J. Biochem. 2010;148(4):517–525 doi:10.1093/jb/mvq097

JB THE JOURNAL OF BIOCHEMISTRY

Solution structure of the HRDC domain of human Bloom syndrome protein BLM

Received June 27, 2010; accepted August 23, 2010; published online August 25, 2010

Akiko Sato[1], Masaki Mishima[1], Aki Nagai[2], Sun-Yong Kim[2], Yutaka Ito[1], Toshio Hakoshima[2], Jun-Goo Jee[2] and Ken Kitano[2,*]

[1]Graduate School of Science and Engineering, Tokyo Metropolitan University, 1-1 Minami-Osawa, Hachioji, Tokyo 192-0397; and [2]Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan

*Ken Kitano, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan. Tel: +81 743 72 5573, Fax: +81 743 72 5579, email: kkitano@is.naist.jp

Accession Number: Atomic coordinates and NMR constraints for BLM HRDC have been deposited in the Protein Data Bank under ID code 2RRD (BMRB ID 11252).

Bloom syndrome is a rare genetic disorder characterized by severe growth retardation and cancer predisposition. The disease is caused by a loss of function of the Bloom syndrome protein (BLM), a member of the RecQ family of DNA helicases. Here we report on the first 3D structure of a BLM fragment, a solution structure of the C-terminal helicase-and-ribonuclease D-C-terminal (HRDC) domain from human BLM. The structure reveals unique features of BLM HRDC that are distinct from the HRDC domain of Werner syndrome protein. In particular, BLM HRDC retains many acidic residues exposed to the solvent, which makes the domain surface extensively electronegative. Consistent with this, fluorescence polarization assays showed an inability of isolated BLM HRDC to interact with DNA substrates. Analyses employing ultracentrifugation, gel-filtration, CD spectroscopy and dynamic light scattering showed that the BLM HRDC domain exists as a stable monomer in solution. The results show that BLM HRDC is a compact, robust and acidic motif which may play a distinct role apart from DNA binding.

Keywords: BLM/Bloom syndrome/HRDC domain/NMR structure/WRN.

Abbreviations: BLM, Bloom syndrome protein; CD, circular dichroism; DHJ, double Holliday junction; DLS, dynamic light scattering; DSB, double-strand break; DTT, dithiothreitol; FITC, fluorescein isothiocyanate; GST, glutathione-*S*-transferase; HRDC, helicase-and-ribonuclease D-C-terminal; pI, isoelectric point; IPTG, isopropyl-β-D-thiogalactoside; MALDI-TOF MS, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; MWCO, molecular weight cut off; r.m.s.d., root mean square deviations; RQC, RecQ C-terminal; WRN, Werner syndrome protein.

The RecQ helicases, a family of DNA unwinding enzymes conserved from prokaryotes to mammals, play a key role in protecting the genome against deleterious changes [for recent reviews (*1–3*)]. Whereas genomes of bacteria and unicellular eukaryotes typically encode a single recQ gene, the human genome contains five recQ genes comprising RECQ1, Bloom syndrome protein (BLM), Werner syndrome protein (WRN), RECQ4 and RECQ5. Mutations in BLM and WRN are associated with hereditary disorders referred to as Bloom and Werner syndromes, respectively. Although the two diseases are commonly characterized by pronounced genomic instability and cancer predisposition, the clinical symptoms of Bloom and Werner syndromes are apparently distinct. Bloom patients display severe growth retardation with short stature, immunodeficiency, sunlight sensitivity, and a predisposition to the development of many different types of malignancies, while Werner patients display features of accelerated aging (*1*, *2*, *4*). The different clinical features of the disorders and the fact that the functional loss of BLM or WRN cannot be compensated for by the presence of the other protein (or other RecQ members) support the notion that BLM and WRN have distinct functions in cells.

BLM and WRN share three structurally-folded domains comprising an ATPase domain, a RecQ C-terminal (RQC) domain, and a helicase-and-ribonuclease D-C-terminal (HRDC) domain (Fig. 1A). WRN also possesses an exonuclease domain at the N-terminus. The ATPase domain is responsible for ATP-dependent DNA translocation, while the RQC winged-helix domain appears to be a major DNA binding site of RecQ family proteins. Our recent crystal structure of a WRN RQC-DNA complex has delineated the importance of the RQC domain in recognition, binding and unwinding of DNA at branched points (*5*). On the other hand, the function of the HRDC domain remains unknown. The HRDC domain is one of the most divergent components of RecQ family proteins. In human RecQs, only BLM and WRN share HRDC sequences at the C-terminus, whereas the other three RecQ members completely lack HRDC sequences. The existing data concerning BLM (*6*) and WRN (*7*) suggest that the HRDC domain is not essential for conventional helicase activity on forked duplexes, and is largely dispensable. Considering that the HRDC sequences of BLM and WRN share only ~20% identity (Fig. 1B), this heterogeneity in HRDCs may confer functional differences between the two proteins and hence be associated with the clinical differences manifested by Bloom and Werner syndromes.

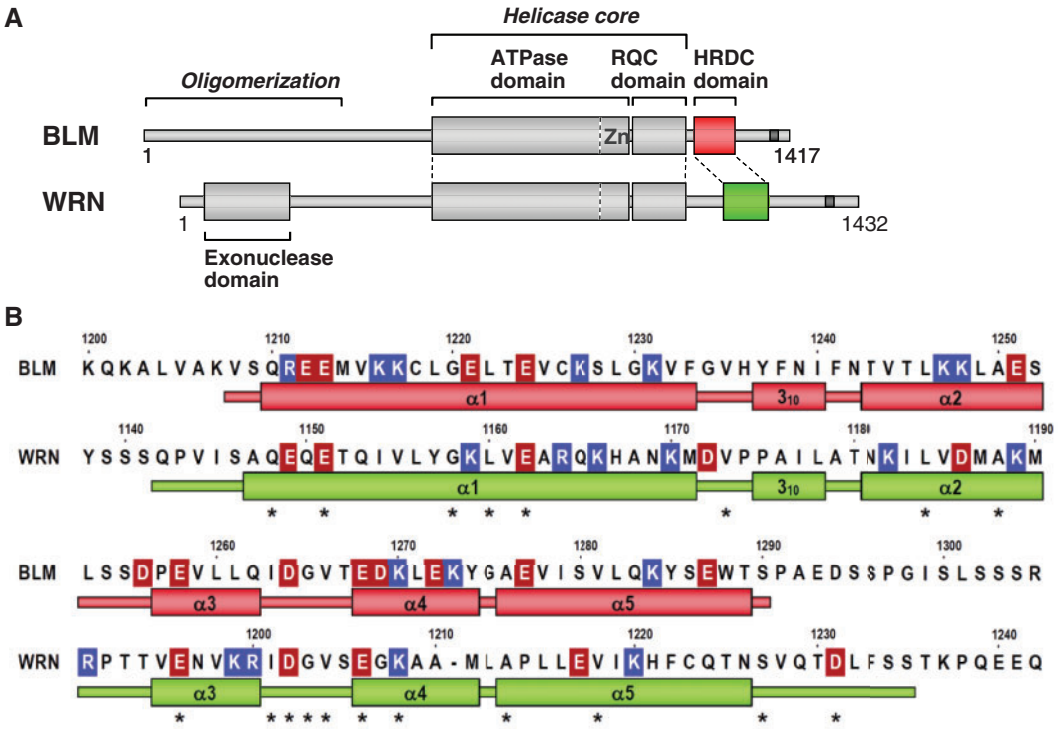

Fig. 1 Amino acid sequences of the HRDC domains of BLM and WRN. (A) Domain diagram of human BLM and WRN. HRDC domains are coloured in red (BLM) and green (WRN). The nuclear localization signal is depicted as a shaded bar. The zinc-binding motif (Zn), which is unique to the RecQ family of proteins and is often combined with the RQC sequence, is depicted in the C-terminus of the ATPase domain based on the latest structural data (*3*, *8*). The N-terminal region of BLM (amino acids 1–431) may be involved in the oligomerization (*22*). The entire sequence identity between BLM and WRN is ~15%, while the identities within the ATPase, RQC and HRDC domains are ~30, ~10 and ~20%, respectively. (B) Sequence alignment of BLM and WRN HRDC domains. Acidic (Asp or Glu) and basic (Arg or Lys) residues are highlighted in red and blue, respectively. Secondary-structure elements (α1–α5) are shown below each sequence as red (BLM; present work) and green [WRN (*20*)] thick bars. Loop regions are shown as thin bars. Residues conserved between the two proteins are denoted by an asterisk.

In contrast to WRN, for which several 3D structures of protein fragments are available [reviewed in refs (*3*, *8*)], structural studies of BLM have been hampered by difficulties in both protein purification and crystallization. Here we report on the first biochemical characterization of the isolated BLM HRDC domain and its structure determination by solution NMR spectroscopy.

Experimental Procedures

Expression and purification of the BLM HRDC domain

DNA encoding BLM HRDC (amino acids 1,200–1,295) was cloned from human BLM and inserted into the glutathione-*S*-transferase (GST)-fusion vector pGEX-5X-1 (GE Healthcare Bio-Sciences). The fidelity of the HRDC coding region was confirmed by DNA sequencing. The plasmid was transformed into *Escherichia coli* strain BL21-CodonPlus RIL (Stratagene), which was grown in LB medium supplemented with 100 μg/ml ampicillin and 50 μg/ml chloramphenicol at 37°C to an OD_{660} of 0.6. Expression was induced by the addition of isopropyl-β-D-thiogalactoside (IPTG) to 1 mM. Following incubation for another 5 h at 25°C, cells were pelleted by centrifugation and washed with 0.9% (w/v) NaCl solution.

For purification, cells were resuspended in 50 mM Tris–HCl (pH 7.5), 200 mM NaCl, 5 mM dithiothreitol (DTT) and 0.5% Triton X-100, and disrupted by sonication in an ice bath. The cell suspension was clarified by ultracentrifugation at 37,000 r.p.m. for 30 min. The supernatant, which contains GST-BLM HRDC, was loaded onto a glutathione sepharose-4B affinity column (GE Healthcare) equilibrated with 50 mM Tris–HCl (pH 7.5), 100 mM NaCl and 5 mM DTT. The column was washed with the same buffer until protein was undetectable in the eluent, and subsequently re-equilibrated with digestion buffer comprising 50 mM Tris–HCl (pH 8.0), 100 mM NaCl, 5 mM DTT and 5 mM $CaCl_2$. Protein bound to the resin was digested on the column with 2 U/ml Factor Xa (Novagen) for 15 h at 20°C. The column effluent containing GST-free HRDC was pooled, loaded onto a HiTrap SP cation-exchange column (GE Healthcare), and protein was eluted using a 100–1,000 mM NaCl gradient. The effluent fraction was pooled and concentrated by centrifugation using an Amicon Ultra 10,000 MWCO (molecular weight cut off) filter. Protein was finally passed through a HiLoad 26/60 Superdex 75 gel filtration column (GE Healthcare) equilibrated with 50 mM Tris–HCl (pH 7.5), 100 mM NaCl and 1 mM DTT (buffer A). Peak fractions were pooled and analysed by SDS–PAGE, which gave one major band corresponding to ~10 kDa (Fig. 2A). Ten milligrams of purified HRDC was obtained from 1 l of culture. The sample was further analysed by N-terminal sequencing and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS; Bruker), which confirmed that the protein had been successfully purified without degradation. Five vector-derived residues (Gly–Ile–Pro–Glu–Phe) were present at the N-terminus. The sample was frozen in liquid nitrogen and stored at −80°C until use.

Circular dichroism

Protein sample was prepared at a concentration of 0.1 mg/ml in buffer A (400 μl), and analysed at 10, 20, 30 and 40°C using a JASCO 720W spectrometer (Japan). The solvent spectrum was subtracted from that of the protein solution. Measurements were repeated eight times and averaged. Analyses of the spectra were performed using Spectra Manager software (JASCO, Japan).

Ultracentrifuge sedimentation equilibrium

Purified BLM HRDC was dialysed against a solution containing 10 mM Tris–HCl (pH 7.5), 100 mM NaCl and 0.1 mM DTT,

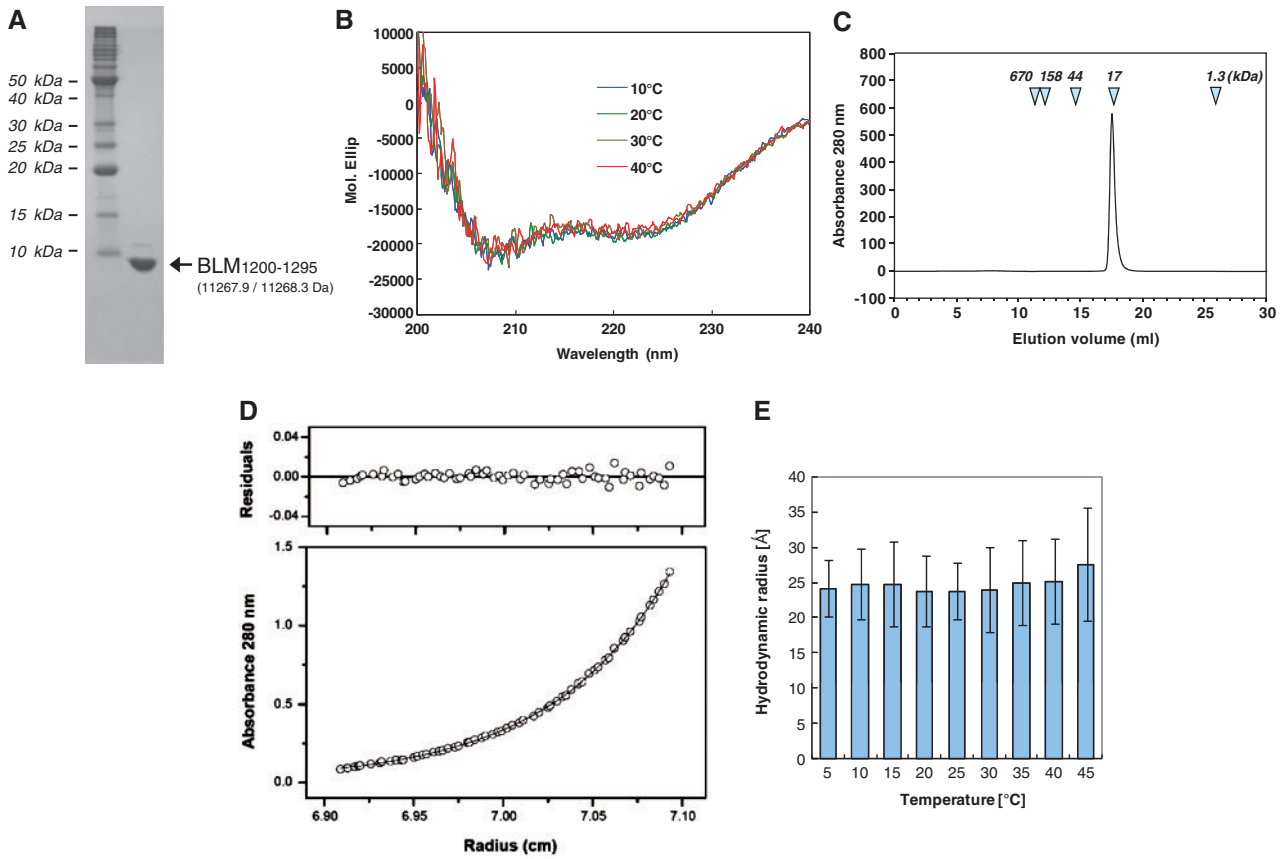

Fig. 2 Characterization of the purified BLM HRDC domain. (A) SDS–PAGE of purified BLM HRDC (amino acids 1,200–1,295). The sample was electrophoresed through a 15% polyacrylamide gel, stained with Coomassie blue, and further characterized by N-terminal sequencing and TOF-MS. The mass calculated from the primary sequence and the mass observed by TOF-MS are shown in parentheses. (B) CD spectrum of BLM HRDC at temperatures ranging from 10 to 40°C. (C) Analytical gel-filtration chromatography of BLM HRDC. Purified BLM HRDC was loaded onto a Superdex 75 10/300 column (GE Healthcare) at 4°C equilibrated with solution containing 10 mM Tris–HCl (pH 7.5), 150 mM NaCl and 1 mM DTT, and elution was monitored at 280 nm. Elution volumes of the molecular-weight markers (1.3, 17, 44, 158 and 670 kDa, Bio-Rad) are indicated by arrows. (D) Ultracentrifuge sedimentation equilibrium of BLM HRDC. Observed absorbance at 33 μM protein concentration at 20°C is plotted against the radial distance. The best-fit curve from a single-species model is overlaid, which yielded a molecular mass of 12.9 kDa. Comparable results were obtained for 77 and 154 μM protein concentrations. (E) Dynamic light scattering of BLM HRDC at temperatures ranging from 5 to 45°C. The hydrodynamic radius obtained from the measurements is shown as bars with standard deviations. The protein precipitated at temperatures beyond 50°C.

followed by adjustment of the absorbance at 280 nm to the three protein concentrations 0.33, 0.77 and 1.54 (corresponding to 33, 77 and 154 μM BLM HRDC, respectively). The sample solution and reference buffer were injected into each side of a six-sector centrepiece. Ultracentrifugation was performed at 20°C and 38,000 r.p.m. using an Optima XL-A analytical ultracentrifuge (Beckman Instruments) equipped with an An-60ti rotor. Following equilibration (50–60 h), the radial absorbance at 280 nm was scanned, which was subsequently fitted to the calculated curve from the ideal single-species model. The partial specific volume of the protein at 20°C was calculated to be 0.742 ml/g based on the amino acid composition, while the solvent density was calculated to be 1.0026 g/ml.

Dynamic light scattering

Dynamic light scattering (DLS) of BLM HRDC was measured using a DynaPro-801 equipped with DYNAMICS software (Protein Solutions). The sample was prepared at 5 mg/ml protein concentration (corresponding to ~450 μM) in buffer A, and filtered using a 0.02 μm membrane. Measurements were performed at temperatures ranging from 5 to 55°C with increments of 5°C. Each measurement was repeated 15 times and averaged.

NMR spectroscopy and structure determination

Uniformly ^{15}N- and ^{15}N/^{13}C-labelled proteins were prepared by growing bacteria in M9 minimal medium containing $^{15}NH_4Cl$ with or without [$^{13}C_6$] glucose. The protein was purified using the same protocol as for the non-labelled proteins, and finally dissolved in 50 mM Tris-buffer (pH 7.5) containing 100 mM NaCl and 1 mM DTT in either 95% (v/v) H_2O/5% ^2H_2O or 99.8% ^2H_2O. For NOESY experiments, d11-Tris was used to eliminate large ^{1}H signals originating from the Tris moiety. Final protein concentrations were 0.5–1.2 mM. NMR spectra were acquired at 30°C on a Bruker AVANCE 600 NMR spectrometer equipped with a cryogenic probe. All NMR spectra were processed with NMRPipe (*9*) and analysed using Sparky (*10*). The ^{1}H, ^{13}C and ^{15}N assignments were mainly obtained from standard multi-dimensional NMR methods (*11*, *12*). HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HN(CA)CO and HNCO experiments were performed for main-chain assignments, and C(CO)NH, H(CCO)NH, HCCH-TOCSY and 4D HC(CO)NH for side-chain assignments. Inter-proton distances were derived from 3D ^{15}N edited NOESY-HSQC and 3D ^{13}C edited NOESY-HSQC. Backbone dihedral ϕ and ψ angles derived from TALOS were also used (*13*). The χ1 rotamers of the side-chain were estimated from HNHB and HN(CO)HB experiments (*14*, *15*).

The CYANA programme was used with the CANDID protocol for the purpose of structural restraint collection and efficient NOESY cross peak assignments (*16*). Finally, an ensemble of 100 HRDC structures was calculated using the programme CNS with a standard simulated annealing protocol (*17*). The final lowest-energy 20 ensemble structures were checked by PROCHECK-NMR (*18*). Detailed statistics of the structures are

summarized in Table I. Figures were prepared using PyMOL (DeLano Scientific, CA, USA). Superimposition for Fig. 3E was performed using LSQMAN (*19*).

Fluorescence polarization assays

Fluorescein isothiocyanate (FITC)-labelled oligonucleotides representing various DNA structures were prepared as described earlier (*5*, *20*). Purified BLM HRDC in the GST-free form (Fig. 2A) was mixed with each substrate in a solution containing 10 mM Tris–HCl (pH 7.5), 25 mM NaCl and 1 mM DTT. The final protein concentration ranged from 1 nM to 10 μM, while the DNA concentration in all samples was 100 pM. The protein–DNA interaction was examined by fluorescence polarization using a full-range Beacon 2000 system (Invitrogen) with 490 and 535 nm excitation and emission wavelengths, respectively. Each measurement was performed at 25°C, repeated five times, and then averaged. Throughout the experiments, the WRN RQC (*5*) and WRN HRDC (*20*) domains were utilized as positive and negative controls, respectively.

Results and Discussion

Purification of BLM HRDC

We previously determined the crystal structure of the WRN HRDC domain (*20*). Based on the sequence alignment with WRN (Fig. 1B), an expression construct for the BLM HRDC domain was generated, which encoded amino acids 1,195–1,304 of human BLM. The fragment was overexpressed in *E. coli* and purified in the GST-cleaved form through several chromatographic steps, however, SDS–PAGE analysis showed that the protein had spontaneously degraded into a smaller fragment during the purification (Supplementary Fig. S1A). The combination of N-terminal sequencing and TOF-MS further identified the resulting fragment as BLM 1,200–1,294, which lacks the original 5 and 10 amino acids at the N- and C-terminus, respectively. Sequence analyses of human, mouse and *Xenopus* BLMs also showed that the central amino acids 1,200–1,294 of BLM represent a conserved cluster region, whereas the outside residues are much less conserved (Supplementary Fig. S1B); the degraded terminal residues are likely to be unstructured in solution. Therefore, the second plasmid encoding BLM 1,200–1,295 was reconstructed, from which protein had been successfully purified to homogeneity without degradation (Fig. 2A).

Table I. Structural statistics for BLM HRDC[a].

Total number of distance constraints	2,314
Long range ($\|i–j\|>4$)	640
Middle range ($\|i–j\|=2,3,4$)	580
Short range ($\|i–j\|=1$)	556
Intra residue	538
Hydrogen bond constraints	27×2
Dihedral constraints	
ϕ, φ	79, 79
$\chi1$	15
r.m.s.d. from experimental constraints[b]	
Distance (Å)	$0.009\pm6\times10^{-4}$
Angle (°)	0.28 ± 0.04
r.m.s.d. from idealized covalent geometry	
Bonds (Å)	$0.001\pm5\times10^{-5}$
Angles (°)	0.274 ± 0.003
Impropers (°)	0.119 ± 0.005
CNS energy terms (kcal/mol)[c]	
E_{bond}	1.7 ± 0.17
E_{angle}	33.6 ± 0.75
E_{imp}	1.7 ± 0.13
$E_{vdw(LJ)}$	-297 ± 4
PROCHECK Ramachandran plot (residues 1,210–1,290)	
Residues in most favored regions (%)	93.1
Residues in additional allowed regions (%)	6.9
Residues in generously allowed regions (%)	0.0
Residues in disallowed regions (%)	0.0
r.m.s.d. of mean structure derived from 20 calculated structures	
Back bone (residues 1,210–1,290) (Å)	0.31
All heavy (residues 1,210–1,290) (Å)	0.77

[a]These statistics comprise an ensemble of the 20 lowest-energy structures obtained from 100 starting structures. Structure calculations were performed using CNS version 1.1.
[b]None of these structures exhibited distance violations >0.5 Å or dihedral angle violations >5°.
[c]E_{vdw} is the Lennerd–Jones energy of the CNS energy terms.

BLM HRDC is a stable monomer in solution

For structure determination, we firstly attempted a crystallization of the purified BLM HRDC domain (amino acids 1,200–1,295) using various commercial screening kits. However, in contrast to WRN HRDC which had been crystallized in two different crystal-forms by the initial screening trial (*20*), no crystals of the BLM HRDC domain could be obtained (only precipitation and/or aggregation appeared). The circular dichroism (CD) measurements at temperatures ranging from 10 to 40°C gave identical spectra (Fig. 2B), suggesting that BLM HRDC remains structured within this temperature range. The spectral analysis further revealed that BLM HRDC possesses a high α-helical (~62%) and low β-strand (~10%) content, not unlike the case with the WRN HRDC domain (*20*).

Full-length BLM in the DNA-free form was previously reported to form a hexameric and/or tetrameric ring structure (*21*), and it was suggested that the N-terminal region might be involved in the oligomerization (*22*). An interesting point here relates to whether the isolated BLM HRDC domain can also undergo oligomerization. We examined the purified BLM HRDC domain by gel-filtration chromatography, from which a monodisperse peak corresponding to a molecular mass of ~17 kDa was obtained (Fig. 2C). This result suggests that the fragment (calculated mass of 11.3 kDa) is present exclusively as a monomer or dimer. The protein was then subjected to ultracentrifugation sedimentation-equilibrium analysis. The resultant absorbance data (Fig. 2D) fitted well to the calculated curve from the ideal single species model, from which the molecular mass was calculated to be 12.9 kDa. Therefore, the BLM HRDC domain adopts a monomeric form in solution. The protein was further analysed by DLS, which showed that BLM HRDC is monodispersed in solution with a hydrodynamic radius of ~25 Å. This dimension is also close to the value determined for the monomeric structure of the WRN HRDC domain (*20*). The DLS measurements further demonstrated that the protein is thermostable up to 40°C (Fig. 2E), consistent with the CD spectroscopic results. In summary, the results obtained from various techniques showed that BLM HRDC is a stable monomer in solution, which prompted us to determine its structure by NMR spectroscopy.

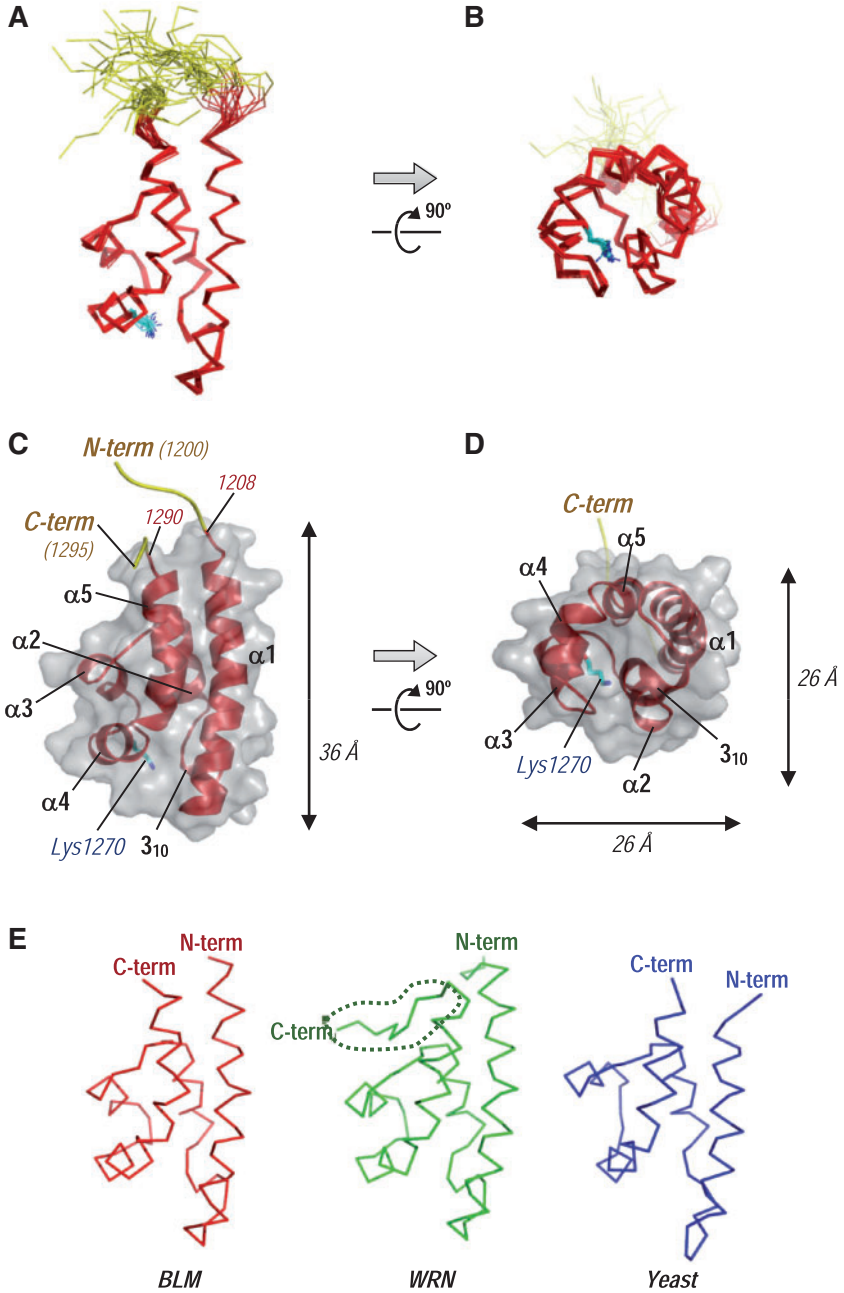

Fig. 3 Solution structure of the BLM HRDC domain. (A) Front view of the backbone superpositions of the final 20 simulated annealing structures of the BLM HRDC domain. The domain core (amino acids 1,208–1,290) is coloured in red, while the distorted N- and C-terminal residues (amino acids 1,200–1,207 and 1,291–1,295, respectively) are in yellow. Lys1270 is shown as a stick model in cyan. (B) Bottom view. (C) Front view of BLM HRDC as a ribbon model. Of all the NMR structures determined, the one with the lowest total energy was selected as being representative. The orientation is the same as in (A). Secondary-structure elements are labelled, and the dimensions of the molecule are indicated. The molecular surface of the domain (amino acids 1,208–1,290) is shown in transparent grey. The height of BLM HRDC is ~36 Å, while the width and depth are 26 Å. (D) Bottom view. (E) Structural comparison of BLM HRDC (red) with the HRDC of WRN (*20*) (green) and *Saccharomyces cerevisiae* Sgs1 (*23*) (blue). The molecules are viewed in the same orientation as in (A). The N and C termini of each molecule are labelled. The HRDC domains of BLM, WRN and Sgs1 comprise 83, 94 and 78 amino acids, respectively. BLM HRDC is superimposed on the WRN and Sgs1 HRDCs with an r.m.s.d. of 2.1 (83 Cα atoms) and 3.0 (78 Cα atoms). BLM and Sgs1 HRDCs lack the C-terminal structured loop that is present in WRN HRDC (encircled by dashed line).

Structure determination by NMR

Protein sample for NMR measurements was prepared as described in 'Experimental Procedures' section. The ^{15}N–^{1}H HSQC spectrum of ^{15}N-labelled BLM HRDC gave a highly dispersed pattern of cross-peaks, supporting the view that the whole protein molecule adopts an ordered monomeric structure in solution (Supplementary Fig. S2). NMR resonance assignments were obtained by performing double- and triple-resonance NMR experiments using ^{15}N- and ^{15}N, ^{13}C-labelled protein samples. Main-chain resonance assignments were obtained from standard multi-dimensional NMR methods. The structure of BLM HRDC was determined from 2,314 distance

and 173 torsion angle restraints (Table I). This relatively large number of angle restraints was due to the φ and χ1 angles obtained from TALOS, and the HNHB and HN(CO)HB experiments.

Figure 3A and B depicts the backbone of the final 20 structures derived from the NMR data. The atomic coordinates throughout the protein including the loop regions are well-defined, except for eight (amino acids 1,200–1,207) N-terminal and five (amino acids 1,291–1,295) C-terminal residues. Most of the NMR signals of the backbone residues in the N-terminal region were missing and the observed signals showed narrow line shapes. These results imply that the terminal regions are unstructured in solution. The root mean square deviations (r.m.s.d.) calculated from the averaged structure of the well-defined region (residues 1,210–1,290) was 0.31 and 0.77 Å for the backbone and all heavy atoms, respectively.

Overall structure

The structure revealed that amino acids 1,208–1,290 of BLM fold into a classical HRDC structure, comprising a globular bundle of five α-helices and one 3_{10}-helix connected by short loop regions (Fig. 3C and D). The BLM HRDC domain begins with a long N-terminal α-helix (α1) comprising 24 amino acids. The α1 helix is flanked by the last helix α5, which adopts an anti-parallel orientation relative to α1. As a result, the N and C termini of the HRDC domain are positioned in close spatial proximity, as is often observed for structurally independent domains. The majority of hydrophobic amino acids in the domain including Leu 1,222, Leu 1,246 and Ala 1,250 assemble to form a folding core. This result is consistent with a previous report where mutation of Ala 1,250 to proline in full-length BLM resulted in the cessation of protein expression (*6*). These hydrophobic residues are conserved in WRN HRDC and are used for domain folding (Fig. 1B).

A structural comparison of the BLM HRDC domain with the HRDC domain of WRN (*20*) and yeast homologue Sgs1 (*23*) is shown in Fig. 3E. WRN and Sgs1 HRDCs are also present as monomers in solution. WRN HRDC possesses a unique C-terminal extended loop that is tightly packed against the upper surface of the classical HRDC fold (encircled by dashed line), whereas the corresponding residues of both BLM and Sgs1 are disordered and not incorporated into the domain architecture. As a consequence, the HRDC domains of BLM and Sgs1 are more compact compared with that of WRN; the solvent accessible surface areas of BLM (5,336 $Å^2$) and Sgs1 (5,229 $Å^2$) HRDCs are smaller by ∼10% compared with that of WRN HRDC (5,835 $Å^2$). With respect to the domain architecture and size, BLM HRDC has greater similarity to yeast HRDC rather than WRN HRDC.

BLM HRDC is highly electronegative and defective in DNA binding

Figure 4A and B shows the electrostatic surface potential of the BLM HRDC domain. The front surface of the domain (Fig. 4A) is highly electronegative given the presence of an acidic residue cluster (one aspartic acid and six glutamic acid residues) that are spread on the domain surface. The back surface (Fig. 4B) also includes an electronegative region, although it is punctuated with electropositive and neutral regions. The acidic domain surface is another property of BLM HRDC that contrasts sharply with that of WRN HRDC (Supplementary Fig. S3A) and Sgs1 HRDC (Supplementary Fig. S3B). The isoelectric point (pI) of BLM HRDC (amino acids 1,208–1,290) was calculated to be 5.1, which is significantly lower than the values for WRN HRDC (8.1; amino acids 1,142–1,235) and Sgs1 HRDC (10.3; amino acids 1,272–1,349). Considering that the acidic residues of BLM are highly conserved among human, mouse and *Xenopus* enzymes (Supplementary Fig. S1B), the electronegative surface property of the BLM HRDC domain may be of functional importance in full-length BLM.

As its name implies, the HRDC domain has been found in several DNA helicases in addition to the RNase D family of nucleases (*24*). Consequently, interest in the HRDC domains has focused on their DNA binding ability. In fact, the isolated HRDC domain of Sgs1 was suggested to bind DNA by an NMR chemical-shift perturbation assay, although the interaction was considerably weak (dissociation constant around 30 μM) (*23*). On the other hand, we recently showed that purified WRN HRDC domain does not display detectable DNA binding ability *in vitro* (*20*).

Importantly, the proposed DNA-binding surface area of Sgs1 HRDC (*23*), which is electropositive (encircled in Supplementary Fig. S3B), is not conserved in BLM HRDC. The corresponding region of BLM HRDC (encircled in Fig. 4A) is highly electronegative and is unlikely to be involved in direct DNA interactions. Therefore, we then examined the DNA binding ability of purified BLM HRDC by fluorescence polarization assay. A forked duplex was chosen as the first binding substrate since it includes both single- and double-strands, and is known as one of the preferred substrates of full-length BLM (*1–3*). For the positive and negative controls, respectively, the WRN RQC (*5*) and WRN HRDC (*20*) domains were purified and used in the experiment. As shown in Fig. 4D, while the WRN RQC domain exhibited strong binding within the nanomolar range, the BLM HRDC domain displayed no detectable binding activity even at the highest protein concentration used (10 μM). We then examined the interaction using various DNA structures including a single-stranded DNA, a blunt-ended duplex, a 5′- or 3′-overhang duplex, a gapped duplex, a bubble and a Holliday junction [for oligo sequences (*5*)]. However, no binding (no increase in the polarization signal) was observed with any of these substrates (data not shown), as was the case with the WRN HRDC domain (*20*).

Possible functions of BLM HRDC

Full-length BLM acts in concert with topoisomerase IIIα to resolve recombination intermediates containing double Holliday junctions (DHJs), a process

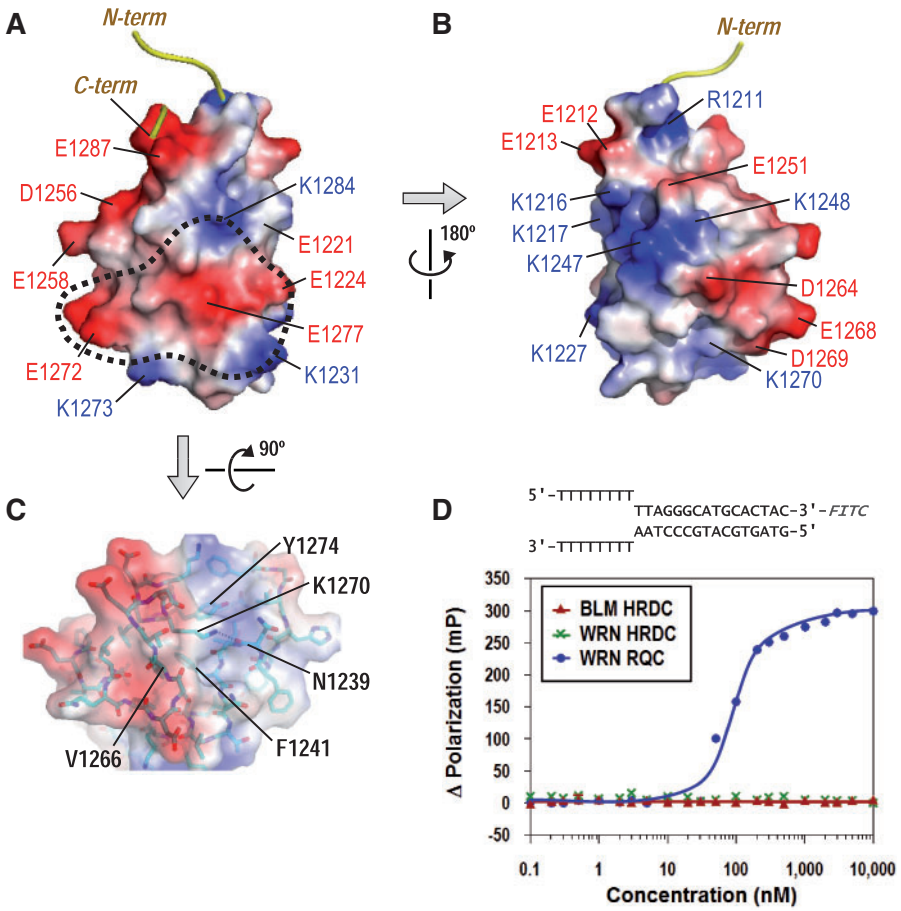

Fig. 4 Surface property of the BLM HRDC domain. (A) Electrostatic surface potential of BLM HRDC in the front view (similar orientation as in Fig. 3A). The surface is coloured for negative (red), positive (blue) and neutral (white) charges. Acidic (Asp or Glu) and basic (Arg or Lys) residues are labelled. The area corresponding to the proposed DNA interaction site of Sgs1 is encircled (*23*). The charge distribution of BLM HRDC differs from that of WRN and Sgs1 HRDCs (Supplementary Fig. S3A and B, respectively). (B) Back view (right). (C) The side-chain of Lys1270 is packed against the hydrophobic core of the domain and is partially involved in domain folding. The molecule is shown in a similar orientation as in Fig. 3D (bottom view). (D) DNA-binding assay employing fluorescence polarization. The forked duplex sequence (FITC-labelled) used in the assay is shown at the top. The interaction between the DNA and BLM HRDC was examined with increasing protein concentrations from 0 to 10 μM (red triangles). The WRN RQC domain (amino acids 949–1079), which possesses a similar domain size to BLM HRDC and exhibits strong DNA-binding activity (*5*), was used as a positive control (blue circles).

referred to as DHJ dissolution (*6*, *25*, *26*). It was reported that a single mutation of Lys1270 within the HRDC domain to valine (K1270V) partially disrupted the dissolution activity of full-length BLM, and this residue was predicted to play a role in mediating interactions with DNA (*6*). In our structure, the side chain of Lys1270, which is located in the middle of the short helix α4, adopts an extended conformation. The non-polar atoms of the side chain ($-CH_2-CH_2-CH_2-CH_2-$) are deeply buried in the hydrophobic patch formed by amino acids Phe1241, Val1266 and Tyr1274, while the amino group ($-NH_3^+$) at the tip is positioned to form a weak hydrogen bond with the main-chain oxygen atom of Asn1239 at a distance of ~3.3 Å (Fig. 4C). This lysine is conserved in WRN (Lys1208) and the side chain adopts the same extended conformation; the non-polar atoms of Lys1208 side-chain are buried in the hydrophobic patch of WRN HRDC, while the amino group forms the hydrogen bond with the main-chain oxygen atom of Ile1177 (corresponds to Asn1239 in BLM) at a distance of ~2.8 Å. Therefore, although the lysine is partially exposed to solvent, it seems to be important in stabilizing the HRDC framework in both BLM and WRN.

On the other hand, another negatively-charged HRDC domain was recently identified in a prokaryotic RecQ protein from the radioresistant bacterium *Deinococcus radiodurans* (*27*). This unusual RecQ protein possesses three tandem HRDC domains at the C-terminus, and a crystal structure of the third (C-terminal most) domain showed that the surface of this domain is also extensively electronegative with a pI of 4.6 (Supplementary Fig. S3C). The domain was proposed to be used for inter-domain interactions with the rest of the protein and assist in directing the protein to specific recombination/repair sites (*27*). In human cells, recruitment of full-length BLM to DNA double-strand break (DSB) sites also requires the HRDC domain (*28*). Repair of DSBs is crucial for tumour suppression, and DHJs are the central intermediates of this process (*29*). It was reported that expression of a mutant BLM which lacks the HRDC domain is unable to complement the

genomic-instability phenotype of Bloom syndrome cells (*30*). Currently, we speculate that BLM HRDC would bind with the rest of the protein and/or other protein(s) that recognize DSBs (especially protein that possesses a positively-charged molecular surface), thus facilitating concomitant recruitment of BLM to the DSB sites.

In conclusion, the present study reported the first structural insights into the HRDC domain of BLM. Although the precise function of this domain remains to be elucidated, its importance in human health is supported by the existing data (*6*, *28*, *30*). Structural differences between the HRDC domains of BLM and WRN, especially with respect to charge distribution, may be important for the recruitment of each enzyme to damaged DNA sites including DSBs and DHJs. In BLM, the HRDC domain is tethered to the RQC domain by a short linker region comprising ~10 a.a. (Fig. 1A). In contrast, WRN possesses a much longer linker comprising 76 a.a. (*20*). It is also tempting to speculate that, in BLM, the HRDC domain may contact with the outer surface of RQC within the same molecule (Supplementary Fig. S4), thus somehow contributing to the binding specificity of RQC to the HJ structure and/or relaxation of the four arms of the junction (also electronegative) by electrostatic repulsion.

Supplementary Data

Supplementary Data are available at *JB* Online.

Acknowledgements

We are grateful to A. Shimamoto and Y. Furuichi (GeneCare Research Institute, Japan) for providing BLM cDNA, N. Yoshihara for assistance with protein purification, J. Tsukamoto for N-terminal sequencing and TOF-MS and C. Kojima for assistance with NMR measurements.

Funding

Inamori Foundation, Uehara Memorial Foundation, Sumitomo Foundation, Japan Society for the Promotion of Science, Nara Institute of Science and Technology (NAIST) Foundation and NAIST Global COE Programme (to K.K.).

Conflict of interest

None declared.

References

1. Chu, W.K. and Hickson, I.D. (2009) RecQ helicases: multifunctional genome caretakers. *Nat. Rev. Cancer* **9**, 644–654
2. Rossi, M.L., Ghosh, A.K., and Bohr, V.A. (2010) Roles of Werner syndrome protein in protection of genome integrity. *DNA Repair* **9**, 331–344
3. Vindigni, A., Marino, F., and Gileadi, O. (2010) Probing the structural basis of RecQ helicase function. *Biophys. Chem.* **149**, 67–77
4. Goto, M. (2000) Werner's syndrome: from clinics to genetics. *Clin. Exp. Rheumatol.* **18**, 760–766
5. Kitano, K., Kim, S.Y., and Hakoshima, T. (2010) Structural basis for DNA strand separation by the unconventional winged-helix domain of RecQ helicase WRN. *Structure* **18**, 177–187
6. Wu, L., Chan, K.L., Ralf, C., Bernstein, D.A., Garcia, P.L., Bohr, V.A., Vindigni, A., Janscak, P., Keck, J.L., and Hickson, I.D. (2005) The HRDC domain of BLM is required for the dissolution of double Holliday junctions. *EMBO J.* **24**, 2679–2687
7. Lee, J.W., Kusumoto, R., Doherty, K.M., Lin, G.X., Zeng, W., Cheng, W.H., von Kobbe, C., Brosh, R.M. Jr, Hu, J.S., and Bohr, V.A. (2005) Modulation of Werner syndrome protein function by a single mutation in the conserved RecQ domain. *J. Biol. Chem.* **280**, 39627–39636
8. Hoadley, K.A. and Keck, J.L. (2010) Werner helicase wings DNA binding. *Structure* **18**, 149–151
9. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293
10. Goddard, T.D. and Kneller, D.G. (1999) *SPARKY3*. University of California, Sanfrancisco
11. Cavanagh, J., Fairbrother, W.J., Palmer III, A.G., and Skelton, N.J. (1996) *Protein NMR Spectroscopy*. Academic Press, San Diego
12. Sattler, M., Schleucher, J., and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**, 93–158
13. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **13**, 289–302
14. Archer, S.J., Ikura, M., Torchia, D.A., and Bax, A. (1991) An alternative 3D NMR technique for correlating backbone ^{15}N with side-chain Hβ resonances in larger proteins. *J. Magn. Reson.* **95**, 636–641
15. Grzesiek, S., Ikura, M., Clore, G.M., Gronenborn, A.M., and Bax, A. (1992) A 3D triple-resonance NMR technique for qualitative measurement of carbonyl-Hβ J coplings in isotopically enriched protein. *J. Magn. Reson.* **96**, 215–221
16. Herrmann, T., Guntert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* **319**, 209–227
17. Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., and Warren, G.L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921
18. Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486
19. Kleywegt, G.J. and Jones, T.A. (1997) Detecting folding motifs and similarities in protein structures. *Methods Enzymol.* **277**, 525–545
20. Kitano, K., Yoshihara, N., and Hakoshima, T. (2007) Crystal structure of the HRDC domain of human Werner syndrome protein, WRN. *J. Biol. Chem.* **282**, 2717–2728
21. Karow, J.K., Newman, R.H., Freemont, P.S., and Hickson, I.D. (1999) Oligomeric ring structure of the Bloom's syndrome helicase. *Curr. Biol.* **9**, 597–600
22. Beresten, S.F., Stan, R., van Brabant, A.J., Ye, T., Naureckiene, S., and Ellis, N.A. (1999) Purification of

overexpressed hexahistidine-tagged BLM N431 as oligomeric complexes. *Protein Expr. Purif.* **17**, 239–248
23. Liu, Z., Macias, M.J., Bottomley, M.J., Stier, G., Linge, J.P., Nilges, M., Bork, P., and Sattler, M. (1999) The three-dimensional structure of the HRDC domain and implications for the Werner and Bloom syndrome proteins. *Structure* **7**, 1557–1566
24. Morozov, V., Mushegian, A.R., Koonin, E.V., and Bork, P. (1997) A putative nucleic acid-binding domain in Bloom's and Werner's syndrome helicases. *Trends Biochem. Sci.* **22**, 417–418
25. Wu, L. and Hickson, I.D. (2003) The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* **426**, 870–874
26. Plank, J.L., Wu, J., and Hsieh, T.S. (2006) Topoisomerase IIIα and Bloom's helicase can resolve a mobile double Holliday junction substrate through convergent branch migration. *Proc. Natl Acad. Sci. USA* **103**, 11118–11123
27. Killoran, M.P. and Keck, J.L. (2008) Structure and function of the regulatory C-terminal HRDC domain from Deinococcus radiodurans RecQ. *Nucleic Acids Res.* **36**, 3139–3149
28. Karmakar, P., Seki, M., Kanamori, M., Hashiguchi, K., Ohtsuki, M., Murata, E., Inoue, E., Tada, S., Lan, L., Yasui, A., and Enomoto, T. (2006) BLM is an early responder to DNA double-strand breaks. *Biochem. Biophys. Res. Commun.* **348**, 62–69
29. Bzymek, M., Thayer, N.H., Oh, S.D., Kleckner, N., and Hunter, N. (2010) Double Holliday junctions are intermediates of DNA break repair. *Nature* **464**, 937–941
30. Yankiwski, V., Noonan, J.P., and Neff, N.F. (2001) The C-terminal domain of the Bloom syndrome DNA helicase is essential for genomic stability. *BMC Cell Biol.* **2**, 11