doi:10.1016/j.jmb.2006.01.077 J. Mol. Biol. (2006) 358, 1071–1080

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

A Conserved G4 DNA Binding Domain in RecQ Family Helicases

Michael D. Huber[1], Michelle L. Duquette[2], Jerome C. Shiels[2] and Nancy Maizels[1,2]*

[1]*Department of Biochemistry University of Washington Medical School, Seattle, WA 98195, USA*

[2]*Department of Immunology University of Washington Medical School, Seattle, WA 98195, USA*

**Corresponding author*

RecQ family helicases play important roles at G-rich domains of the genome, including the telomeres, rDNA, and immunoglobulin switch regions. This appears to reflect the unusual ability of enzymes in this family to unwind G4 DNA. How RecQ family helicases recognize this substrate has not been established. Here, we show that G4 DNA is a preferred target for BLM helicase within the context of long DNA molecules. We identify the RQC domain, found only in RecQ family enzymes, as an independent, high affinity and conserved G4 DNA binding domain; and show that binding to Holliday junctions involves both the RQC and the HRDC domains. These results provide mechanistic understanding of differences and redundancies of function and activities among RecQ family helicases, and of how deficiencies in human members of this family may contribute to genomic instability and disease.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Bloom syndrome; cancer; DNA binding domain; G-quadruplex; genomic instability

Introduction

RecQ helicases comprise a small family of enzymes critical for genomic stability.[1,2] Members of this family include the eponymous RecQ from *Escherichia coli*; Sgs1p from *Saccharomyces cerevisiae*; Rqh1 of *Schizosaccharomyces pombe*; and human BLM, WRN, RecQ1, RecQ4, and RecQ5 (Figure 1). Deficiencies in human RecQ family members BLM, WRN and RECQ4 cause Bloom syndrome, Werner syndrome and Rothmund-Thomson syndrome, respectively.[2–4] The shared hallmark of these three human diseases is genomic instability leading to early onset of cancer. Genomic instability in the absence of BLM is profound, evident at the cellular level as greatly elevated sister-chromatid exchanges and distinctive quadriradial structures. Individuals affected with Bloom syndrome develop a broad spectrum of malignancies, and typically die before reaching the age of 30; and predisposition to malignancy extends even to heterozygotes.[5] Werner syndrome is characterized by accelerated aging, which is accompanied by development of malignancies, predominantly sarcomas, which are an unusual tumor in the general population. Rothmund-Thomson syndrome is also characterized by genomic instability, development of osteosarcomas, and developmental defects. Deficiencies in other RecQ family helicases have not yet been shown to contribute to human disease, although biochemical and genetic analyses suggest that RecQ5, in particular, may have overlapping functions with BLM and WRN.[6–9] The intracellular functions of BLM and WRN may in some cases be redundant, despite the clear distinctions in disease symptomology.[10–12] It is therefore of considerable interest to identify the specific molecular mechanisms by which RecQ family enzymes identify and unwind their DNA targets.

Human RecQ helicases BLM and WRN and *S. cerevisiae* Sgs1 are all critical for stability and function of G-rich genomic regions, including the telomeres,[11,13–20] the rDNA,[21–27] and the immunoglobulin heavy chain switch regions.[3,28] A distinguishing biochemical property of some RecQ family helicases is the ability to unwind G4 DNA.[29–31] In G4 DNA, interactions among four DNA strands are

Present addresses: M. D. Huber, Department of Cell Biology, The Scripps Research Institute, La Jolla, CA 92037, USA; M. L. Duquette, Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA.

Abbreviations used: HJ, Holliday junction; EM, electron microscopy; MBP, maltose binding protein.

E-mail address of the corresponding author: maizels@u.washington.edu

0022-2836/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.

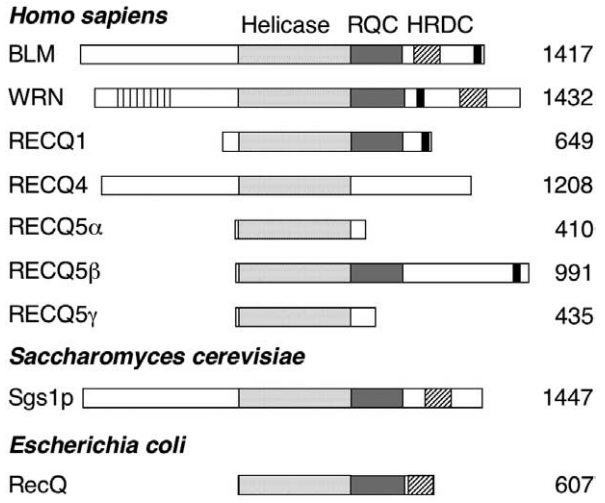

Figure 1. Conserved domains in RecQ family helicases. Alignment of human RecQ family helicases with related enzymes from *S. cerevisiae* and *E. coli*. Conserved regions are indicated by shading. The exonuclease domain (vertical stripes) is unique to WRN. Nuclear localization signals (NLS) are indicated by black boxes. Numbers on the right indicate the total number of amino acid residues in each helicase. Human RecQ5α, β and γ are encoded by a single gene; alternative splicing results in formation of the distinct polypeptides diagrammed.

stabilized by Hoogsteen base-pairing to form quartets of guanine bases.[32,33] In the canonical double-helix, G/C base-pairing prevents G4 DNA formation; however this structure may form as a result of transient strand separation during transcription[34,35] or replication[17,36–38] of G-rich sequences. G4 DNA unwinding activity may explain the importance of RecQ family members for these genomic regions.

Helicases are modular enzymes composed of distinct domains that combine to confer specificity and functionality.[39] The helicase domain provides the ATPase motor module critical to DNA unwinding. RecQ helicases are related by a high degree of homology within this domain, which carries the seven signature helicase motifs[40] and a motif important for ATP hydrolysis that is reminiscent of the Q-motif found in DEAD helicases.[41–43] Some but not all RecQ family helicases contain two additional domains, the RQC (RecQ C-terminal) domain, which is restricted to enzymes in this family; and the HRDC (helicase and RNase D C-terminal) domain (Figure 1). In addition to G4 DNA, RecQ helicases also unwind a variety of other non-duplex DNA structures, including forked DNA, Holliday junctions (HJ), D-loops, 5′ flaps, and triplex DNA.[1,4] The RQC domain of WRN has been shown to bind bubble and forked DNA;[44,45] and interaction between a positively charged surface on the HRDC domain and DNA has been proposed based on structural analysis of the HRDC domains from *S. cerevisiae* Sgs1p and *E. coli* RecQ.[46,47] However, it is not known whether a specific helicase domain confers high affinity G4 DNA binding.

Better understanding of how RecQ family helicases recognize and unwind their DNA targets will help explain how the enzymes in this family recognize their substrates in living cells, and provide mechanistic understanding of the distinct biochemical properties and disease phenotypes of human RecQ family enzymes, which differ in their domain composition (Figure 1). Here, we show that G4 DNA is a preferred target for BLM helicase within the context of long DNA molecules representative of active genomic regions. We identify the RQC domain as an autonomous and conserved high-affinity G4 DNA binding domain, which has clear preference for binding to G4 DNA and forked DNA relative to HJ. These results provide mechanistic understanding of both differences and redundancies in activity and function of RecQ family helicases, and provide new insights into how human members of this family may contribute to genomic stability and disease.

Results

BLM binds to G4 DNA formed by transcription within long DNA molecules

BLM helicase binds to short synthetic G4 DNAs with high affinity,[48] but it is not known how this helicase or other RecQ family members interact with longer molecules more representative of substrates formed in living cells. Transcription of a G-rich region within a plasmid results in the formation of G-loops, characteristic structures that contain G4 DNA interspersed with single-stranded regions on the G-rich (non-template) strand, and a stable RNA/DNA hybrid on the template strand.[34,35] Electron microscopic (EM) imaging allows ready visualization of G-loops, and it can also reveal proteins bound to specific structures within G-loops, including the mismatch repair factor, MutSα, which binds G4 DNA;[49] and the B cell-specific cytidine deaminase, AID, which binds single-stranded regions within G-loops.[35] We therefore used EM to analyze binding of BLM helicase to G4 DNA formed by *in vitro* transcription of pPH600, which carries a 0.6 kb G-rich region derived from the murine immunoglobulin switch regions within a longer plasmid backbone.[34] Free RNA was digested with RNase A, and the DNA was linearized by restriction digestion and then incubated with a mutant derivative of BLM, BLM^{K695A}, which has unaltered binding affinity (data not shown) but lacks unwinding activity due to a point mutation in the ATPase domain.[50] Protein/DNA complexes were visualized following Kleinschmidt spreading, as the large size of BLM helicase (1417 residues) permits direct EM visualization. As documented previously, following *in vitro* transcription approxi-

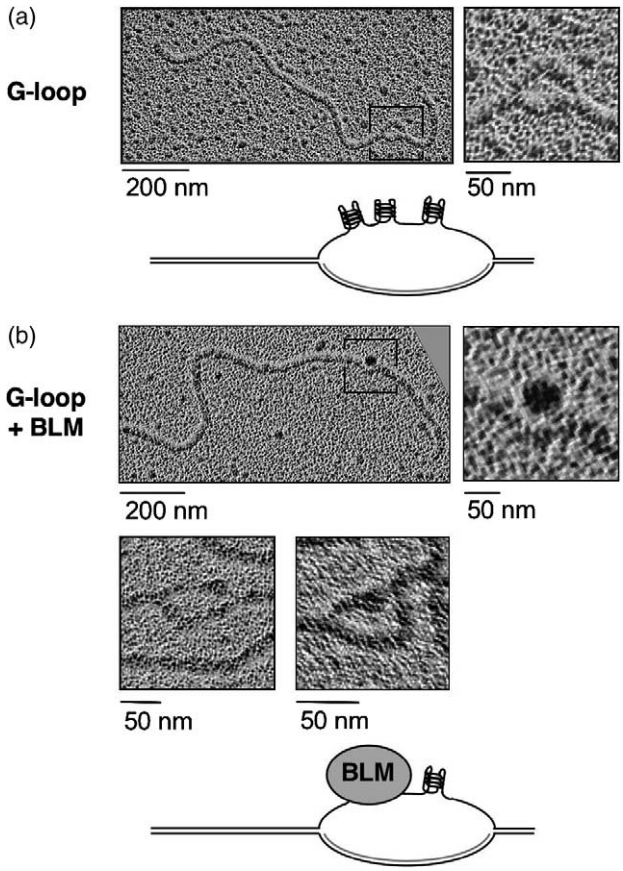

Figure 2. Binding of BLM to G-loops. (a) Representative electron micrographic image of a G-loop formed within transcribed, G-rich sequences from the murine Sγ3 switch region, in plasmid pPH600. The region containing the G-loop (framed) is shown enlarged adjacent to the image. A diagram of the G-loop structure is shown below. G4 DNA is shown as connected stacked squares on the top strand of the loop, and the RNA in the RNA–DNA heteroduplex is drawn in gray at the bottom of the loop. (b) G-loops bound by BLM^{K695}. BLM^{K695} bound to the G-loop is seen as a dark mass tethered to the DNA filament within the loop. A diagram of a BLM decorated G-loop is shown below. Scale bars are shown below each image.

mately half the transcribed DNAs contained G-loops[34,35] (e.g. Figure 2(a)). Following incubation with 15 nM BLM^{K695A} and protein/DNA crosslinking with glutaraldehyde, BLM could be directly visualized bound to the G-loops (e.g. Figure 2(b)). At this protein concentration, BLM^{K695A} was bound to 23% of G-loops ($n=40$), and in all cases binding was to the thinner filament of the G-loops. This is the filament that contains G4 DNA.[34] BLM^{K695A} was not observed bound to duplex regions of the plasmid ($n>200$), or to the thicker filament of the G-loops, which contains the RNA/DNA hybrid. Moreover, BLM was in nearly all cases observed at the central region of the loop rather than at the branch points at the ends of the loop. These results confirm the specificity of BLM for G4 DNA, consistent with biochemical analysis[29,48]

The RQC domain is an independent, high affinity domain for binding to G4 DNA

RecQ family helicases are composed of multiple domains. To identify the domain(s) key to recognition of G4 DNA, we carried out structure–function analysis of *E. coli* RecQ. Based on the high-resolution structure of the catalytically active core of *E. coli* RecQ,[51] we generated and purified maltose binding protein (MBP) fusion polypeptides bearing the conserved helicase domain (MBP-Hel), the RQC domain (MBP-RQC), and the HRDC domain (MBP-HRDC; Figure 3(a)). MBP alone does not bind single-stranded, fork, HJ or G4 DNA, at concentrations up to 750 nM (data not shown).[52] Gel mobility shift analysis of binding to synthetic, parallel G4 DNA showed that full-length RecQ and MBP-RQC bound to G4 DNA with comparably high affinity (apparent $k_D=2.9$ and 7.5 nM, respectively; Figure 3(b) and Table 1). MBP-HRDC also bound G4 DNA, but with much lower affinity (apparent $k_D=102$ nM; Figure 3(b) and Table 1) than the RQC domain or full-length RecQ. MBP-Hel did not bind G4 DNA (data not shown). This analysis identified the RQC domain of *E. coli* RecQ as a high affinity and autonomous G4 DNA binding domain.

The RQC domain binds with high affinity to forked substrates, but not HJ

Forked DNAs are good substrates for unwinding by RecQ helicase family members,[1] particularly *E. coli* RecQ.[31] MBP-RQC bound with very high affinity to forked DNA, comparable to full-length RecQ (apparent $k_D=0.4$ nM and 0.5 nM, respectively; Figure 3(b) and Table 1). In contrast, MBP-HRDC bound poorly to forked DNA substrates (Figure 3(b) and Table 1). The RQC domain is therefore an independent, high affinity binding domain for forked substrates.

HJ are unwound by RecQ family helicases, and are bound with high affinity by full-length RecQ (apparent $k_D=2.6$ nM; Figure 3(b) and Table 1). However, neither MPB-RQC nor MBP-HRDC bound to HJ with affinity comparable to full-length RecQ (apparent $k_D=55$ nM and 52 nM, respectively; Figure 3(b) and Table 1). Thus, neither the RQC nor the HRDC explain the high affinity of full-length RecQ for HJ.

G4 DNA binding by the RQC domain is conserved

To establish whether the RQC is a conserved G4 DNA binding domain, we analyzed binding by the RQC domain of human BLM. The polypeptide tested spanned BLM residues 994–1202, equivalent to the RQC domain fragment of *E. coli* RecQ (Figure 4(a)). Gel mobility shift analysis showed

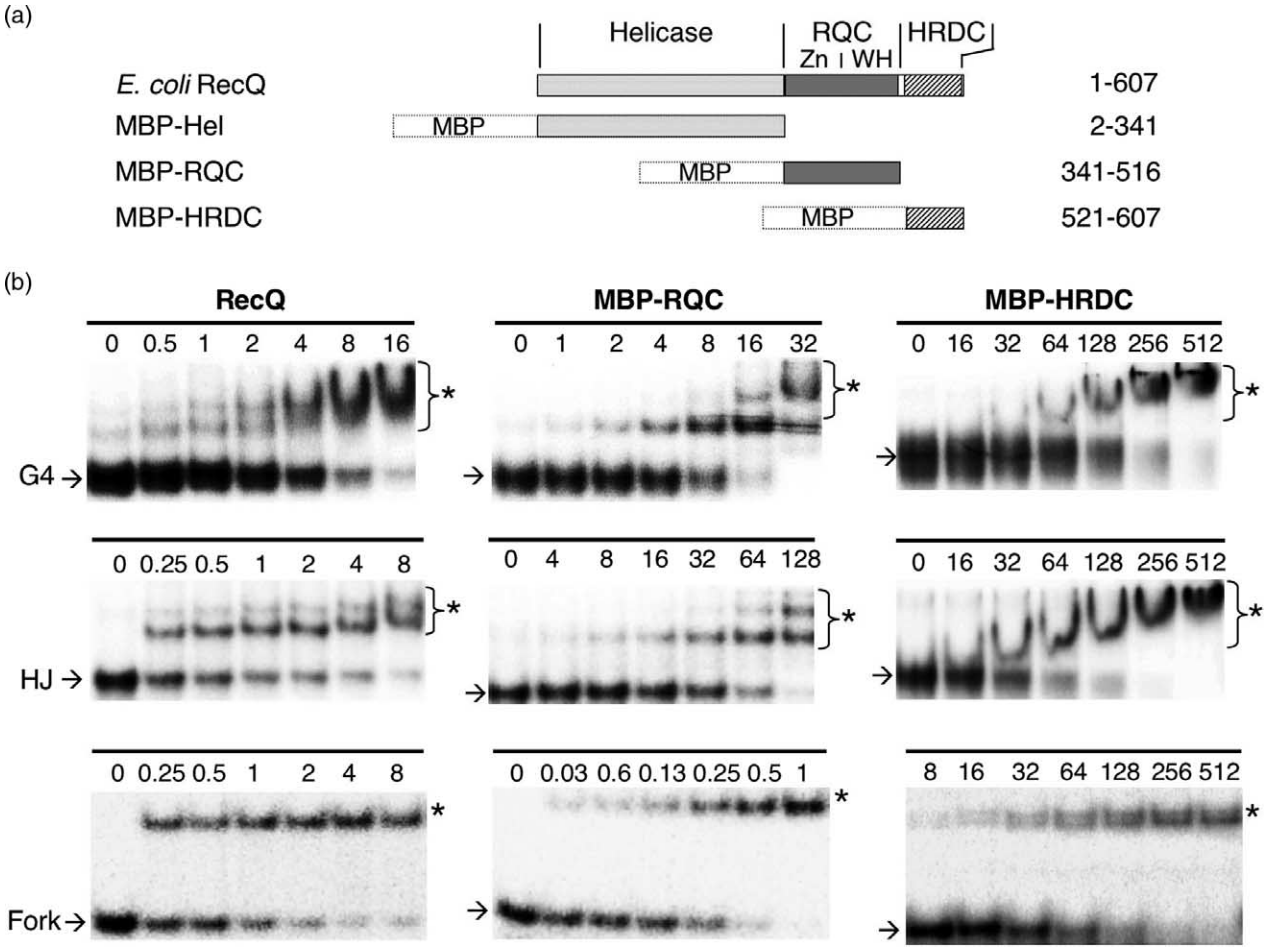

Figure 3. The RQC domain is an independent, high affinity binding domain for forked and G4 DNA substrates. (a) Recombinant fusions of *E. coli* RecQ fragments with MBP. Numbers on the right indicate amino acid residues of *E. coli* RecQ represented in the MBP fusions. (b) Gel mobility shift analysis of *E. coli* RecQ and MBP-RQC and MBP-HRDC fusion proteins binding to ^{32}P-labeled G4 DNA, HJ and forked DNA substrates. Protein concentrations (nM) are indicated above each image; position of free DNA by arrows at the left; and bound DNA by asterisks and brackets at the right. See Table 1 for summary.

that the BLM RQC domain binds with high affinity to G4 DNA (apparent $k_D = 3$ nM; Figure 4(b), Table 2). Like the *E. coli* RQC domain, the BLM RQC domain also bound to forked DNA (apparent $k_D = 9$ nM; Figure 4(c), Table 2), and HJ (apparent $k_D = 26$ nM; Figure 4(d), Table 2). The apparent affinities (k_D) and relative affinities (k_D/k_D G4 DNA) of the independent BLM RQC domain for G4 DNA and HJ, and the affinities relative to G4 DNA are broadly comparable to those of full-length BLM,[48] as shown in Table 2. These results show that the RQC domain is a conserved G4 DNA binding domain, which confers most of the substrate specificity to BLM.

The RQC domain contains a single binding site for G4 DNA and forked DNA

Identification of the RQC domain as an independent, conserved and high affinity domain for binding both G4 DNA and forked DNA raised the question of whether these structured substrates occupy the same or distinct binding sites. We addressed this by competition experiments. Essentially identical competition profiles were obtained if binding of MBP-RQC to labeled G4 DNA was carried out in the presence of either unlabeled G4 DNA or forked DNA; and *vice versa* (Figure 5(a)). In each case, 50% competition was observed at a 1:1 ratio of labeled to unlabeled DNA. The RQC domain therefore contributes a single functional binding site for G4 DNA and forked DNA.

Table 1. Relative binding of *E. coli* RecQ and its RQC and HRDC domains to structured DNAs (nm)

	RecQ	RQC	HRDC
G4	2.9 (±0.9)	7.5 (±2.7)	102 (±18)
Fork	0.4 (±0.2)	0.5 (±0.2)	42 (±10)
HJ	2.6 (±0.9)	55 (±5)	52 (±11)

Summary of analysis of binding by RecQ and its recombinant derivatives. Apparent affinities (k_D) were determined from the results of at least three separate experiments. Margin of error is shown in parentheses.

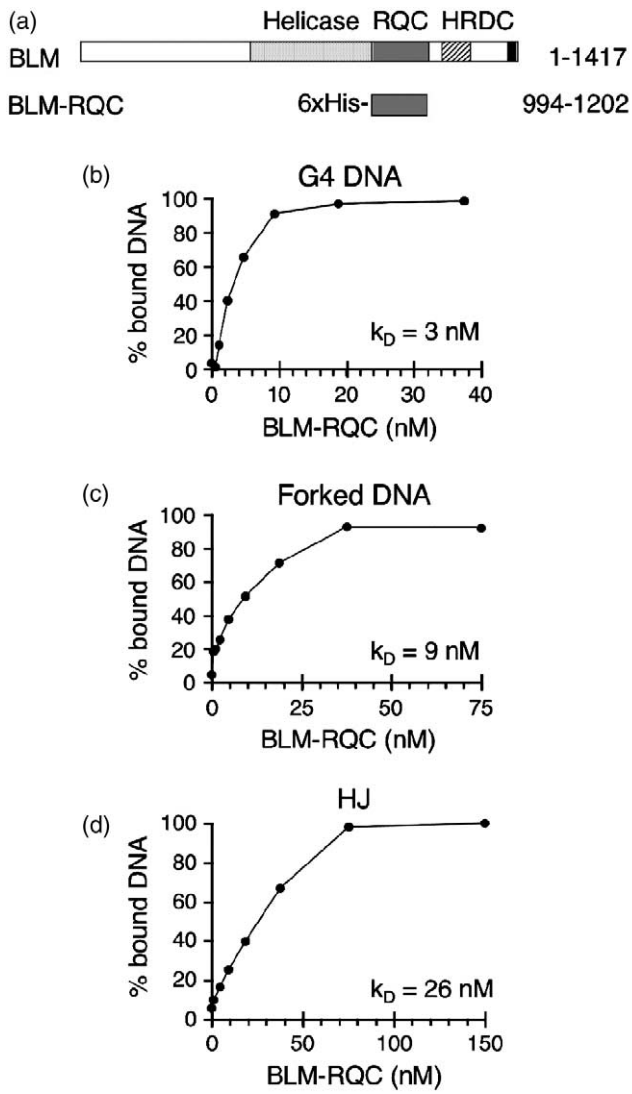

Figure 4. The RQC domain is a conserved, independent G4 DNA binding domain. (a) The recombinant BLM-RQC domain fragment. Full-length BLM is shown above. Recombinant fusions of *E. coli* RecQ fragments with MBP. (b)–(d) Quantifications of representative gel mobility shift analyses of BLM-RQC binding to (b) G4 DNA, (c) forked DNA, and (d) HJ. Apparent affinities (k_D) are shown within each graph. See Table 2 for summary.

Table 2. Binding of the BLM RQC domain to structured DNAs

	BLM-RQC		BLM	
	k_D (nM)	k_D/k_D^{G4}	k_D (nM)	k_D/k_D^{G4}
G4	2.7 (±0.5)	1.0	4	1.0
HJ	24.5 (±1.5)	9.1	78	19.5
Fork	9.5 (±1.1)	3.5	n.d.	n.d.

Summary of analysis of binding by BLM RQC (Figure 4) compared to published results for binding by full-length BLM.[48] Apparent affinities (k_D) and affinities relative to affinities for G4 DNA (k_D/k_D G4 DNA) are shown.

Full-length RecQ contains a single binding site for G4 DNA and HJ

In contrast to forked substrates, HJ did compete G4 DNA binding by full-length RecQ, but did not compete effectively for G4 DNA binding by MBP-RQC (Figure 5(b)). These experiments confirmed specificity of the RQC for G4 DNA and not HJ, consistent with relative affinities for these substrates (Table 1). Moreover, the evidence that G4 DNA and HJ compete comparably for binding of full-length RecQ to G4 DNA suggests that full-length RecQ contains a single binding site for both these structured substrates, which is probably composed of elements from different domains.

The RQC domain recognizes distinct features of G4 DNA and forked substrates

In both G4 DNA and forked DNA, the 3′ tail forms a branch point at the juncture of the structured and single-stranded DNA regions. The demonstration that the RQC domain is not only a high affinity binding domain for both G4 DNA and forked DNA, but also appears to contain a single site for both substrates (Figure 5(a)), raised the question of whether the RQC domain recognizes this same feature on both substrates. To answer this question, we assayed binding of MBP-RQC and *E. coli* RecQ to G4 DNA with and without a 3′ single-stranded DNA tail. MBP-RQC bound to the blunt G4 DNA (OX1-G4, apparent k_D=3 nM) and tailed G4 DNA (OX1-T G4 DNA, apparent k_D= 6 nM) with comparable affinities (Figure 6(a)). Full-length RecQ similarly bound both OX1-G4 and OX1T-G4 with comparable affinities (k_D=6.5 nM and 4 nM, respectively; Figure 6(b)). Hence the RQC domain recognizes G-quadruplex structures directly.

Discussion

The ability to bind and unwind G4 DNA is a distinguishing feature of a subset of RecQ helicase family members. We have shown that the human RecQ family member, BLM, binds to G4 DNA within G-loops, a physiological target for this enzyme. We have identified the RQC domain as a conserved, independent and high affinity domain for binding forked DNA and G4 DNA. The RQC domain is restricted to RecQ family helicases, and appears key to the functions of enzymes in this family in maintenance of G-rich genomic regions, where G-quadruplexes may form in the course of transcription, replication, or recombination.

Other proteins and conserved domains have been shown to bind to G4 DNA, but they are structurally distinct from the RQC. RRM (RBD) domains and RGG domains, which are common among RNA binding proteins but also found in DNA binding proteins, both bind to G4 DNA[52,53] but have distinct modes of G4 DNA binding.[54] Hop1p, required for

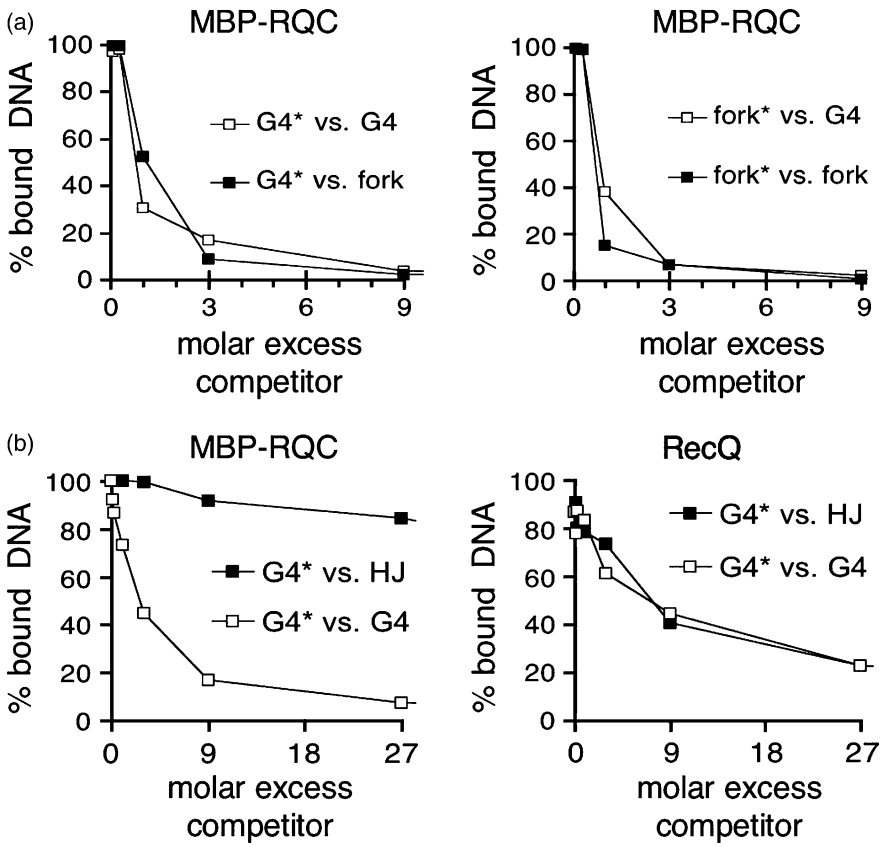

Figure 5. Overlapping binding sites for structured DNA substrates. (a) Quantifications of representative competition analysis of MBP-RQC binding to G4 DNA and forked substrates are graphed. An asterisk in the graph legend indicates that this DNA substrate was radiolabeled. Essentially identical results were obtained in three independent experiments. (b) Quantifications of competition analysis of binding of MBP-RQC or full-length RecQ to G4 DNA or HJ are graphed. Notations are as in (a). Essentially identical results were obtained in three independent experiments.

meiosis in *S. cerevisiae*, contains a zinc-finger motif and binds G4 DNA with high affinity.[55] Telomere end-binding proteins that contain OB folds regulate formation of G4 DNA at the chromosome ends.[37,38,56] The mismatch repair factor MutSα (the MSH2/MSH6 heterodimer) binds with nanomolar affinity to G4 DNA.[49] Despite the shared ability to bind G4 DNA, these factors are structurally distinct. It will be of great interest to learn how all these very different factors recognize G4 DNA with considerable specificity and affinity.

The RQC domain is composed of a Zn^{2+}-binding subdomain and a winged helix (WH) domain with similarities to those found in other DNA binding proteins.[51] A single point mutation in the WH domain of the human WRN protein drastically decreases the affinity of this helicase for forked DNA.[45] Helix-turn-helix motifs like the WH subdomain play their major role in sequence-specific recognition of *B*-form DNA,[57] but WH domains can also bind DNA in a sequence-independent, structure-specific manner as exemplified by the DNA repair enzyme *O*6-alklyguanine-DNA alkytransferase (AGT).[58] AGT interrogates the minor groove for damage with its WH domain, where an arginine within a recognition helix in the WH domain flips out damaged bases for *in situ* repair.[58] If RecQ unwinds DNA by an active mechanism, the WH domain might not only bind DNA in a structure-specific manner, but also promote strand separation by destabilizing base-pairing in the structured stem of the DNA substrate.

In addition to the RQC domain, most RecQ family helicases contain a second DNA binding domain, the HRDC domain. This domain has been shown to bind Holliday junctions and ssDNA.[44,46,47] The HRDC domain can also bind G4 DNA and forked DNA, but in all cases binding affinity is not comparable to that of full-length RecQ (Table 1). The HRDC is important for dissolution of double HJ,[59] and recombinant Sgs1p lacking the HRDC domain bound forked DNA[60] and G4 DNA[48] better than HJ, pointing to an important role of this auxiliary domain in HJ recognition. Competition experiments indicated only a single functional binding site for both G4 DNA and HJ in *E. coli* RecQ (Figure 5(b)), supporting the view that the HRDC domain may be an auxiliary DNA binding domain, which might interact with other portions of the helicase to form a single, high-affinity binding site for HJ.[46,59,61] By analogy to the p53 tumor suppressor, where modification of the C-terminal domain regulates sequence-specific DNA binding by the p53 central core domain,[62] such an interplay between HRDC and the helicase core domain could be regulated by protein–protein interactions or

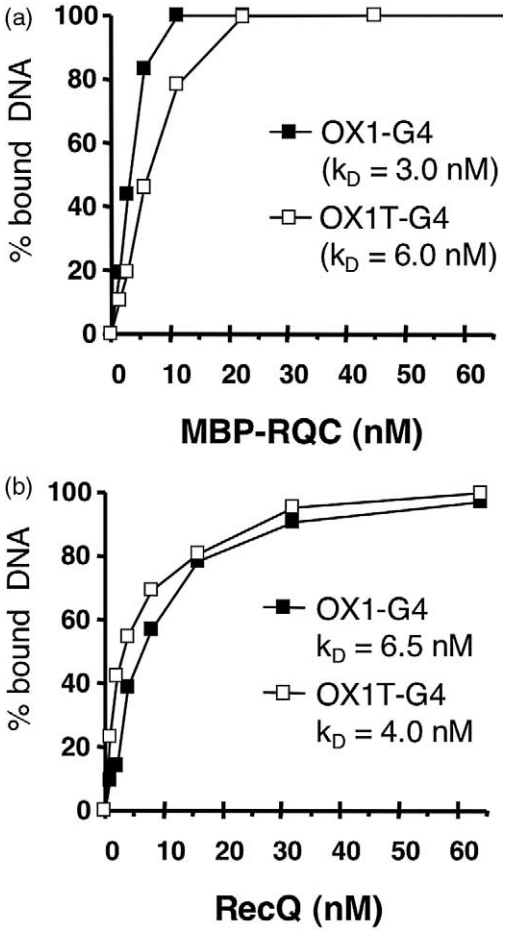

Figure 6. The RQC domain recognizes distinct features of G4 DNA. Quantification of representative analyses of binding of MBP-RQC and full-length *E. coli* RecQ to G4 DNA with (OX1T) and without (OX1) a 7 nt single-stranded 3′ tail. Apparent affinities (k_D) are shown as part of the graph legend. Essentially identical results were obtained in three independent experiments analyzing binding by MBP-RQC; and two analyses of binding by RecQ.

post-translational modifications. Future experiments will determine if such mechanisms can modulate substrate preference by RecQ helicases.

Implications for our understanding of other RecQ family members

Assignment of substrate-specific binding to the RQC and HRDC domains may provide mechanistic insight into two related RecQ family helicases, RecQ4 and RecQ5, which have not been as intensely studied as BLM and WRN. RecQ4 lacks both the RQC and HRDC domains (Figure 1). This provides a probable explanation for the inability of purified, recombinant RecQ4 to unwind a panel of DNA substrates assayed *in vitro*.[63] Nonetheless, RecQ4 is likely to participate in DNA transactions, as it has a DNA-dependent ATPase activity,[63] and is associated with the human genetic disease Rothmund-Thomson syndrome, which is characterized by genomic instability.[6,64] RecQ4 may have a DNA binding domain distinct from those identified in other RecQ family helicases; or it may interact with other factors that provide a domain (or domains) for DNA recognition. The *RECQ5* gene gives rise to several polypeptides as the result of alternative RNA processing (Figure 1). The largest, RecQ5β, contains the RQC domain, but not the HRDC domain. *RECQ5*-deficiency has not yet been associated with human disease, although mutational analysis in both chicken[8] and nematode[7] suggests functions in maintenance of genomic stability. RecQ5β may be specialized to unwind G4 DNA or forked substrates, a possibility that can be addressed as more is learned about the function of RecQ5 in human disease.

Materials and Methods

Oligonucleotides and DNA substrates

Oligonucleotides purchased from Operon Biotechnologies, Inc. (Huntsville, AL) were:

HJS1 (CACCCTTTTCTATACCTATCATGATGTGACC),
HJS2 (TCATTTGTTTTTTTTCACACTGCTGGAATTTTC),
HMH131 (AAAAGCAAGCTTTCAATCGTAAAACAGC ATCGCTTCC),
HMH132 (AAAAGCAAGCTTTCACACGATACGAGGC ACGGC), and
HMH133 (AAAAGGATCCGATCCGGCTGATATGGCG).

Oligonucleotides purchased from the Keck facility (Yale University, New Haven, CT) were:

HLD5 (AAAAAGGATCCAGAAATCGTTCGGCG),
HLD6 (AAAAAAAGCTTTCACTCTTCGTCATCGCCATC),
HMH3 (CAGGCCTGATTCACTTGATGGC),
HMH4 (AAAAAGCGCTAGCACCACCTCCAGTCGGCATC AGG),
HMH5 (AAAAACGGCTAGCTTGTGTTACCAGCTCCCT GC),
HMH6 (GGAACAGAAGGAAACTTCTGGCG), and
HMH72 (AAAAAGCGGATCCGCTCAGGCGGAAGTGTTG AATC).

G4 DNA was formed from oligonucleotides TP, OX1 and OX1T, as described.[29] Experiments used G4 DNA generated from the TP oligonucleotide unless otherwise indicated. Partial-duplex (forked) H1K1 substrate and the synthetic Holliday junction substrate HJ were formed, purified and stored as published.[48] Substrates for binding and unwinding assays were 5′ end-labeled with [γ-^{32}P]ATP (Amersham, Piscataway, NJ) and T4 polynucleotide kinase (MBI Fermentas, Hanover, MD), and unincorporated label was removed by purification over Sephadex G-50 resin.

Protein expression constructs

Plasmid pJK1-K695A for expression of BLM^{K695A} was created by PCR amplification of two fragments of BLM overlapping at the residue targeted with primer pairs HMH3/4 and HMH5/6. The PCR products were digested with NheI, gel purified, ligated, and re-amplified

with flanking primers HMH3 and HMH6. The resulting mutagenic cassette was digested with AvrII and DraIII to replace the corresponding wild-type fragment in pJK1.[65]

To generate plasmids for the expression of MBP fusions, DNA products encoding the conserved helicase domain from *E. coli* RecQ (Hel; RecQ residues 2–341), the RQC domain (RQC; RecQ residues 341–516), or the HRDC domain (HRDC; RecQ residues 521–607; Figure 3(a)) were PCR amplified with primer pairs HMH72/131, HMH132/133, and HLD5/6, respectively, using pEG88 plasmid DNA as the template.[31] Forward primers incorporated a BamHI site positioned to allow N-terminal in-frame fusion of the amplified coding region with MBP in pMAL-c2x (NEB, Ipswich, MA). Reverse primers contained a stop codon to terminate the open reading frame, and a HindIII restriction site for directional cloning. Ligation of digested, gel-purified PCR fragments into the BamHI/HindIII backbone of pMAL-c2x generated plasmids pMAL-Hel, pMAL-RQC, and pMAL-HRDC for expression of MBP-Hel, MBP-RQC, and MBP-HRDC fusion polypeptides (Figure 3(a)).

To express the BLM RQC domain, a DNA fragment encoding BLM residues 994–1202 was PCR amplified using primers HJS1 and HJS2. The resulting PCR fragment was transferred by directional topoisomerase ligation into expression vector pET100 (Invitrogen, Carlsbad, CA) to allow for expression of the N-terminally hexahistidine tagged BLM RQC fragment in *E. coli*. All constructs were verified by DNA sequencing.

Enzyme preparation

Full-length *E. coli* RecQ was expressed in *E. coli* strain BL21(DE3)-RP (Stratagene, La Jolla, CA) and purified as described.[31] Wild-type BLM and catalytically inactive BLM^{K695A} were expressed and purified from yeast as described.[48] The ability of BLM^{K695A} to bind DNA in the absence of ATPase and helicase activity was confirmed by standard DNA binding, helicase, and ATPase assays (data not shown).

MBP-fusion proteins (Figure 3(a)) were expressed in *E. coli* strain XL-1 blue (Stratagene, La Jolla, CA) and purified by affinity chromatography with amylose resin (NEB, Ipswich, MA), as recommended by the vendor. The BLM RQC fragment was expressed at 16 °C for 8 h after induction with 0.2 mM IPTG and purified under non-denaturing conditions. Briefly, cells were lysed by sonication in 50 mM phosphate buffer containing 400 mM NaCl, 0.5% (v/v) Triton X-100, 10 mM imidazole, and 5% (v/v) glycerol. The lysate was cleared by centrifugation at 25,000*g* for 1 h, then 500 μl of HIS-Select nickel resin (Sigma, St. Louis, MO) was added to the cleared lysate and the suspension was incubated at 4 °C for 1 h. The resin was separated from the liquid by gravity flow through a Poly-Prep chromatography column (BioRad, Hercules, CA) and washed extensively with buffer containing 20 mM imidazole. Bound protein was eluted with buffer containing 200 mM imidazole and fractions were collected. Purity of protein preparations was assessed by SDS-PAGE, and protein concentrations were determined by Bradford assays using BSA as the standard.

Electron microscopic analysis

Plasmid pPH600, containing a portion of the murine Sγ3 switch region sequence, was prepared and transcribed *in vitro* as described.[34] Transcribed pPH600 was separated from mononucleotides and T7 RNA polymerase, and the transcription buffer was replaced with 20 mM Hepes–NaOH (pH 7.4), 40 mM KCl, by filtration and washing of the reactions over a 100,000 Da molecular mass cut-off Microcon column (Millipore, Bedford, MA) according to the manufacturer's protocol. Binding reactions with BLM^{K695A} were carried out by addition of 15 nM helicase in the presence of 2 mM DTT. Reactions were incubated at 37 °C for 30 min and crosslinked in 0.2% EM grade glutaraldehyde for 20 min at 37 °C (Sigma-Aldrich, St. Louis, MO). Glutaraldehyde was removed by Microcon centrifugation, as described above, exchanging the buffer with 10 mM Tris–HCl (pH 7.4), 10 mM KCl. Plasmids were linearized with restriction endonuclease AflIII (NEB, Ipswich, MA) in buffer containing 100 mM NaCl, 10 mM Tris–HCl, 10 mM $MgCl_2$, 1 mM DTT. Samples were visualized by transmission electron microscopy.[34]

DNA binding assays

Binding reactions[48] were carried out in 20 μl reactions containing 50 mM Tris–HCl (pH 7.4), 2 mM $MgCl_2$, 50 mM NaCl, 100 μg/ml of BSA, 1 mM DTT, 0.5 nM ^{32}P-labeled DNA, and 10 nM enzyme (unless otherwise indicated), and were incubated at room temperature for 30 min, chilled for 5 min on ice and, following addition of 0.1 volume of buffered 50% glycerol, free nucleic acids and nucleoprotein complexes were resolved by electrophoresis on 6% (w/v) gels in 0.5×TBE, 10 mM KCl, at 6.5 V/cm for 2.5 h. Binding was quantified by phosphorimaging for calculation of apparent binding affinities (k_D). Experiments were repeated at least two and typically three or more times, and representative results are shown in each Figure.

Acknowledgements

We thank our colleagues for many helpful discussions. This research was supported by NIH R01 GM65988 and NCI P01 CA77852.

References

1. Bachrati, C. Z. & Hickson, I. D. (2003). RecQ helicases: suppressors of tumorigenesis and premature aging. *Biochem. J.* **374**, 577–606.
2. Hickson, I. D. (2003). RecQ helicases: caretakers of the genome. *Nature Rev. Cancer*, **3**, 169–178.
3. Kaneko, H. & Kondo, N. (2004). Clinical features of Bloom syndrome and function of the causative gene, BLM helicase. *Expert Rev. Mol. Diagn.* **4**, 393–401.
4. Opresko, P. L., Cheng, W. H. & Bohr, V. A. (2004). Junction of RecQ helicase biochemistry and human disease. *J. Biol. Chem.* **279**, 18099–18102.
5. Goss, K. H., Risinger, M. A., Kordich, J. J., Sanz, M. M., Straughen, J. E., Slovek, L. E. *et al.* (2002). Enhanced tumor formation in mice heterozygous for Blm mutation. *Science*, **297**, 2051–2053.
6. Shimamoto, A., Nishikawa, K., Kitao, S. & Furuichi, Y. (2000). Human RecQ5beta, a large isomer of RecQ5 DNA helicase, localizes in the nucleoplasm and interacts with topoisomerases 3alpha and 3beta. *Nucl. Acids Res.* **28**, 1647–1655.

7. Jeong, Y. S., Kang, Y., Lim, K. H., Lee, M. H., Lee, J. & Koo, H. S. (2003). Deficiency of *Caenorhabditis elegans* RecQ5 homologue reduces life span and increases sensitivity to ionizing radiation. *DNA Repair (Amst.)*, **2**, 1309–1319.
8. Wang, W., Seki, M., Narita, Y., Nakagawa, T., Yoshimura, A., Otsuki, M. *et al*. (2003). Functional relation among RecQ family helicases RecQL1, RecQL5, and BLM in cell growth and sister chromatid exchange formation. *Mol. Cell. Biol.* **23**, 3527–3535.
9. Garcia, P. L., Liu, Y., Jiricny, J., West, S. C. & Janscak, P. (2004). Human RECQ5beta, a protein with DNA helicase and strand-annealing activities in a single polypeptide. *EMBO J.* **23**, 2882–2891.
10. Imamura, O., Fujita, K., Itoh, C., Takeda, S., Furuichi, Y. & Matsumoto, T. (2002). Werner and Bloom helicases are involved in DNA repair in a complementary fashion. *Oncogene*, **21**, 954–963.
11. Du, X., Shen, J., Kugan, N., Furth, E. E., Lombard, D. B., Cheung, C. *et al*. (2004). Telomere shortening exposes functions for the mouse Werner and Bloom syndrome genes. *Mol. Cell. Biol.* **24**, 8437–8446.
12. Spillare, E. A., Wang, X. W., von Kobbe, C., Bohr, V. A., Hickson, I. D. & Harris, C. C. (2005). Redundancy of DNA helicases in p53-mediated apoptosis. *Oncogene*, E-pub ahead of print.
13. Cohen, H. & Sinclair, D. A. (2001). Recombination-mediated lengthening of terminal telomeric repeats requires the Sgs1 DNA helicase. *Proc. Natl Acad. Sci. USA*, **98**, 3174–3179.
14. Huang, P., Pryde, F. E., Lester, D., Maddison, R. L., Borts, R. H., Hickson, I. D. & Louis, E. J. (2001). SGS1 is required for telomere elongation in the absence of telomerase. *Curr. Biol.* **11**, 125–129.
15. Johnson, F. B., Marciniak, R. A., McVey, M., Stewart, S. A., Hahn, W. C. & Guarente, L. (2001). The *Saccharomyces cerevisiae* WRN homolog Sgs1p participates in telomere maintenance in cells lacking telomerase. *EMBO J.* **20**, 905–913.
16. Chang, S., Multani, A. S., Cabrera, N. G., Naylor, M. L., Laud, P., Lombard, D. *et al*. (2004). Essential role of limiting telomeres in the pathogenesis of Werner syndrome. *Nature Genet.* **36**, 877–882.
17. Crabbe, L., Verdun, R. E., Haggblom, C. I. & Karlseder, J. (2004). Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. *Science*, **306**, 1951–1953.
18. Opresko, P. L., Otterlei, M., Graakjaer, J., Bruheim, P., Dawut, L., Kolvraa, S. *et al*. (2004). The Werner syndrome helicase and exonuclease cooperate to resolve telomeric D loops in a manner regulated by TRF1 and TRF2. *Mol. Cell.* **14**, 763–774.
19. Laud, P. R., Multani, A. S., Bailey, S. M., Wu, L., Ma, J., Kingsley, C. *et al*. (2005). Elevated telomere-telomere recombination in WRN-deficient, telomere dysfunctional cells promotes escape from senescence and engagement of the ALT pathway. *Genes Dev.* **19**, 2560–2570.
20. Mandell, J. G., Goodrich, K. J., Bahler, J. & Cech, T. R. (2005). Expression of a RecQ helicase homolog affects progression through crisis in fission yeast lacking telomerase. *J. Biol. Chem.* **280**, 5249–5257.
21. Sinclair, D. A., Mills, K. & Guarente, L. (1997). Accelerated aging and nucleolar fragmentation in yeast sgs1 mutants. *Science*, **277**, 1313–1316.
22. Sinclair, D. A. & Guarente, L. (1997). Extrachromosomal rDNA circles–a cause of aging in yeast. *Cell*, **91**, 1033–1042.
23. Marciniak, R. A., Lombard, D. B., Johnson, F. B. & Guarente, L. (1998). Nucleolar localization of the Werner syndrome protein in human cells. *Proc. Natl Acad. Sci. USA*, **95**, 6887–6892.
24. Cobb, J. A., Bjergbaek, L., Shimada, K., Frei, C. & Gasser, S. M. (2003). DNA polymerase stabilization at stalled replication forks requires Mec1 and the RecQ helicase Sgs1. *EMBO J.* **22**, 4325–4336.
25. Schawalder, J., Paric, E. & Neff, N. F. (2003). Telomere and ribosomal DNA repeats are chromosomal targets of the bloom syndrome DNA helicase. *BMC Cell Biol.* **4**, 15.
26. Versini, G., Comet, I., Wu, M., Hoopes, L., Schwob, E. & Pasero, P. (2003). The yeast Sgs1 helicase is differentially required for genomic and ribosomal DNA replication. *EMBO J.* **22**, 1939–1949.
27. Weitao, T., Budd, M. & Campbell, J. L. (2003). Evidence that yeast SGS1, DNA2, SRS2, and FOB1 interact to maintain rDNA stability. *Mutat. Res.* **532**, 157–172.
28. Kondo, N., Motoyoshi, F., Mori, S., Kuwabara, N., Orii, T. & German, J. (1992). Long-term study of the immunodeficiency of Bloom's syndrome. *Acta Paediatr.* **81**, 86–90.
29. Sun, H., Karow, J. K., Hickson, I. D. & Maizels, N. (1998). The Bloom's syndrome helicase unwinds G4 DNA. *J. Biol. Chem.* **273**, 27587–27592.
30. Sun, H., Bennett, R. J. & Maizels, N. (1999). The *Saccharomyces cerevisiae* Sgs1 helicase efficiently unwinds G-G paired DNAs. *Nucl. Acids Res.* **27**, 1978–1984.
31. Wu, X. & Maizels, N. (2001). Substrate-specific inhibition of RecQ helicase. *Nucl. Acids Res.* **29**, 1765–1771.
32. Gellert, M., Lipsett, M. N. & Davies, D. R. (1962). Helix formation by guanylic acid. *Proc. Natl Acad. Sci. USA*, **48**, 2013–2018.
33. Sen, D. & Gilbert, W. (1988). Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature*, **334**, 364–366.
34. Duquette, M. L., Handa, P., Vincent, J. A., Taylor, A. F. & Maizels, N. (2004). Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev.* **18**, 1618–1629.
35. Duquette, M. L., Pham, P., Goodman, M. F. & Maizels, N. (2005). AID binds to transcription-induced structures in c-MYC that map to regions associated with translocation and hypermutation. *Oncogene*, **24**, 5791–5798.
36. Schaffitzel, C., Berger, I., Postberg, J., Hanes, J., Lipps, H. J. & Pluckthun, A. (2001). In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei. *Proc. Natl Acad. Sci. USA*, **98**, 8572–8577.
37. Paeschke, K., Simonsson, T., Postberg, J., Rhodes, D. & Lipps, H. J. (2005). Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nature Struct. Mol. Biol.* **12**, 847–854.
38. Zaug, A. J., Podell, E. R. & Cech, T. R. (2005). Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc. Natl Acad. Sci. USA*, **102**, 10864–10869.
39. Singleton, M. R. & Wigley, D. B. (2002). Modularity and specialization in superfamily 1 and 2 helicases. *J. Bacteriol.* **184**, 1819–1826.
40. Gorbalenya, A. E. & Koonin, E. V. (1993). Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.* **3**, 419–429.

41. Bernstein, D. A. & Keck, J. L. (2003). Domain mapping of *Escherichia coli* RecQ defines the roles of conserved N- and C-terminal regions in the RecQ family. *Nucl. Acids Res.* **31**, 2778–2785.
42. Tanner, N. K., Cordin, O., Banroques, J., Doere, M. & Linder, P. (2003). The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol. Cell.* **11**, 127–138.
43. Cordin, O., Tanner, N. K., Doere, M., Linder, P. & Banroques, J. (2004). The newly discovered Q motif of DEAD-box RNA helicases regulates RNA-binding and helicase activity. *EMBO J.* **23**, 2478–2487.
44. von Kobbe, C., Thoma, N. H., Czyzewski, B. K., Pavletich, N. P. & Bohr, V. A. (2003). Werner syndrome protein contains three structure-specific DNA binding domains. *J. Biol. Chem.* **278**, 52997–53006.
45. Lee, J. W., Kusumoto, R., Doherty, K. M., Lin, G. X., Zeng, W., Cheng, W. H. *et al.* (2005). Modulation of Werner syndrome protein function by a single mutation in the conserved RQC domain. *J. Biol. Chem.* **280**, 39627–39636.
46. Liu, Z., Macias, M. J., Bottomley, M. J., Stier, G., Linge, J. P., Nilges, M. *et al.* (1999). The three-dimensional structure of the HRDC domain and implications for the Werner and Bloom syndrome proteins. *Struct. Fold. Des.* **7**, 1557–1566.
47. Bernstein, D. A. & Keck, J. L. (2005). Conferring substrate specificity to DNA helicases: role of the RecQ HRDC domain. *Structure (Camb)*, **13**, 1173–1182.
48. Huber, M. D., Lee, D. C. & Maizels, N. (2002). GH DNA unwinding by BLM and Sgs1p: substrate specificity and substrate-specific inhibition. *Nucl. Acids Res.* **30**, 3954–3961.
49. Larson, E. D., Duquette, M. L., Cummings, W. J., Streiff, R. J. & Maizels, N. (2005). MutSalpha binds to and promotes synapsis of transcriptionally activated immunoglobulin switch regions. *Curr. Biol.* **15**, 470–474.
50. Neff, N. F., Ellis, N. A., Ye, T. Z., Noonan, J., Huang, K., Sanz, M. & Proytcheva, M. (1999). The DNA helicase activity of BLM is necessary for the correction of the genomic instability of bloom syndrome cells. *Mol. Biol. Cell*, **10**, 665–676.
51. Bernstein, D. A., Zittel, M. C. & Keck, J. L. (2003). High-resolution structure of the *E. coli* RecQ helicase catalytic core. *EMBO J.* **22**, 4910–4921.
52. Hanakahi, L. A., Sun, H. & Maizels, N. (1999). High affinity interactions of nucleolin with G-G-paired rDNA. *J. Biol. Chem.* **274**, 15908–15912.
53. Dempsey, L. A., Sun, H., Hanakahi, L. A. & Maizels, N. (1999). G4 DNA binding by LR1 and its subunits, nucleolin and hnRNP D, A role for G-G pairing in immunoglobulin switch recombination. *J. Biol. Chem.* **274**, 1066–1071.
54. Cocco, M. J., Hanakahi, L. A., Huber, M. D. & Maizels, N. (2003). Specific interactions of distamycin with G-quadruplex DNA. *Nucl. Acids Res.* **31**, 2944–2951.
55. Anuradha, S. & Muniyappa, K. (2004). *Saccharomyces cerevisiae* Hop1 zinc finger motif is the minimal region required for its function in vitro. *J. Biol. Chem.* **279**, 28961–28969.
56. Lei, M., Zaug, A. J., Podell, E. R. & Cech, T. R. (2005). Switching human telomerase on and off with hPOT1 protein in vitro. *J. Biol. Chem.* **280**, 20449–20456.
57. Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M. & Iyer, L. M. (2005). The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol. Rev.* **29**, 231–262.
58. Daniels, D. S., Woo, T. T., Luu, K. X., Noll, D. M., Clarke, N. D., Pegg, A. E. & Tainer, J. A. (2004). DNA binding and nucleotide flipping by the human DNA repair protein AGT. *Nature Struct. Mol. Biol.* **11**, 714–720.
59. Wu, L., Lung Chan, K., Ralf, C., Bernstein, D. A., Garcia, P. L., Bohr, V. A. *et al.* (2005). The HRDC domain of BLM is required for the dissolution of double Holliday junctions. *EMBO J.* **24**, 2679–2687.
60. Bennett, R. J., Keck, J. L. & Wang, J. C. (1999). Binding specificity determines polarity of DNA unwinding by the Sgs1 protein of *S. cerevisiae*. *J. Mol. Biol.* **289**, 235–248.
61. Janscak, P., Garcia, P. L., Hamburger, F., Makuta, Y., Shiraishi, K., Imai, Y. *et al.* (2003). Characterization and mutational analysis of the RecQ core of the bloom syndrome protein. *J. Mol. Biol.* **330**, 29–42.
62. Appella, E. & Anderson, C. W. (2001). Post-translational modifications and activation of p53 by genotoxic stresses. *Eur. J. Biochem.* **268**, 2764–2772.
63. Macris, M. A., Krejci, L., Bussen, W., Shimamoto, A. & Sung, P. (2005). Biochemical characterization of the RECQ4 protein, mutated in Rothmund-Thomson syndrome. *DNA Repair* (*Amst.*), **5**, 172–180.
64. Kitao, S., Ohsugi, I., Ichikawa, K., Goto, M., Furuichi, Y. & Shimamoto, A. (1998). Cloning of two new human helicase genes of the RecQ family: biological significance of multiple species in higher eukaryotes. *Genomics*, **54**, 443–452.
65. Karow, J. K., Chakraverty, R. K. & Hickson, I. D. (1997). The Bloom's syndrome gene product is a 3′-5′ DNA helicase. *J. Biol. Chem.* **272**, 30611–30614.

Edited by J. O. Thomas

(Received 8 December 2005; received in revised form 17 January 2006; accepted 20 January 2006)
Available online 7 February 2006