however, encode only a single Pif1 helicase. In many instances, little is known about these enzymes other than that they share essentially equal sequence similarity to ScPif1 and ScRrm3 and, where examined, seem to be found in both nuclei and mitochondria (Bochman *et al.*, 2010).

## PROKARYOTIC PIF1 HELICASES

Searches of the National Center for Biotechnology Information (NCBI) Protein database reveal that many prokaryotic genomes also encode Pif1-like helicases. Due to sequence similarity with RecD (ScPif1 and ScRrm3 are 18.9 and 16.4% identical, respectively, to *Escherichia coli* RecD), however, many of the prokaryotic Pif1s are misannotated as RecDs (or ambiguously annotated using a variety of other terms). The situation in bacteria is further complicated by the existence of multiple RecD (Montague *et al.*, 2009) and/or Pif1 homologues in some species (see Figure 1 and Table 1 for examples).

Pif1s and RecDs are all superfamily 1b helicases that share seven conserved motifs (I, Ia, II, III, IV, V, and VI) common to SFI enzymes, as well as three additional motifs (A, B, and C) found only in Pif1 and RecD family helicases (see Figure 2). Only Pif1-like helicases,

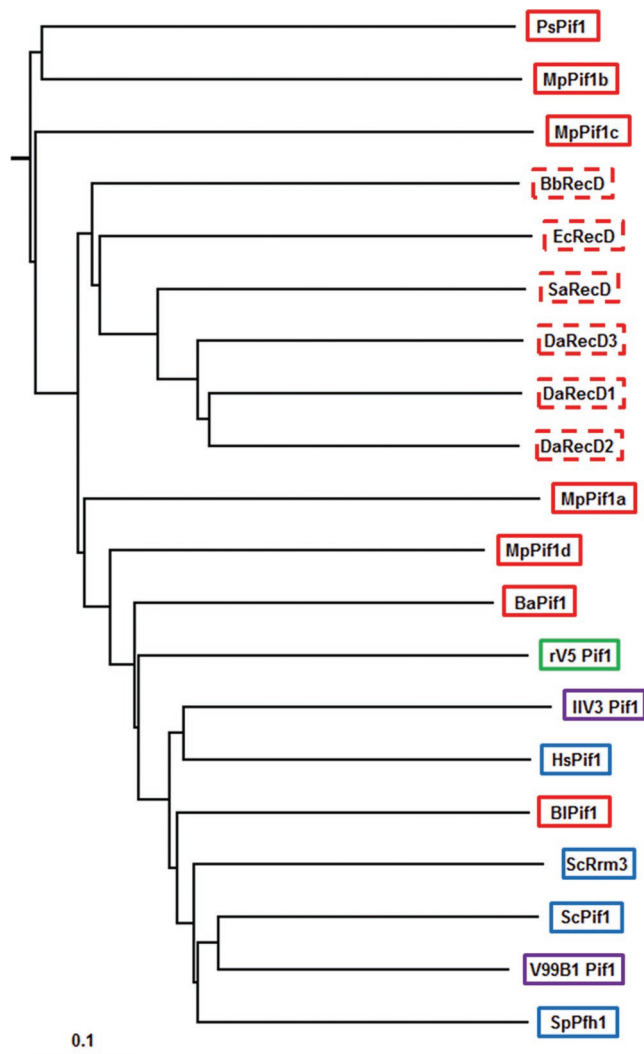
DOI: 10.1091/mbc.E11-01-0045

Address correspondence to: Virginia A. Zakian (vzakian@molbio.princeton.edu).

Abbreviations used: G4, G-quadruplex; HR, homologous recombination; mtDNA, mitochondrial DNA; ScPif1, *Saccharomyces cerevisiae* Pif1 helicase; ScRrm3, *S. cerevisiae* Rrm3 helicase.

© 2011 Bochman *et al.* This article is distributed by The American Society for Cell Biology under license from the author(s). Two months after publication it is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

“ASCB®,” “The American Society for Cell Biology®,” and “Molecular Biology of the Cell®” are registered trademarks of The American Society of Cell Biology.



**FIGURE 1:** Phylogenetic tree revealing the relationships between selected Pif1 and RecD helicases. Pif1 (solid boxes) and RecD (dashed boxes) protein sequences were aligned using CuxtalX (v. 2.0.12) by iterating each step of the alignment. The same program was used to both draw and bootstrap a neighbor-joining tree. TreeView (v. 1.6.6) was used to visualize the phylogenetic tree and root it with the outgroup (human  $\beta$ -actin, unpublished data). The scale bar indicates the number of substitutions per site. Red boxes indicate prokaryotes, blue boxes denote eukaryotes, the green box indicates a bacteriophage, and the purple boxes designate viruses. Ps, *Psychrobacter* sp. PRwf-1; Mp, *Mucilaginibacter paludis* DSM 18603; Bb, *Borrelia burgdorferi* N40; Ec, *E. coli*; Sa, *Staphylococcus aureus*; Da, *Desulfobacterium autotrophicum* HRM2; Ba, *Bacteroides* sp. 2\_1\_16; rV5, *E. coli* phage rV5; IIV3, Invertebrate Iridescent Virus-3; Hs, *Homo sapiens*; Bl, *Bifidobacterium longum* subsp. *longum* F8; Sc, *S. cerevisiae*; V99B1, *Emiliania huxleyi* virus 99B1; and Sp, *Schizosaccharomyces pombe*. Protein sequences are available upon request.

however, contain the Pif1 family signature sequence (DKLeXvARai-RkqXkPFGGIQ), a degenerate sequence located between motifs II and III, the function of which is currently unknown (Bochman *et al.*, 2010). Aside from CLUSTAL alignments of protein sequences to determine phylogenies, one way to differentiate between Pif1 and RecD helicases is with the signature sequence. Using the sequence above as a Basic Local Alignment Search Tool (BLAST) query of the NCBI database returns only Pif1 homologues, although the more

distantly related a Pif1 family helicase is from the canonical ScPif1, the more degenerate the signature sequence becomes (compare BaPif1 and BbPif1 in Figure 2). Careful analysis using CLUSTAL alignments, however, reveals that prokaryotic Pif1 enzymes are generally more closely related to eukaryotic Pif1 helicases than they are to any of the various RecD subgroups (Figure 1).

### Why do prokaryotes need Pif1 family helicases?

We hypothesize that Pif1 family helicases are important for resolving common issues that arise during DNA metabolism rather than performing eukaryotic-specific functions. Thus, based on what we know about eukaryotic Pif1 helicases, prokaryotic Pif1s may serve multiple functions, including maintaining prokaryotic telomeres (in bacteria with long linear chromosomes), resolving DNA and DNA–RNA secondary structures, and complementing the lack of other helicases in vivo. Each of these possibilities is discussed separately.

**Prokaryotic telomere maintenance.** Whereas *Escherichia coli* and *Bacillus* have circular chromosomes, many bacteria have linear chromosomes, plasmids, and/or phages that replicate without a circular intermediate (reviewed in Stewart *et al.*, 2004, and Casjens and Huang, 2008). Well-known examples of such organisms include *Borrelia burgdorferi* (the causative agent of Lyme disease), many *Streptomyces* species (sources of various antibiotics), and the  $\phi$ 29 *Bacillus subtilis* bacteriophage. To solve the end replication problem, at least two different types of prokaryotic telomeres have evolved: 1) those with ends protected by a protein that is covalently linked to the terminal nucleotide (e.g., *Streptomyces* spp. and  $\phi$ 29), and 2) those with covalently closed terminal hairpins (e.g., *Borrelia* spp.). In eukaryotes, various Pif1 family helicases are known or hypothesized to associate with telomeres. For instance, ScPif1 is a catalytic inhibitor of telomerase, ScRrm3 is important for replication through telomeric DNA, and mammalian Pif1 helicases interact with telomerase in vivo (Bochman *et al.*, 2010). Although the structures and sequences of prokaryotic telomeres are unrelated to the well-known eukaryotic telomerase-generated telomeres, Pif1 family helicases may nevertheless be involved in prokaryotic telomere maintenance.

The ends of linear chromosomes and plasmids are inherently recombinogenic (Stewart *et al.*, 2004; Casjens and Huang, 2008). In some instances, such as generating antigen variability in *Borrelia* spp., this recombination is beneficial, and bacterial Pif1 helicases that resolve recombination intermediates may promote such processes. Indeed, ScPif1 promotes recombination in the linear mtDNA of *S. cerevisiae* (Foury and Dyck, 1985). Alternatively, prokaryotic Pif1 helicases may possess a nucleoprotein disruption activity similar to that hypothesized for ScRrm3 (Ivessa *et al.*, 2003; Azvolinsky *et al.*, 2009) that is used to displace DNA end-binding proteins during the replication of linear chromosomes and plasmids. Such an activity could be important for the complete replication of DNA ends.

More broadly, it is widely accepted that the genomic DNA in eukaryotic organelles, such as mitochondria, has a prokaryotic origin. It was long overlooked and often disputed, however, that some eukaryotic mtDNA molecules, including *S. cerevisiae* mtDNA, are linear (Nosek *et al.*, 1998), perhaps reflecting an origin from a prokaryote species with a linear genome. Because Pif1 family helicases are important for mtDNA maintenance (Bochman *et al.*, 2010), it is tempting to speculate that the first eukaryotic Pif1 helicase may have had a mitochondrial (and hence prokaryotic) origin, with nuclear isoforms evolving later. Indeed, there are several examples of mitochondrial proteins with putative prokaryotic origins (e.g., mitochondrial ribosomal proteins; Henze and Martin, 2001) that are encoded in the nuclear genome.

	I	Ia	II	Signature sequence	III
ScPif1	<b>I</b> F <b>Y</b> T <b>G</b> S <b>A</b> G <b>T</b> G <b>K</b> S	<b>L</b> A <b>A</b> C <b>N</b> I <b>G</b>	<b>L</b> V <b>V</b> D <b>E</b> I <b>S</b>	<b>D</b> K <b>L</b> D <b>F</b> I <b>A</b> R <b>K</b> I <b>R</b> K <b>N</b> H <b>Q</b> P <b>F</b> G <b>G</b> I <b>Q</b>	<b>F</b> C <b>G</b> D <b>F</b> F <b>Q</b> L
ScRrm3	<b>V</b> F <b>F</b> T <b>G</b> S <b>A</b> G <b>T</b> G <b>K</b> S	<b>L</b> A <b>A</b> V <b>T</b> I <b>G</b>	<b>L</b> I <b>I</b> D <b>E</b> I <b>S</b>	<b>D</b> K <b>L</b> E <b>Q</b> I <b>A</b> R <b>R</b> I <b>R</b> K <b>N</b> D <b>D</b> P <b>F</b> G <b>G</b> I <b>Q</b>	<b>L</b> T <b>G</b> D <b>F</b> F <b>Q</b> L
HsPif1	<b>I</b> F <b>F</b> T <b>G</b> S <b>A</b> G <b>T</b> G <b>K</b> S	<b>V</b> A <b>A</b> C <b>H</b> I <b>G</b>	<b>L</b> V <b>I</b> D <b>E</b> I <b>S</b>	<b>D</b> K <b>L</b> E <b>A</b> V <b>A</b> R <b>A</b> V <b>R</b> Q <b>N</b> K <b>P</b> F <b>G</b> G <b>I</b> Q	<b>I</b> C <b>G</b> D <b>F</b> L <b>Q</b> L
BaPif1	<b>V</b> F <b>V</b> T <b>G</b> K <b>A</b> G <b>S</b> G <b>K</b> T	<b>I</b> A <b>A</b> I <b>N</b> A <b>G</b>	<b>L</b> I <b>I</b> D <b>E</b> I <b>S</b>	<b>D</b> T <b>I</b> D <b>R</b> K <b>L</b> R <b>V</b> Y <b>E</b> S <b>D</b> E <b>P</b> F <b>G</b> G <b>V</b> Q	<b>M</b> F <b>G</b> D <b>L</b> F <b>Q</b> L
BbPif1	<b>V</b> F <b>L</b> T <b>G</b> G <b>A</b> G <b>S</b> G <b>K</b> S	<b>A</b> A <b>A</b> V <b>L</b> L <b>G</b>	<b>V</b> I <b>I</b> D <b>E</b> I <b>S</b>	<b>M</b> I <b>A</b> E <b>A</b> L <b>S</b> Q <b>R</b> A <b>R</b> E <b>S</b> K <b>L</b> P <b>W</b> G <b>G</b> M <b>R</b>	<b>A</b> V <b>G</b> D <b>F</b> A <b>Q</b> L
EcRecD	<b>S</b> V <b>I</b> S <b>G</b> G <b>P</b> G <b>T</b> G <b>K</b> T	<b>K</b> A <b>A</b> A <b>R</b> L <b>T</b>	<b>L</b> V <b>V</b> D <b>E</b> A <b>S</b>	<b>s</b> r <b>l</b> i <b>d</b> a <b>l</b> p <b>d</b> h <b>a</b> r <b>v</b> i <b>f</b> l <b>g</b> d <b>r</b> d <b>q</b>	<b>F</b> L <b>G</b> D <b>R</b> D <b>Q</b> L

	IV	A	B	C	V	VI
ScPif1	<b>M</b> L <b>Q</b> K <b>V</b> F <b>R</b>	<b>D</b> M <b>L</b> N <b>R</b> M <b>R</b> L <b>G</b>	<b>V</b> G <b>A</b> Q <b>V</b> M <b>M</b> V <b>K</b> N	<b>V</b> N <b>G</b> S <b>L</b> G	<b>L</b> S <b>I</b> H <b>K</b> S <b>Q</b> G <b>Q</b> T <b>L</b> P <b>K</b> V	<b>Y</b> V <b>A</b> L <b>S</b> R
ScRrm3	<b>L</b> L <b>T</b> K <b>V</b> F <b>R</b>	<b>D</b> I <b>L</b> N <b>A</b> I <b>R</b> Y <b>G</b>	<b>E</b> D <b>A</b> Q <b>V</b> M <b>L</b> K <b>N</b>	<b>V</b> N <b>G</b> S <b>L</b> G	<b>L</b> S <b>I</b> H <b>K</b> A <b>Q</b> G <b>Q</b> T <b>I</b> Q <b>R</b> L	<b>Y</b> V <b>A</b> L <b>S</b> R
HsPif1	<b>E</b> L <b>T</b> K <b>V</b> W <b>R</b>	<b>S</b> L <b>L</b> Q <b>A</b> V <b>R</b> L <b>G</b>	<b>L</b> G <b>A</b> Q <b>V</b> M <b>L</b> V <b>K</b> N	<b>V</b> N <b>G</b> A <b>R</b> G	<b>M</b> S <b>I</b> H <b>K</b> S <b>Q</b> G <b>M</b> T <b>L</b> D <b>C</b> V	<b>Y</b> V <b>A</b> L <b>S</b> R
BaPif1	<b>E</b> L <b>T</b> K <b>I</b> F <b>R</b>	<b>N</b> V <b>L</b> N <b>N</b> I <b>R</b> N <b>Y</b>	<b>V</b> G <b>A</b> R <b>V</b> M <b>S</b> L <b>V</b> N	<b>Y</b> N <b>G</b> M <b>L</b> G	<b>I</b> T <b>I</b> H <b>K</b> S <b>Q</b> G <b>L</b> T <b>F</b> D <b>K</b> I	<b>Y</b> V <b>A</b> L <b>S</b> R
BbPif1	<b>M</b> L <b>S</b> H <b>N</b> Q <b>R</b>	<b>D</b> V <b>L</b> S <b>D</b> V <b>R</b> H <b>G</b>	<b>I</b> G <b>C</b> Q <b>V</b> M <b>F</b> L <b>Q</b> N	<b>V</b> N <b>H</b> T <b>R</b> G	<b>T</b> T <b>I</b> H <b>K</b> S <b>Q</b> G <b>A</b> T <b>L</b> D <b>D</b> L	<b>Y</b> V <b>A</b> L <b>S</b> R
EcRecD	<b>L</b> L <b>Q</b> K <b>S</b> Y <b>R</b>	<b>Q</b> L <b>A</b> A <b>A</b> I <b>N</b> R <b>G</b>	<b>E</b> G <b>R</b> E <b>V</b> M <b>I</b> A <b>R</b> N	<b>F</b> N <b>G</b> D <b>I</b> G	<b>M</b> T <b>V</b> H <b>K</b> S <b>Q</b> G <b>S</b> E <b>F</b> D <b>H</b> A	<b>Y</b> T <b>A</b> V <b>T</b> R

**FIGURE 2:** Sequence conservation of helicase motifs in the Pif1 family and *E. coli* RecD proteins. The sequences of the conserved SFI helicase motifs I–VI, the Pif1/RecD-specific motifs A–C, and the Pif1 family signature sequence are shown (protein sequences were aligned as in Figure 1). Residues that are completely conserved in all six sequences are bold. Residues that are identical in  $\geq 3$  sequences are red, and conserved similarities (i.e., either similar to the red residues in the same column or  $\geq 3$  similar residues in the same column) are green. *E. coli* RecD does not display conservation of the Pif1 family signature sequence, so the residues aligned by ClustalX in this region are shown in lowercase. The amino acid similarity groups were defined as FYW, IVLM, RK, DE, GA, TS, and NQ. Here Bb indicates *B. bacteriovorus* instead of *B. burgdorferi*; the remaining abbreviations are the same as in Figure 1.

**DNA/RNA secondary structure resolution.** Eukaryotic Pif1 helicases have recently been implicated in the processing of G4 structures (Ribeyre *et al.*, 2009; Sanders, 2010; Paeschke *et al.*, 2011). A G4 structure is a four-stranded RNA or DNA secondary structure that is held together by Hoogsteen G-G base pairing. Human Pif1 and ScPif1 both bind and efficiently unwind G4 structures in vitro (Ribeyre *et al.*, 2009; Sanders, 2010). It is known that guanine-rich sequences that have the potential to form G4 structures are enriched in telomeric DNA and rDNA, and G4 structures are also found at transcriptional regulatory regions and preferred meiotic double-stranded break sites (Neidle and

Balasubramanian, 2006; Capra *et al.*, 2010). Thus G4 sequences are likely to have roles in diverse cellular functions, such as telomere maintenance (Paeschke *et al.*, 2008) and transcriptional regulation (Rawal *et al.*, 2006; Huppert *et al.*, 2008). Moreover, a recent analysis of 18 prokaryotic genomes revealed that the association of G4 motifs with transcriptional regulatory regions is not limited to eukaryotes, as evidenced by conserved G4 motifs in the promoters of gene orthologues in distantly related bacteria (Rawal *et al.*, 2006). Although helicases such as *E. coli* RecQ are also known to unwind G4 structures (Wu and Maizels, 2001), not all bacterial genomes encode RecQ homologues (Table 1). It is therefore plausible that prokaryotes that lack RecQ helicases may instead encode Pif1 helicases to resolve G4 structures. Furthermore, as might be expected, bacteria with GC-rich genomes are predicted to contain more G4-forming sequences than are organisms with lower GC content (Rawal

*et al.*, 2006). It is therefore intriguing that Pif1 family helicases are found throughout the Bifidobacteriales (a large order of Actinobacteria) and the phylum Bacteroidetes (which includes human gut commensals), the genomes of which are particularly GC-rich.

Previously, it was shown that ScPif1 preferentially unwinds DNA–RNA hybrids relative to DNA–DNA substrates (Boule and Zakian, 2007), much like the *E. coli* UvrD helicase (Matson, 1989). It is hypothesized that ScPif1 inhibits telomerase activity by unwinding the DNA–RNA hybrid formed between telomeric DNA and the telomerase RNA template (Boule *et al.*, 2005). Although

Organism	Phylum	Helicase homologues per genome							
		Pif1	RecD	DinG	RecQ	Rep	UvrD	PcrA	
<i>Mucilaginibacter paludis</i> DSM 18603	Bacteroidetes	4	0	0	3	1	1	0	
<i>Psychrobacter</i> sp. PRwf-1	$\gamma$ -proteobacteria	1	1	0	0	2	1	0	
<i>Rhodomicrobium vannielii</i> ATCC 17100	$\alpha$ -proteobacteria	1	0	0	0	0	1	0	
<i>Sulfurovum</i> sp. NBC37–1	$\epsilon$ -proteobacteria	1	0	0	1	1	2	0	
<i>Desulfobacterium autotrophicum</i> HRM2	$\delta$ -proteobacteria	1	3	0	0	0	0	2	
<i>Gardnerella vaginalis</i> AMD	Actinobacteria	1	0	0	0	0	0	1	
<i>Deferribacter desulfuricans</i> SSM1	Deferribacteres	1	1	0	0	0	0	1	
<i>Anaerobaculum hydrogeniformans</i> ATCC BAA-1850	Synergistetes	1	0	0	0	0	1	1	
<i>Thermodesulfovibrio yellowstonii</i> DSM 11347	Nitrospirae	1	0	0	0	0	1	2	
<i>Chloroherpeton thalassium</i> ATCC 35110	Chlorobi	1	0	0	1	0	1	0	
<i>Borrelia burgdorferi</i> N40	Spirochaetes	0	1	0	0	0	1	1	
<i>Escherichia coli</i>	$\gamma$ -proteobacteria	0	1	1	1	1	1	0	
<i>Staphylococcus aureus</i>	Firmicutes	0	1	1	1	1	1	1	

Due to sequence divergence from their better-characterized homologues, it was difficult to definitively determine the difference between Pif1 and RecD, as well as Rep and UvrD, family helicases in some organisms. Thus, some of the proteins listed in these categories may represent intermediate forms.

**TABLE 1:** Repertoire of helicases in evolutionarily diverse prokaryotes as revealed by BLAST searches and ClustalX alignments.

the prokaryotic telomeres just discussed are not replicated by a conventional telomerase-based mechanism, perhaps Pif1 helicases function in bacteria to resolve DNA–RNA hybrids (i.e., R-loops) that form during transcription. Because R-loop accumulation leads to genomic instability in organisms from *E. coli* to mammals, resolution of R-loops is important for maintaining genome integrity (Li and Manley, 2006).

**Complementing the lack of other helicases.** Most genomes encode multiple distinct helicases that have specialized functions in DNA and RNA metabolism. For instance, there are 134 open reading frames in the *S. cerevisiae* genome that encode proteins containing helicase structural motifs (Shiratori *et al.*, 1999). The *E. coli* genome also encodes a variety of predicted helicases, with more than dozen such enzymes verified by direct biochemical assays. A universal collection of helicase types that is required for viability, however, has not been identified in bacteria or eukaryotes, likely due to functional redundancy *in vivo*. Thus, although *E. coli* encodes one member each of the DinG, RecD, RecQ, Rep, and UvrD helicase families (Table 1), other bacteria contain multiple homologues of these helicases, are devoid of them altogether, and/or possess other types of helicases (e.g., PcrA and Pif1) that may complement the lack of one or more of these enzymes. In *E. coli*, DinG is a 5′-3′ helicase with a putative role in homologous recombination (HR), RecD is the 5′-3′ member of the RecBCD complex (which is also involved in HR), RecQ (3′-5′) is involved in HR and double-stranded break repair, Rep is a 3′-5′ accessory replicative helicase needed for timely DNA replication, and UvrD is a 3′-5′ repair helicase that can remove RecA from single-stranded DNA (see Wu and Maizels, 2001; Montague *et al.*, 2009; Boubakri *et al.*, 2010; and references therein).

Given their sequence similarity to RecD enzymes, one might predict that Pif1 family helicases are found in the subset of bacterial species that lack a RecD homologue. Indeed, this situation is true in species such as *Rhodomicrobium vannielii* and *Gardnerella vaginalis*, which lack RecD. Other bacterial species, however, encode one or more homologues of both Pif1 and RecD (e.g., *Psychrobacter* sp. PRwf-1 and *Desulfobacterium autotrophicum*; Table 1). Additionally, across the diverse prokaryotic phyla (Ciccarelli *et al.*, 2006), many prokaryotes that encode Pif1 family helicases appear to lack both RecQ and DinG helicase homologues. As stated earlier in text, Pif1 helicases may function in place of RecQs to resolve G4 DNA and/or to remove R-loops, as DinG does in *E. coli* (Boubakri *et al.*, 2010).

Additionally, DinG acts in concert with either Rep or UvrD in *E. coli* to remove RNA polymerase from the path of replication forks when replication and transcription complexes collide (Boubakri *et al.*, 2010). This activity is reminiscent of the nucleoprotein disruption activity of ScRrm3 (Ivessa *et al.*, 2003), and the increased replication fork pausing seen at inverted ribosomal operons in *E. coli* *dinG rep* mutants is similar to the fork pausing observed at rDNA repeats in *rrm3Δ S. cerevisiae* (Ivessa *et al.*, 2000). Therefore, because eukaryotic Pif1 helicases have functions similar to those observed for the *E. coli* DinG, UvrD, and Rep helicases, prokaryotic Pif1 helicases may complement the lack of one or more of these enzymes in bacteria that do not encode DinG, UvrD, and/or Rep homologues.

## Outlook

The possibilities presented here are by no means comprehensive, and none of these hypotheses has been tested. Additionally, if prokaryotic Pif1-like helicases were acquired by horizontal gene transfer, they may not be expressed or active *in vivo*. Transcriptome data from *Bdellovibrio bacteriovorus* indicate, however, that BbPif1

mRNA is expressed (Lambert *et al.*, 2010), and biochemical results from a number of purified prokaryotic Pif1 helicases demonstrate that these enzymes are active *in vitro* (Bochman, Judge, and Zakian, unpublished). In any event, more bioinformatic analyses are needed to determine exactly how widespread Pif1 family helicases are among prokaryotes. Furthermore, molecular-genetic examination of nontraditional model organisms that encode Pif1-like helicases (e.g., *Bifidobacteria* and *Bacteroides* spp.) is required to begin examining these enzymes *in vivo*. To the best of our knowledge, no *in vivo* functional research has been performed on a prokaryotic Pif1 family helicase, but we hope that this article will generate interest in experimental tests of the roles of Pif1 helicases in bacterial cells.

## ACKNOWLEDGMENTS

We apologize to those researchers whose work was not discussed and/or cited due to space constraints. This work was supported by a postdoctoral research fellowship from the American Cancer Society (#PF-10-145-01-DMC to M.L.B.) and by a grant from the National Institutes of Health (GM26938 to V.A.Z.).

## REFERENCES

- Azvolinsky A, Dunaway S, Torres JZ, Bessler JB, Zakian VA (2006). The *S. cerevisiae* Rrm3p DNA helicase moves with the replication fork and affects replication of all yeast chromosomes. *Genes Dev* 20, 3104–3116.
- Azvolinsky A, Giresi PG, Lieb JD, Zakian VA (2009). Highly transcribed RNA polymerase II genes are impediments to replication fork progression in *Saccharomyces cerevisiae*. *Mol Cell* 34, 722–734.
- Bochman ML, Sabouri N, Zakian VA (2010). Unwinding the functions of the Pif1 family helicases. *DNA Repair (Amst)* 9, 237–249.
- Boubakri H, de Septenville AL, Viguera E, Michel B (2010). The helicases DinG, Rep and UvrD cooperate to promote replication across transcription units *in vivo*. *EMBO J* 29, 145–157.
- Boule JB, Vega LR, Zakian VA (2005). The yeast Pif1p helicase removes telomerase from telomeric DNA. *Nature* 438, 57–61.
- Boule JB, Zakian VA (2007). The yeast Pif1p DNA helicase preferentially unwinds RNA DNA substrates. *Nucleic Acids Res* 35, 5809–5818.
- Capra JA, Paeschke K, Singh M, Zakian VA (2010). G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in *Saccharomyces cerevisiae*. *PLoS Comput Biol* 6, e1000861.
- Casjens SR, Huang WM (2008). Prokaryotic telomeres: Replication mechanisms and evolution. In: *Origin and Evolution of Telomeres*, ed. J Nosek and L Tomáška, Austin, TX: Landes Bioscience.
- Ciccarelli FD, Doerks T, von Mering C, Creevey CJ, Snel B, Bork P (2006). Toward automatic reconstruction of a highly resolved tree of life. *Science* 311, 1283–1287.
- Foury F, Dyck EV (1985). A PIF-dependent recombinogenic signal in the mitochondrial DNA of yeast. *EMBO J* 4, 3525–3530.
- Henze K, Martin W (2001). How do mitochondrial genes get into the nucleus? *Trends Genet* 17, 383–387.
- Huppert JL, Bugaut A, Kumari S, Balasubramanian S (2008). G-quadruplexes: the beginning and end of UTRs. *Nucleic Acids Res* 36, 6260–6268.
- Ivessa AS, Lenzmeier BA, Bessler JB, Goudsouzian LK, Schnakenberg SL, Zakian VA (2003). The *Saccharomyces cerevisiae* helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes. *Mol Cell* 12, 1525–1536.
- Ivessa AS, Zhou JQ, Zakian VA (2000). The *Saccharomyces* Pif1p DNA helicase and the highly related Rrm3p have opposite effects on replication fork progression in ribosomal DNA. *Cell* 100, 479–489.
- Keil RL, McWilliams AD (1993). A gene with specific and global effects on recombination of sequences from tandemly repeated genes in *Saccharomyces cerevisiae*. *Genetics* 135, 711–718.
- Lambert C, Chang CY, Capeness MJ, Sockett RE (2010). The first bite—profiling the predatosome in the bacterial pathogen *Bdellovibrio*. *PLoS One* 5, e8599.
- Li X, Manley JL (2006). Cotranscriptional processes and their influence on genome stability. *Genes Dev* 20, 1838–1847.
- Matson SW (1989). *Escherichia coli* DNA helicase II (uvrD gene product) catalyzes the unwinding of DNA:RNA hybrids *in vitro*. *Proc Natl Acad Sci USA* 86, 4430–4434.

- Montague M, Barnes C, Smith HO, Chuang RY, Vashee S (2009). The evolution of RecD outside of the RecBCD complex. *J Mol Evol* 69, 360–371.
- Neidle S, Balasubramanian S (2006). *Quadruplex Nucleic Acids*, Cambridge, UK: Royal Society of Chemistry.
- Nosek J, Tomaska L, Fukuhara H, Suyama Y, Kovac L (1998). Linear mitochondrial genomes: 30 years down the line. *Trends Genet* 14, 184–188.
- Paeschke K, Capra JA, Zakian VA (2011). DNA replication through G-quadruplex motifs is promoted by the *Saccharomyces cerevisiae* Pif1 DNA helicase. *Cell* 145, 678–691.
- Paeschke K, Juranek S, Simonsson T, Hempel A, Rhodes D, Lipps HJ (2008). Telomerase recruitment by the telomere end binding protein-beta facilitates G-quadruplex DNA unfolding in ciliates. *Nat Struct Mol Biol* 15, 598–604.
- Pike JE, Henry RA, Burgers PM, Campbell JL, Bambara RA (2010). An alternative pathway for Okazaki fragment processing: resolution of fold-back flaps by Pif1 helicase. *J Biol Chem* 285, 41712–41723.
- Rawal P, Kummarasetti VB, Ravindran J, Kumar N, Halder K, Sharma R, Mukerji M, Das SK, Chowdhury S (2006). Genome-wide prediction of G4 DNA as regulatory motifs: role in *Escherichia coli* global regulation. *Genome Res* 16, 644–655.
- Ribeyre C, Lopes J, Boule JB, Piazza A, Guedin A, Zakian VA, Mergny JL, Nicolas A (2009). The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences in vivo. *PLoS Genet* 5, e1000475.
- Sanders CM (2010). Human Pif1 helicase is a G-quadruplex DNA-binding protein with G-quadruplex DNA-unwinding activity. *Biochem J* 430, 119–128.
- Shiratori A, Shibata T, Arisawa M, Hanaoka F, Murakami Y, Eki T (1999). Systematic identification, classification, and characterization of the open reading frames which encode novel helicase-related proteins in *Saccharomyces cerevisiae* by gene disruption and Northern analysis. *Yeast* 15, 219–253.
- Stewart P, Rosa PA, Tilly K (2004). Linear Plasmids in Bacteria: Common origins, Uncommon Ends. In: *Plasmid Biology*, ed. BE Funnell and GJ Phillips, Washington, DC: ASM Press, 291–301.
- Wu X, Maizels N (2001). Substrate-specific inhibition of RecQ helicase. *Nucleic Acids Res* 29, 1765–1771.
- Zhou J, Monson EK, Teng SC, Schulz VP, Zakian VA (2000). Pif1p helicase, a catalytic inhibitor of telomerase in yeast. *Science* 289, 771–774.