The three-dimensional structure of the HRDC domain and implications for the Werner and Bloom syndrome proteins

Z Liu, MJ Macias, MJ Bottomley, G Stier, JP Linge, M Nilges, P Bork and M Sattler*

Background: The HRDC (helicase and RNaseD C-terminal) domain is found at the C terminus of many RecQ helicases, including the human Werner and Bloom syndrome proteins. RecQ helicases have been shown to unwind DNA in an ATP-dependent manner. However, the specific functional roles of these proteins in DNA recombination and replication are not known. An HRDC domain exists in both of the human RecQ homologues that are implicated in human disease and may have an important role in their function.

Results: We have determined the three-dimensional structure of the HRDC domain in the *Saccharomyces cerevisiae* RecQ helicase Sgs1p by nuclear magnetic resonance (NMR) spectroscopy. The structure resembles auxiliary domains in bacterial DNA helicases and other proteins that interact with nucleic acids. We show that a positively charged region on the surface of the Sgs1p HRDC domain can interact with DNA. Structural similarities to bacterial DNA helicases suggest that the HRDC domain functions as an auxiliary domain in RecQ helicases. Homology models of the Werner and Bloom HRDC domains show different surface properties when compared with Sgs1p.

Conclusions: The HRDC domain represents a structural scaffold that resembles auxiliary domains in proteins that are involved in nucleic acid metabolism. In Sgs1p, the HRDC domain could modulate the helicase function via auxiliary contacts to DNA. However, in the Werner and Bloom syndrome helicases the HRDC domain may have a role in their functional differences by mediating diverse molecular interactions.

Address: European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany.

*Corresponding author.
E-mail: sattler@EMBL-Heidelberg.de

Key words: Bloom syndrome, HRDC domain, NMR, RecQ helicases, Werner syndrome

Received: 29 June 1999
Revisions requested: 10 August 1999
Revisions received: 22 September 1999
Accepted: 1 October 1999

Published: 30 November 1999

Structure December 1999, 7:1557–1566

0969-2126/99/$ – see front matter

Introduction

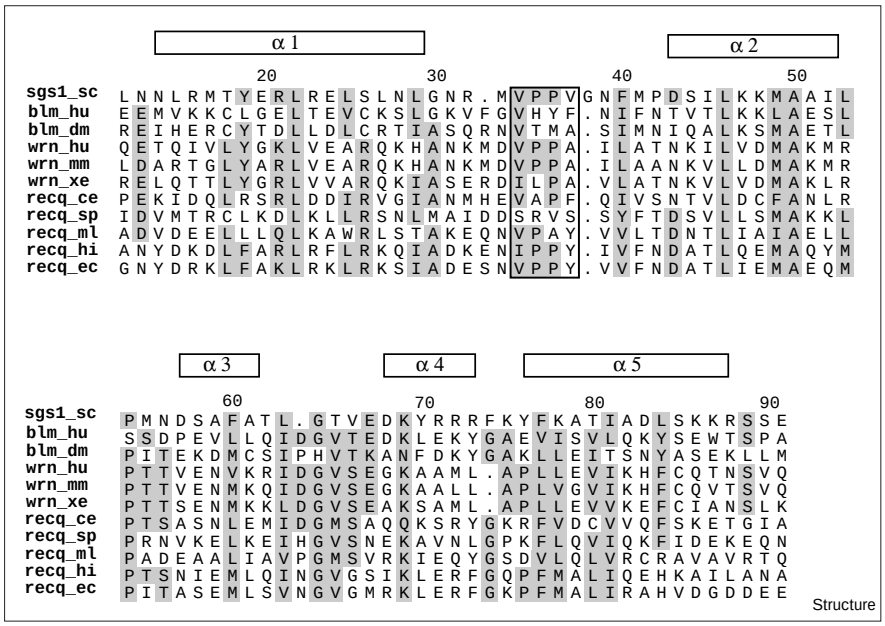
Werner syndrome (WS) and Bloom syndrome (BS) are inherited autosomal-recessive diseases characterized by chromosomal aberrations, impaired replication and genomic instability. Although a predisposition to cancer is common to both diseases, the specific phenotypes are different [1]. For example, only WS patients show symptoms of premature aging. *Sgs1*, the *Saccharomyces cerevisiae* homolog of the *WS* gene, is important for genomic stability in yeast [2], and deletion of *Sgs1* causes premature aging in yeast cells [3].

The proteins encoded by the *WS*, *BS* and *Sgs1* genes (WRN, BLM and Sgs1p, respectively) belong to the RecQ family of DNA helicases. They contain a DEAH (single-letter amino acid code) helicase domain [4–6] and a C-terminal domain (RecQ-Ct) that is unique to the RecQ family [7,8]. A number of RecQ helicases, including human and *Xenopus* WRN, BLM, Sgs1p and RecQL, have been shown to unwind double-stranded DNA with 3′→5′ directionality in an ATP-dependent manner [9–12,8,13,14].

The specific functional roles of the RecQ helicases and their involvement in the Werner and Bloom syndromes are still under investigation. Sgs1p can interact with type II and III topoisomerases, which function in DNA replication [15,16]. Sgs1p, *Escherichia coli* RecQ, and the human proteins WRN and BLM have been shown to suppress illegitimate recombination [17,18]. An essential role in the formation of replication foci was recently discovered for FFH-1, the *Xenopus laevis* ortholog of WRN [19]. Moreover, the wild-type WRN protein, but not a mutant with a deletion in the helicase domain, was found to co-purify with a multiprotein–DNA replication complex [20]. These findings imply that the RecQ proteins function in the regulation of DNA recombination or replication, and correlate with the observed genetic instability and increased mutation rates found in Werner and Bloom syndromes.

The HRDC (helicase and RNaseD C-terminal) domain has been recently described as an 80 amino acid protein domain found at the C terminus of RecQ helicases and RNaseD homologs from various organisms, including humans, yeast and bacteria [7]. An HRDC domain exists in at least three genes linked to human diseases. These genes encode the RecQ helicases WRN and BLM, and the human PM-Scl autoantigen, an RNaseD homolog

Figure 1



Sequence alignment of HRDC domains in RecQ helicases. The multiple alignment was performed with the program CLUSTAL X [65] and manually refined. Similar amino acids according to a slightly modified Blosum62 matrix are shadowed in gray. The helices found in the 3D structure of the Sgs1p HRDC domain are shown as boxes on top of the alignment. A stretch of conserved hydrophobic residues in the loop connecting helices α1 and α2 is boxed.

found in patients affected by polymyositis and scleroderma [21]. Multiple mutations affecting the HRDC domain have been mapped to the WRN and BLM genes [7]. For example, 6 of 19 mutations are found within the exons encoding the HRDC domain [22]. A sequence alignment of the HRDC domains found in RecQ helicases is shown in Figure 1.

To shed light on the possible functions of the HRDC domain in RecQ helicases we have determined the three-dimensional (3D) structure of the *S. cerevisiae* Sgs1p HRDC domain using heteronuclear multidimensional nuclear magnetic resonance (NMR) spectroscopy. Our study presents the first structure of an HRDC domain. Structural similarity searches revealed that the 3D structure of the HRDC domain resembles auxiliary domains in bacterial DNA helicases and other proteins that interact with DNA. In NMR titrations we found that a basic patch on the surface of the Sgs1p HRDC domain can interact with DNA. We then calculated homology models of the WRN and BLM HRDC domains. Distinct charge distributions on the surfaces of these models suggest that the HRDC domain could have a role in functional differences between the WRN and BLM proteins.

Results

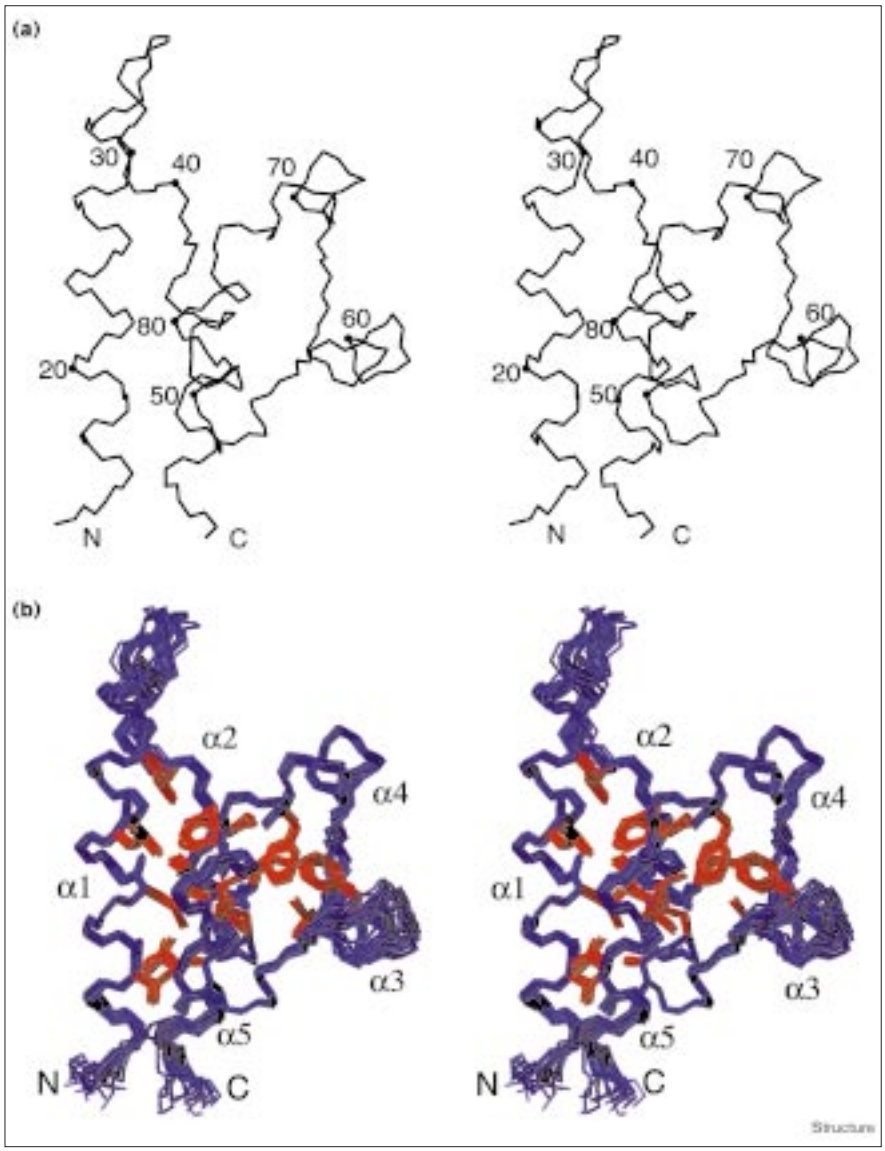
HRDC domain structure

The backbone trace of the lowest-energy structure of the Sgs1p HRDC domain is shown in Figure 2a. An ensemble of the 15 lowest-energy conformations is shown in Figure 2b. The experimental restraints and structural statistics are summarized in Table 1. The 3D structure of the HRDC domain consists of five helices, α1–α5, comprising residues 14–32, 43–51, 59–62, 68–74 and 77–88 (Figure 3a). Helices α1, α2 and α5 form a right-hand twisted three-helix bundle with interhelical angles of 130° (α1/α2), 160° (α1/α5) and 35° (α2/α5). The three-helix bundle is flanked by a short 3_{10} helix (α3) and helix α4. The two N-terminal helices α1 and α2 are connected by a long hydrophobic loop (residues Met33–Pro42), whereas helices α2/α3 and α3/α4 are connected by short turns. Helix α4 is linked to α5 by a kink at Lys75, which changes the orientation of the helical axis by ~60°. The N and C termini of the HRDC fold are in close spatial proximity, as is often observed for structurally independent domains.

Conserved residues in all five helices (Figure 1) are important for the packing of the HRDC fold. Whereas conserved hydrophobic residues in helices α1, α2 and α5 contribute to the hydrophobic core of the HRDC domain, helices α3 and α4 form additional contacts with α2 and α5 via Phe60, Leu63 and Phe74. Notably, Leu22, Leu46 and Met49 are key residues for hydrophobic interactions and are highly conserved in the HRDC family, suggesting a common 3D fold for all HRDC domains.

The conformation of the ten-residue loop connecting α1 and α2 is well defined by the NMR data. A stretch of four solvent-exposed hydrophobic residues, Val34–Pro35–Pro36–Val37, is found in this loop and conserved in many other HRDC domains (Figure 1). This hydrophobic region on the surface of the HRDC domain may therefore constitute a site for intramolecular or intermolecular interactions.

Figure 2



Stereoviews of the NMR structure of the Sgs1p HRDC domain. Residues 12–89 are shown (corresponding to residues 1272–1349 in Sgs1p). (a) Backbone trace (N, C^{α}, C′ atoms) of the lowest-energy structure. (b) Ensemble of the 15 lowest-energy NMR structures. The sidechains of conserved hydrophobic residues are shown in red.

The Sgs1p HRDC domain is highly positively charged because of numerous basic sidechains located at the surface of the protein (Figure 3b). Specifically, a region comprising residues of helix α4 (Lys69, Arg72, Arg73), the N terminus of α5 (Lys75, Lys78) and the C-terminal end of helix α1 (Arg32), exhibits a cluster of positive charges. Among those, Lys69 is highly conserved in the HRDC domains found in RecQ helicases (Figure 1).

The HRDC domain resembles auxiliary DNA-binding domains

To test whether the HRDC domain resembles a known structural motif, we searched the Protein Data Bank (PDB) using the Dali program [23]. Only PDB entries that were found with a Z score ≥ 3 were considered and further analyzed as described in the Materials and methods section. The Dali search revealed structural similarities between the HRDC domain and subdomains in DNA helicases, recombinases and polymerases.

The 3D structure of the HRDC domain resembles the domain 1B in bacterial DExx-box helicases (where x is any amino acid), such as *E. coli* Rep helicase [24] and the bacterial PcrA helicase [25,26]. The structural similarity to Rep helicase in complex with single-stranded DNA (ssDNA) is shown in Figure 4. The positional root mean square deviation (rmsd) for backbone atoms of 54 residues is 3.1 Å. Note that the primary sequences of the HRDC domain and domain 1B have no detectable sequence similarity. Both the PcrA and Rep proteins consist of four subdomains

Table 1

Structural statistics for the Sgs1p HRDC domain.

	<SA>	<SA>$^{water\text{-}refined}$
Rmsd (Å) from experimental distance restraints*		
Unambiguous (1807 total)	0.009 ± 0.003	0.019 ± 0.002
Intraresidue (959)	0.014 ± 0.007	0.020 ± 0.004
Sequential (|i – j|) = 1 (280)	0.002 ± 0.002	0.005 ± 0.003
Medium range 1 < (|i – j|) ≤ 4 (269)	0.007 ± 0.0008	0.019 ± 0.003
Long range (|i – j|) ≥ 5 (299)	0.008 ± 0.001	0.021 ± 0.002
Ambiguous (83)	0.012 ± 0.001	0.023 ± 0.004
Hydrogen bonds (34)	0.010 ± 0.002	0.015 ± 0.005
Rmsd (°) from experimental torsion restraints†		
Dihedral angles (44 ϕ, 7 χ_1)	0.23 ± 0.05	0.43 ± 0.17
Rmsd from experimental ψ restraints‡		
$\Gamma^{H\alpha C\alpha,C'}$ [Hz] (39)	1.8 ± 0.08	3.3 ± 0.16
$^3\Delta C^{\alpha}_{(ND)}$ [ppb] (39)	8.2 ± 0.03	8.1 ± 0.13
Coordinate precision (Å)§		
N, C^{α}, C′ (residue 13–88)	0.43 ± 0.07	0.61 ± 0.09
All heavy atoms (residue 13–88)	0.91 ± 0.10	1.07 ± 0.11
Structural quality¶		
$E_{L.-J.}$¥	–303 ± 8	–444 ± 15
No. of bad contacts (PROCHECK)	0 ± 0	0.07 ± 0.24
Quality index (WHATIF)	–1.03 ± 0.08	–0.76 ± 0.09
Ramachandran plot		
% in most-favored region	88.0 ± 1.5	91.5 ± 0.9
% in additionally allowed region	10.5 ± 1.1	8.4 ± 0.8

<SA> is an ensemble of the 15 lowest-energy solution structures of the Sgs1p HRDC domain; <SA>$^{water\text{-}refined}$ is an ensemble of 15 water-refined [57] structures. The CNS [50] E_{repel} function was used to simulate van der Waals interactions with an energy constant of 25.0 kcal mol^{-1} Å^{-4} using PROLSQ van der Waals radii as described in [57]; Rmsd values for bond lengths, bond angles and improper dihedral angles are 0.0013 ± 0.00004 Å, 0.291 ± 0.003° and 0.132 ± 0.003° for <SA>, and 0.004 ± 0.00009 Å, 0.510 ± 0.016° and 0.433 ± 0.019° for <SA>$^{water\text{-}refined}$, respectively. 1 kcal = 4.18 kJ. *Distance restraints were employed with a soft square-well potential [51] using an energy constant of 50 kcal mol^{-1} Å^{-2}. For hydrogen bonds, 34 distance restraints with bounds of 1.8–2.3 Å (H–O), and 2.8–3.3 Å (N–O) were derived for 17 slow-exchanging amide protons. No distance restraint was violated by more than 0.5 Å in any of the final structures. †Dihedral-angle restraints were applied to 44 ϕ (–60° ± 40°) and 7 χ_1 (180° ± 40°) dihedral angles using energy constants of 200 kcal mol^{-1} rad^{-2}. No dihedral angle restraint was violated by more than 5°. ‡The structures were directly refined against the experimentally determined cross-correlated relaxation rates $\Gamma^{H\alpha C\alpha,C'}$ and three-bond H/D isotope effects on the C^{α} chemical shifts $^3\Delta C^{\alpha}_{(ND)}$. Final energy constants of 0.05 kcal mol^{-1} Hz^{-2} and 0.02 kcal mol^{-1} ppb^{-2} were used. §Coordinate precision is given as the Cartesian coordinate rmsd of the 15 lowest-energy structures with respect to their mean structure. ¶Structural quality checks were performed for residues 13–88 using the programs PROCHECK [58] and WHATIF [59]. ¥E_{L-J} in kcal mol^{-1} is the Lennard–Jones van der Waals energy calculated using the CHARMM PARMALLH6 parameters. E_{L-J} was not included in the target function during the structure calculations.

(1A, 2A, 1B, 2B) and adopt a similar 3D fold. In the Rep helicase–DNA complex a positively charged groove between domains 1A and 1B interacts with the single-stranded DNA [24]. In the substrate and product complexes of PcrA with a DNA substrate, the domains 1B and 2B alter their positions relative to each other and define two protein conformations with low and high affinity for duplex DNA [26]. Thus, domain 1B functions as an auxiliary domain in bacterial DNA helicases and contributes additional contacts to the DNA substrate. The structural similarity to domain 1B suggests that the HRDC domain may have a similar role in RecQ helicases.

The HRDC domain also resembles the N-terminal 8 kDa domain of human DNA polymerase β (Pol β) [27]. The proteolytically cleavable N-terminal domain of Pol β is not found in other DNA polymerases. It contributes about 70% of the ssDNA-binding affinity of the Pol β holoenzyme and the domain alone binds to ssDNA with a dissociation constant in the low micromolar range [28]. In the crystal structure of Pol β complexed to a gapped DNA substrate [27], the 8 kDa domain makes additional contacts to the DNA.

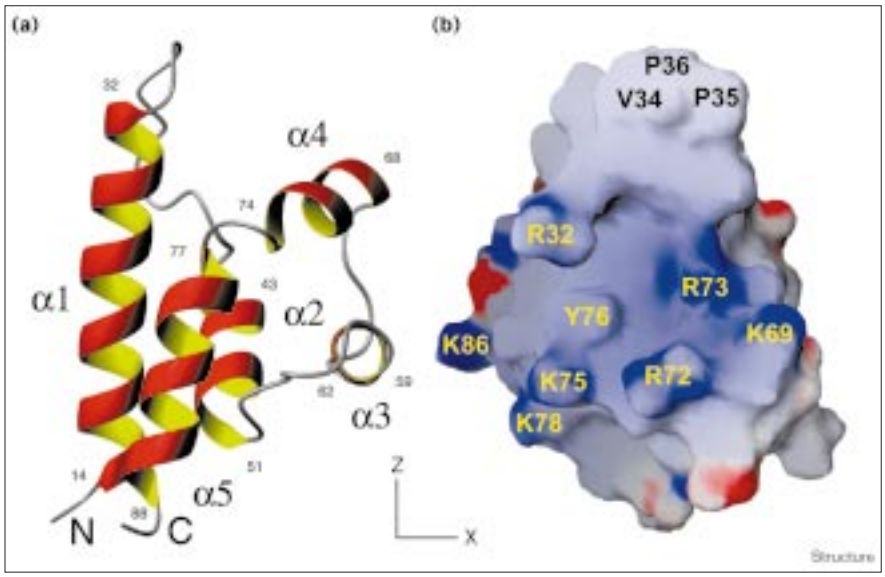
The Dali searches also revealed structural similarities to the N-terminal domain of the site-specific recombinase XerD [29] and to a helical domain extending the finger domain in T7 RNA polymerase [30]. For each protein it has been suggested that the subdomains resembling HRDC contribute to DNA binding in their respective enzyme–substrate complexes. These structural similarities show that the HRDC fold resembles auxiliary DNA-binding domains in proteins that are involved in nucleic acid metabolism and points to a similar role for the HRDC domain.

The Sgs1p HRDC domain can interact with DNA

To test whether the Sgs1p HRDC domain can bind to DNA we measured chemical-shift changes in the NMR spectra of ^{15}N-labeled Sgs1p HRDC domain upon addition of DNA. Significant spectral changes were observed with different DNA ligands, indicating that the Sgs1p HRDC domain interacts with DNA. The backbone amides of Gly30, Asn31, Arg32, Met33, Lys69, Arg73, Tyr76 and Ala79 display the largest chemical-shift changes upon addition of a partially double-stranded DNA (dsDNA) (8/16-mer) (Figure 5a). These residues are located at a basic patch on the surface of the protein around helix α4 and the C-terminal end of helix α1 (Figure 5b). We found significant and comparable chemical-shift changes for partially dsDNA with a single-stranded overhang (7/13-mer, 8/16-mer) and for ssDNA ligands (8-mer and 20-mer). Smaller spectral changes were observed with dsDNA. In a titration experiment with the ssDNA 20-mer we measured chemical-shift changes as a function of DNA concentration and determined a dissociation constant of ~30 μM (data not shown).

To test for DNA binding we also attempted gel-shift retardation experiments using both the yeast Sgs1p and the human WRN HRDC domains. However, because of precipitation of the sample we were not able to observe DNA band shifts. Nevertheless, upon adding increasing amounts of either of the proteins the signal of the

Figure 3



Ribbon and electrostatic surface representation of the Sgs1p HRDC domain. (a) 3D structure of the Sgs1p HRDC domain in the same orientation as in Figure 2. The start and end residues of the helices are labeled by residue numbers. (b) Surface depiction of the Sgs1p HRDC domain, viewed along the hydrophobic loop connecting helices α1 and α2 and helix α4 (rotated counter-clockwise by 90° around the z axis and clockwise by 90° around the x axis compared with (a)). Positive and negative charges are colored in blue and red, respectively.

unshifted DNA probe was reduced proportionally. This suggests that the precipitation is due to an interaction between the protein and DNA.

These data indicate that the Sgs1p HRDC domain can interact with DNA. The binding is not sequence specific, as we observed comparable chemical-shift changes in the NMR spectra when using different DNA ligands. Together with the relatively low binding affinity, this suggests that the interaction of the Sgs1p HRDC domain with DNA is largely mediated by electrostatic contacts between the basic patch on its surface and the phosphate backbone of the DNA.

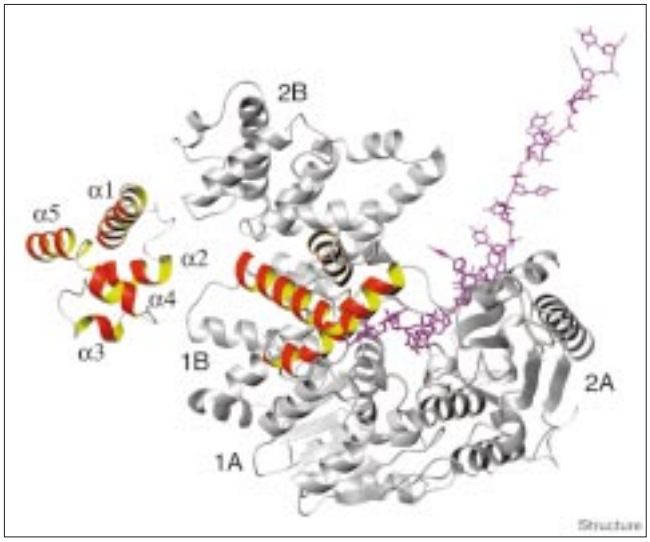
Homology models of the Werner and Bloom syndrome HRDC domains

In order to compare the charge distribution at the surfaces of the WRN and BLM HRDC domains to Sgs1p we calculated homology models using the program Modeler [31]. The hydrophobicity of the loop between α1 and α2 is conserved in most HRDC domains (Figure 1). Consequently, a hydrophobic patch is also found in this region of the WRN and BLM HRDC domains (Figure 6). It is therefore likely that this region is involved in intramolecular or intermolecular interactions in the RecQ proteins. However, of the amino acids that contribute to the basic patch at the surface of the Sgs1p HRDC domain only Lys69 is conserved. Therefore, in the human RecQ HRDC domains the corresponding region of the surface is much less positively charged. Moreover, in the BLM HRDC domain a negatively charged patch is found in a distant region involving helix α3. This suggests that auxiliary molecular interactions involving the HRDC domain are likely to be different in Sgs1p, WRN and BLM.

Discussion

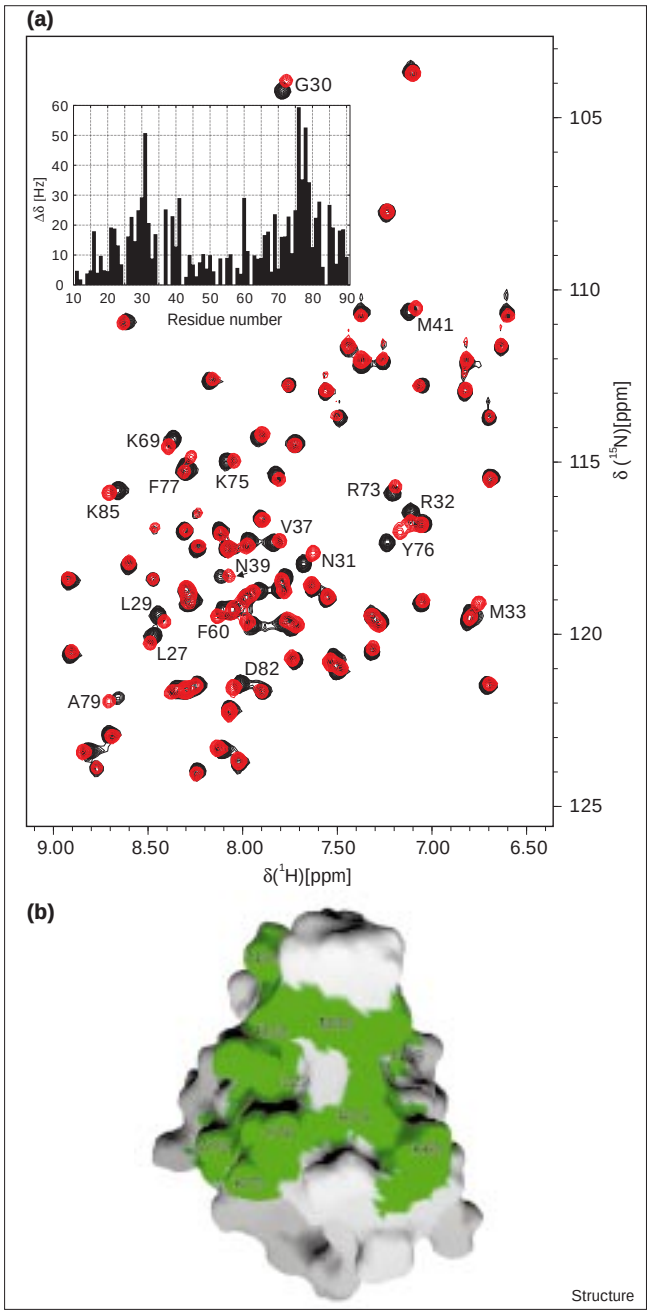
The structural similarity to domain 1B indicates that the HRDC fold is structurally conserved between the DExx-box and RecQ helicases even though they belong

Figure 4



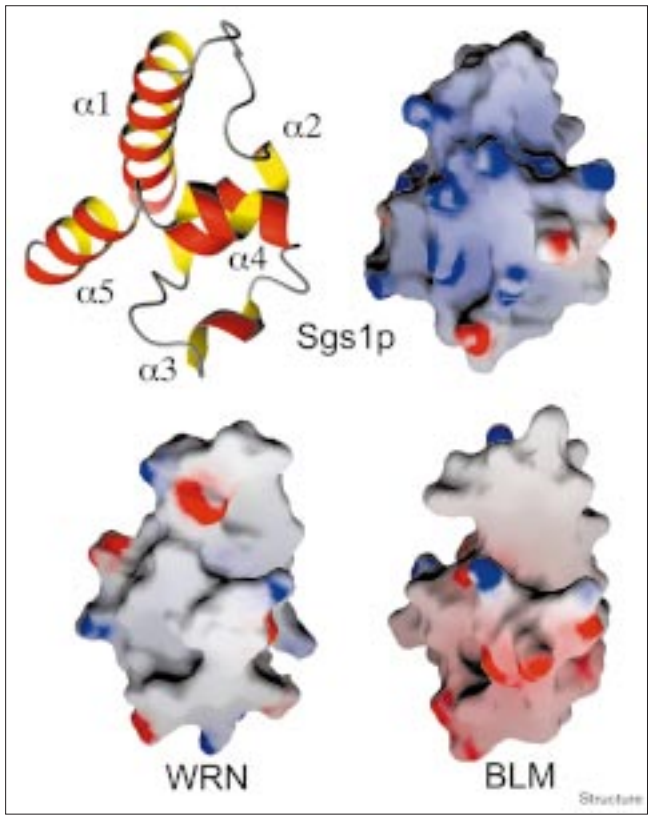
Structural similarity between the Sgs1p HRDC domain and the *E. coli* Rep helicase in complex with ssDNA. The structure of the Rep helicase corresponds to the closed conformation found in the crystal structure. The four subdomains (1A, 2A, 1B, 2B) that are also found in the PcrA DNA helicase are labeled. The HRDC domain and the structurally similar domain 1B in Rep are colored in red and yellow. The HRDC domain is shown in the same orientation as when superimposed on domain 1B. For the superposition, residues 12–31, 42–49, 55–58, 67–73 and 74–88 of the Sgs1p HRDC domain were aligned with residues 107–126, 131–138, 149–152, 157–163 and 165–179 of domain 1B in the Rep helicase.

Figure 5



DNA binding of the Sgs1p HRDC domain. **(a)** Superposition of 2D $^{1}H,^{15}N$-HSQC spectra of 1 mM ^{15}N-labeled Sgs1p HRDC domain, with (red) and without (black) 1:1 partially dsDNA (8/16-mer) recorded on a 600 MHz spectrometer. Residues displaying large chemical-shift changes are labeled. The free and bound forms of the protein are in fast exchange on the NMR timescale. The insert shows the observed chemical-shift change $\Delta\delta = |\Delta\delta(^{15}N)| + |\Delta\delta(^{1}H)|$ in Hz versus residue number. **(b)** DNA-binding surface of the Sgs1p HRDC domain with the partially dsDNA 8/16-mer. Residues for which a chemical-shift difference $|\Delta\delta(^{15}N)| + |\Delta\delta(^{1}H)| > 20$ Hz is observed are colored in green. The DNA-binding site involves helix α4 and the C-terminal end of helix α1. Some residues in the hydrophobic loop between α1 and α2 are also affected. The same view as in Figure 3b is shown.

Figure 6

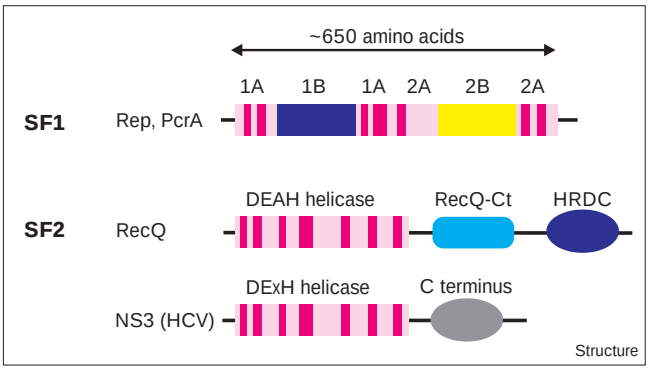


Electrostatic surface representation of homology-modeled HRDC domains of the Bloom and Werner syndrome proteins compared with Sgs1p. A ribbon representation of the Sgs1p HRDC domain is shown in the same orientation as the surfaces.

to different superfamilies: SF1 and SF2, respectively [32]. In DExx-box helicases, the 1A and 2A domains contain the seven helicase signature motifs and comprise the minimal structural unit required for helicase activity [33]. The auxiliary domains 1B and 2B are inserted into the primary sequences of the 1A and 2A domains, respectively. In contrast, in the primary sequence of RecQ proteins the HRDC domain is located C-terminal to a contiguous stretch of ~350 residues comprising the helicase domain (Figure 7). Nevertheless, in the 3D structure the relative orientation of the auxiliary domains may be similar [34]. A similar structural arrangement of minimal helicase unit and auxiliary domains may therefore be conserved in SF1 and SF2 helicases.

The structural similarities also suggest that the HRDC domain is associated with the RecQ helicase and may function as an auxiliary domain. Thus, similar to domain 1B in bacterial DNA helicases, the HRDC domain may contribute to positioning a helicase substrate by steric interactions or via secondary contacts to the DNA. This notion is supported by our finding that a highly positively

Figure 7



Domain composition of helicases belonging to superfamilies SF1 and SF2. The bacterial DExx-box helicases PcrA and Rep (SF1) are compared to RecQ helicases and to the hepatitis C virus (HCV) RNA helicase NS3, both of which belong to SF2 [32,34]. The minimal helicase unit is shown in light pink, the seven helicase signature motifs are colored in dark pink. The structurally related auxiliary domains, HRDC in RecQ and domain 1B in DExx-box helicases, are shown in dark blue. RecQ-Ct (cyan) is a conserved domain in RecQ helicases. The C-terminal domain in the HCV NS3 helicase is shown in gray.

charged patch at the surface of the Sgs1p HRDC domain can interact with DNA in a non-sequence-specific manner. Therefore, in the Sgs1p protein the HRDC domain may contribute to the RecQ helicase function by mediating auxiliary electrostatic contacts to the phosphate backbone of a DNA substrate.

The HRDC domain is not a specific DNA-binding domain, as is indicated by the low DNA-binding affinity and the non-sequence-specific interaction with different DNA ligands. Furthermore, most residues in the basic patch of the Sgs1p HRDC domain are not conserved in other HRDC domains. However, the HRDC domain is found in helicases and nucleases and resembles subdomains in helicases, polymerases and recombinases. This indicates that the HRDC fold represents a common structural motif in the context of nucleic acid metabolism. All these subdomains have been implicated in auxiliary substrate recognition by direct or indirect contacts to the respective nucleic acid substrates, suggesting a similar function for the HRDC domain.

The distinct surface properties of the WRN and BLM HRDC domains could have a role in mediating diverse molecular interactions to modulate the helicase function. For example, the highly negatively charged patch at the surface of the BLM HRDC domain might provide electrostatic interactions with other protein components in a multiprotein–DNA complex that specifically involves the BLM protein. This is consistent with the hypothesis that the HRDC domain may have a role in the functional specificities of the Werner and Bloom syndrome proteins.

All RecQ proteins exhibit a conserved helicase domain [35], and so it has been suggested that the human RecQ homologs may partially replace each other, even though the individual homologs are involved in distinct cellular processes [36]. It is conceivable that functional specificities are conferred by auxiliary domains, like HRDC, which have less well conserved primary sequences. In a similar way, the 1B domains of bacterial DNA helicases, which structurally resemble the HRDC domain, have been suggested to have a role in substrate specificity [26], and the primary sequences of the 1B domains are also more divergent than those of the helicase domains.

The HRDC domain may function as an auxiliary domain by mediating specific intermolecular contacts in multiprotein–DNA complexes that are implicated in the regulation of DNA replication and recombination [20,22,37]. Distinct auxiliary interactions of the HRDC domain may thus correlate with the functional differences of the Werner and Bloom syndrome proteins. For example, mutations affecting the HRDC domain in Werner syndrome may alter its surface properties, or lead to the complete loss of the WRN helicase as a result of intracellular degradation of truncated protein products. This may abolish molecular interactions involving the WRN HRDC domain and hence the formation of WRN-specific replication or recombination complexes.

The results of our structural studies suggest an important role for the HRDC domain in the specific functions of the Werner and Bloom syndrome proteins. It will be interesting to compare the biological activity of recombinant RecQ helicases with and without wild-type HRDC domains, and to study recombinant chimeric RecQ helicases where the HRDC domain has been replaced by this domain from other RecQ homologs.

Biological implications

The HRDC domain is a conserved protein domain found at the C terminus of many RecQ helicases, including the human Werner and Bloom syndrome proteins. RecQ helicases have been shown to unwind DNA in an ATP-dependent manner. However, the specific roles of these proteins in the regulation of recombination and replication are not yet understood. It is also not known how the functional differences of the human RecQ homologs are conferred.

Here, we present the first 3D structure of an HRDC domain. The structure resembles auxiliary domains in bacterial DNA helicases and other proteins that are involved in nucleic acid metabolism. For Sgs1p we found that the HRDC domain can interact with DNA in a non-sequence-specific manner. The structural similarity to domain 1B of bacterial DNA helicases suggests that the HRDC domain functions as an auxiliary domain in

RecQ helicases. Thus, as in the bacterial DNA helicases, the helicase activity and the principal DNA-binding affinity reside in the RecQ helicase domain whereas the HRDC domain may contribute auxiliary molecular interactions.

Distinct surface properties for homology models of the WRN and BLM HRDC domains suggest that they are involved in diverse molecular interactions. This could have a role in the formation of specific molecular contacts in multiprotein–DNA replication or recombination complexes. The HRDC domain may thereby contribute to the functional differences between the Werner and Bloom syndrome proteins.

Materials and methods

Sample preparation

The *S. cerevisiae* Sgs1 HRDC domain DNA was amplified by polymerase chain reaction (PCR) from yeast genomic DNA as template using the following two primers: (5′-GAGTCCATGGAACTTAATAATCTGCGAATGACATACG) and (5′-GAGTGAATTCACTC-GCTTGATCTCTTTTTGCTAAG). The amplified fragment was inserted into the *Nco*I/*Eco*RI sites of a modified pET9D expression vector (Novagen). The sequence was confirmed by DNA sequencing. The 91-residue recombinant protein used for the structural studies comprises residues 1271–1351 of Sgs1p and ten additional residues from an N-terminal histidine tag. The molecular weight of the protein was confirmed by mass spectroscopy. Uniformly ^{15}N-labeled and ^{15}N,^{13}C-labeled proteins were prepared by growing the *E. coli* strain BL21(DE3) overexpressing the Sgs1p HRDC domain in a minimal medium containing $^{15}NH_4Cl$ with or without uniformly ^{13}C-labeled glucose. The recombinant protein was purified by affinity chromatography on a Ni-chelating sepharose fast-flow column (Pharmacia). Further purification was achieved by ion-exchange chromatography on a Q-sepharose column and gel filtration on a Superdex 75 column. NMR samples were exchanged to 20 mM sodium phosphate buffer (pH 6.5) in H_2O or ^2H_2O at 1.5–1.8 mM concentration.

For the NMR titrations with DNA a non-histidine-tagged version of the Sgs1p HRDC domain (residues 1271–1351) was overexpressed as a His_6–GST–TEV–HRDC (GST is glutathione-*S* transferase; TEV is tobacco etch virus) fusion protein with a TEV protease [38] cleavage site. The fusion protein was cleaved by TEV protease and purified using Ni-chelating, glutathione- and Q-sepharose columns.

The HRDC domain of human WRN (residues 1152–1233) was expressed as a GST fusion protein. The GST tag was cleaved by TEV protease before separation by gel filtration. The purified protein was folded on the basis of results from 1D and 2D NMR spectra.

NMR spectroscopy

NMR spectra were acquired at 22°C on a Bruker DRX500 or DRX600 NMR spectrometer equipped with triple-resonance probes and pulsed-field gradients. Spectra were processed with NMRPipe [39] and analyzed using XEASY [40]. The ^{1}H, ^{13}C and ^{15}N assignments of the backbone resonances were obtained from sensitivity-enhanced 3D HNCA, CBCA(CO)NH and CBCANH experiments using a uniformly ^{15}N,^{13}C-labeled protein sample [41,42]. The sidechain signals were assigned from 3D HCCONH-TOCSY, CCONH-TOCSY and HCCH-TOCSY experiments [41,42]. Assignments for aromatic H^δ and H^ε protons were obtained from 2D experiments correlating C^β chemical shifts of aromatic residues to the corresponding H^δ and H^ε resonances [43]. Proline residues, including Pro35–Pro36, were sequentially assigned using a proline-edited CDCA(NCO)CAHA experiment [44]. Distance restraints were derived from ^{15}N- or ^{13}C-resolved 3D nuclear Overhauser effect (NOE) spectra [42,45]. $^1J(H^N,H^\alpha)$ coupling constants were measured in an HNHA-J experiment [46], and used to derive restraints for the backbone angle ϕ. $^3J(N,C^\gamma)$ were measured in a 2D spin-echo-difference experiment [47] to determine the dihedral angle χ_1. Cross-correlated relaxation rates ($\Gamma^{H\alpha C\alpha,C'}$) and the three-bond H/D isotope effect on the C^α chemical shifts ($^3\Delta C^\alpha_{(ND)}$) were measured in order to derive restraints for the backbone angle ψ. Slow-exchanging amide protons were identified from ^{1}H,^{15}N heteronuclear single-quantum coherence (HSQC) experiments [42,48] recorded after the addition of ^2H_2O to a lyophilized protein sample. The methyl groups of valine and leucine residues were stereospecifically assigned from a high-resolution ^{1}H,^{13}C -HSQC experiment recorded on a fractionally (10%) ^{13}C-labeled sample [49].

Structure calculations

The experimentally determined distance and dihedral-angle restraints (Table 1) were applied in a simulated annealing protocol using the program CNS [50]. Ambiguous distance restraints were employed using ARIA [51]. The simulated annealing protocol consisted of four stages: a high-temperature torsion angle simulated annealing phase (1100 steps at 10,000K with a time step of 45 fs) [52,53], a first torsion angle dynamics cooling stage from 10,000K to 2000K (550 steps), a Cartesian dynamics cooling stage from 2000K to 1000K (5000 steps) and a Cartesian dynamics cooling stage from 1000K to 0K. A time step of 5 fs was used for the Cartesian dynamics. During the high-temperature stage only two atoms per residue were included in the non-bonded interactions [51], during the first and second cooling stages hydrogens were not explicitly taken into account, and during the third cooling stage all atoms were considered. Van der Waals interactions were scaled essentially as described for the different cooling stages. The final structure calculations included distance restraints based on 1643 manually assigned, unambiguous NOEs and 164/83 unambiguous/ambiguous NOEs which were assigned by ARIA. When stereospecific assignments were not available, a floating chirality approach was used during the structure calculations [54]. Dihedral angle restraints for ϕ and χ_1 were derived from the experimentally measured $^3J(H^N,H^\alpha)$ and $^3J(N,C^\gamma)$, respectively. The structures were directly refined against experimentally determined cross-correlated relaxation rates ($\Gamma^{H\alpha C\alpha,C'}$) [55] and the three-bond H/D isotope effect on the C^α chemical shifts ($^3\Delta C^\alpha_{(ND)}$) [56] which define the backbone angle ψ (R Sprangers, MJB, JPL, J Schultz, MN and MS, unpublished results). In brief, energy constants were 0.02 kcal mol^{-1} Hz^{-2} and 0.01 kcal mol^{-1} ppb^{-2} during the first and second cooling, and 0.05 kcal mol^{-1} Hz^{-2} and 0.02 kcal mol^{-1} ppb^{-2} during the second cooling for $\Gamma^{H\alpha C\alpha,C'}$ and $^3\Delta C^\alpha_{(ND)}$, respectively. Eight iterations of ARIA were run with parameters as in previous calculations. The ψ restraints were switched on in iteration 6. Water-refined structures were calculated as described previously [57]. Structural quality was evaluated using the programs PROCHECK [58] and WHATIF [59].

Figures showing 3D structures were prepared with MOLMOL [60]. The electrostatic surface potential was calculated in GRASP [61], which was also used for the surface representations.

DNA binding: NMR titration experiments

To test for DNA binding, NMR samples were prepared by combining equimolar (~100 μM) solutions of DNA and the non-histidine-tagged ^{15}N-labeled Sgs1p HRDC domain and concentrating to 1 mM. Chemical shifts were recorded with ^{1}H,^{15}N-HSQC experiments using WATERGATE [62] combined with water-flip-back [63] for solvent suppression. The chemical shifts of the histidine-tagged and the non-histidine-tagged proteins were very similar. The assignments were confirmed by comparing 3D ^{15}N-edited NOESY experiments on both recombinant proteins.

The following oligonucleotides were used for the DNA-binding experiments. dsDNA: ds(GCGCGCGCGC); ds(CGCTTCGGGCCTTTC-CGGAAGC). ssDNA: ss(TTTTTTTT) (8-mer); ss(AGCCATGAG-CTAAGCTTGGT) (20-mer). Partially dsDNA: 5′-CAGCGTG-3′ and

5′-CACGCTGTACTAG-3′ (7/13-mer); 5′-ACAGCGTG-3′ and 5′-CACGCTGTACTATTGA-3′ (8/16-mer). For the single-stranded DNA 20-mer a titration curve was obtained by measuring ^{1}H and ^{15}N chemical shifts of the protein at DNA concentrations of L_0 = 0.1, 0.2, 0.4, 0.8 and 1.5 mM. Some precipitation was observed after adding the first aliquot of DNA. However, no further precipitation occurred when adding more DNA. The protein concentration P_0 was ~0.25 mM. Dissociation constants were obtained from a least-squares fit to the equation $\Delta\delta = \chi \times \Delta\delta_b$, where $\chi = PL/P_0$ is the mole fraction of bound protein, $\Delta\delta = (\delta - \delta_f)$ is the difference between the measured chemical shift (δ) and the chemical shift of free protein (δ_f), and $\Delta\delta_b = (\delta_b - \delta_f)$ is the chemical-shift difference between fully bound (δ_b) and free protein (δ_f). PL is the concentration of bound protein and is related to K_D through the definition for the dissociation constant.

DNA binding: gel-shift retardation

We attempted gel-shift retardation experiments using the recombinant HRDC domains of yeast Sgs1p and the human WRN helicase. The experiments were performed by incubating 10 ng of 5′-^{32}P-labeled DNA oligonucleotides with 0.01 to 1 μg of HRDC protein at room temperature for 30 min [64]. Various binding conditions were tested; however, in all experiments both the Sgs1p and WRN HRDC proteins precipitated after adding the DNA so that band shifts could not be observed.

Searches in the Protein Data Bank

For the searches in the PDB, coordinates of the lowest-energy water-refined structure (residues 12–90) of the Sgs1p HRDC domain were submitted to the Dali server at http://www2.ebi.ac.uk/dali [23]. Structural similarities with a Dali Z-score ≥ 3.0 were further analyzed. Structures were only considered if the regions of structural similarity approximately correspond to the five helices of the HRDC domain structure without any large insertions. The structural similarities were visually inspected by superimposing the coordinates to the structure of the HRDC domain. The rmsds for the superposition of backbone atoms (50–60 residues) with respect to the Sgs1p HRDC domain were in the order of 3 Å. The PDB accession numbers for the bacterial DNA helicases Rep and PcrA, DNA polymerase β, the XerD recombinase and T7 RNA polymerase are 1uaa, 1pjr, 1bpy, 1a0p, and 1aro, respectively.

Homology modeling

The primary sequences of the Sgs1p HRDC domain and of the HRDC domain to be modeled were aligned based on the sequence alignment of Figure 1 and the 3D structure of the Sgs1p HRDC domain. Homology models for the WRN, BLM and other HRDC domains were calculated with the program Modeler [31] using standard protocols. The lowest-energy conformer out of ten calculated models was used to calculate the electrostatic potential surface.

Accession numbers

The atomic coordinates and restraint lists have been deposited in the PDB with accession code 1D8B. The chemical shifts and NOE peak lists have been deposited to the BMRB with accession code 4445.

Acknowledgements

We thank Ramon Eritja for DNA oligonucleotides, Valérie Botquin and Hans Schöler for helpful suggestions and Matti Saraste for useful discussions. ZL acknowledges support by EMBL. MJB is grateful for an EMBO long-term postdoctoral fellowship. JPL thanks the Boehringer Ingelheim Fonds for a Ph. D. fellowship. This work has been supported by the DFG.

References

1. Ellis, N.A. (1997). DNA helicases in inherited human disorders. *Curr. Opin. Genet. Dev.* **7**, 354-363.
2. Watt, P.M., Hickson, I.D., Borts, R.H. & Louis, E.J. (1996). SGS1, a homologue of the Bloom's and Werner's syndrome genes, is required for maintenance of genome stability in Saccharomyces cerevisiae. *Genetics* **144**, 935-945.
3. Sinclair, D.A., Mills, K. & Guarente, L. (1997). Accelerated aging and nucleolar fragmentation in yeast sgs1 mutants. *Science* **277**, 1313-1316.
4. Ellis, N.A., *et al.,* & German, J. (1995). The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell* **83**, 655-666.
5. Lu, J., Mullen, J.R., Brill, S.J., Kleff, S., Romeo, A.M. & Sternglanz, R. (1996). Human homologues of yeast helicase. *Nature* **383**, 678-679.
6. Yu, C.E., *et al.,* & Schellenberg, G.D. (1996). Positional cloning of the Werner's syndrome gene. *Science* **272**, 258-262.
7. Morozov, V., Mushegian, A.R., Koonin, E.V. & Bork, P. (1997). A putative nucleic acid-binding domain in Bloom's and Werner's syndrome helicases. *Trends Biochem. Sci.* **22**, 417-418.
8. Bahr, A., De Graeve, F., Kedinger, C. & Chatton, B. (1998). Point mutations causing Bloom's syndrome abolish ATPase and DNA helicase activities of the BLM protein. *Oncogene* **17**, 2565-2571.
9. Umezu, K., Nakayama, K. & Nakayama, H. (1990). *Escherichia coli* RecQ protein is a DNA helicase. *Proc. Natl Acad. Sci. USA* **87**, 5363-5367.
10. Umezu, K. & Nakayama, H. (1993). RecQ DNA helicase of *Escherichia coli.* Characterization of the helix-unwinding activity with emphasis on the effect of single-stranded DNA-binding protein. *J. Mol. Biol.* **230**, 1145-1150.
11. Gray, M.D., *et al.*, & Loeb, L.A. (1997). The Werner syndrome protein is a DNA helicase. *Nat. Genet.* **17**, 100-103.
12. 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.
13. Bennett, R.J., Sharp, J.A. & Wang, J.C. (1998). Purification and characterization of the Sgs1 DNA helicase activity of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **273**, 9644-9650.
14. Shen, J.C., Gray, M.D., Oshima, J. & Loeb, L.A. (1998). Characterization of Werner syndrome protein DNA helicase activity: directionality, substrate dependence and stimulation by replication protein A. *Nucleic Acids Res.* **26**, 2879-2885.
15. Watt, P.M., Louis, E.J., Borts, R.H. & Hickson, I.D. (1995). Sgs1: a eukaryotic homolog of *E. coli* RecQ that interacts with topoisomerase II in vivo and is required for faithful chromosome segregation. *Cell* **81**, 253-260.
16. Chakraverty, R.K. & Hickson, I.D. (1999). Defending genome integrity during DNA replication: a proposed role for RecQ family helicases. *BioEssays* **21**, 286-294.
17. Hanada, K., Ukita, T., Kohno, Y., Saito, K., Kato, J. & Ikeda, H. (1997). RecQ DNA helicase is a suppressor of illegitimate recombination in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **94**, 3860-3865.
18. Yamagata, K., Kato, J., Shimamoto, A., Goto, M., Furuichi, Y. & Ikeda, H. (1998). Bloom's and Werner's syndrome genes suppress hyperrecombination in yeast sgs1 mutant: implication for genomic instability in human diseases. *Proc. Natl Acad. Sci. USA* **95**, 8733-8738.
19. Yan, H., Chen, C.Y., Kobayashi, R. & Newport, J. (1998). Replication focus-forming activity 1 and the Werner syndrome gene product. *Nat. Genet.* **19**, 375-378.
20. Lebel, M. & Leder, P. (1998). A deletion within the murine Werner syndrome helicase induces sensitivity to inhibitors of topoisomerase and loss of cellular proliferative capacity. *Proc. Natl Acad. Sci. USA* **95**, 13097-13102.
21. Ge, Q., Frank, M.B., O'Brien, C. & Targoff, I.N. (1992). Cloning of a complementary DNA coding for the 100 kD antigenic protein of the PM-Scl autoantigen. *J. Clin. Invest.* **90**, 559-570.
22. Moser, M.J., Oshima, J. & Monnat, R.J., Jr (1999). WRN mutations in Werner syndrome. *Hum. Mutat.* **13**, 271-279.
23. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123-138.
24. Korolev, S., Hsieh, J., Gauss, G.H., Lohman, T.M. & Waksman, G. (1997). Major domain swiveling revealed by the crystal structures of complexes of *E. coli* Rep helicase bound to single-stranded DNA and ADP. *Cell* **90**, 635-647.
25. Subramanya, H.S., Bird, L.E., Brannigan, J.A. & Wigley, D.B. (1996). Crystal structure of a DExx box DNA helicase. *Nature* **384**, 379-383.
26. Velankar, S.S., Soultanas, P., Dillingham, M.S., Subramanya, H.S. & Wigley, D.B. (1999). Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* **97**, 75-84.
27. Sawaya, M.R., Prasad, R., Wilson, S.H., Kraut, J. & Pelletier, H. (1997). Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* **36**, 11205-11215.
28. Liu, D., Prasad, R., Wilson, S.H., DeRose, E.F. & Mullen, G.P. (1996). Three-dimensional solution structure of the N-terminal domain of DNA polymerase beta and mapping of the ssDNA interaction interface. *Biochemistry* **35**, 6188-6200.

29. Subramanya, H.S., Arciszewska, L.K., Baker, R.A., Bird, L.E., Sherratt, D.J. & Wigley, D.B. (1997). Crystal structure of the site-specific recombinase, XerD. *EMBO J.* **16**, 5178-5187.
30. Jeruzalmi, D. & Steitz, T.A. (1998). Structure of T7 RNA polymerase complexed to the transcriptional inhibitor T7 lysozyme. *EMBO J.* **17**, 4101-4113.
31. Sali, A. & Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779-815.
32. Gorbalenya, A.E. & Koonin, E.V. (1993). Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.* **3**, 419-429.
33. Bird, L.E., Subramanya, H.S. & Wigley, D.B. (1998). Helicases: a unifying structural theme? *Curr. Opin. Struct. Biol.* **8**, 14-18.
34. Korolev, S., Yao, N., Lohman, T.M., Weber, P.C. & Waksman, G. (1998). Comparisons between the structures of HCV and Rep helicases reveal structural similarities between SF1 and SF2 super-families of helicases. *Protein Sci.* **7**, 605-610.
35. Kusano, K., Berres, M.E. & Engels, W.R. (1999). Evolution of the RECQ family of helicases: a *Drosophila* homolog, Dmblm, is similar to the human Bloom syndrome gene. *Genetics* **151**, 1027-1039.
36. Fry, M. & Loeb, L.A. (1998). The three faces of the WS helicase. *Nat. Genet.* **19**, 308-309.
37. Matson, S.W., Bean, D.W. & George, J.W. (1994). DNA helicases: enzymes with essential roles in all aspects of DNA metabolism. *BioEssays* **16**, 13-22.
38. Dougherty, W.G., Carrington, J.C., Cary, S.M. & Parks, T.D. (1988). Biochemical and mutational analysis of a plant virus polyprotein cleavage site. *EMBO J.* **7**, 1281-1287.
39. Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX Pipes. *J. Biomol. NMR* **6**, 277-293.
40. Bartels, C., Xia, T.-H., Billeter, M., Güntert, P. & Wüthrich, K. (1995). The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* **5**, 1-10.
41. Clore, G.M. & Gronenborn, A.M. (1998). Determining the structures of large proteins and protein complexes by NMR. *Tibtech* **16**, 22-34.
42. Sattler, M., Schleucher, J. & Griesinger, C. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. NMR Spectrosc.* **34**, 93-158.
43. Yamazaki, T., Forman-Kay, J.D. & Kay, L.E. (1993). Two-dimensional NMR experiments for correlating $^{13}C^{\beta}$ and $^{1}H^{\delta/\varepsilon}$ chemical shifts of aromatic residues in ^{13}C-labeled proteins via scalar couplings. *J. Am. Chem. Soc.* **115**, 11054-11055.
44. Bottomley, M.J., Macias, M.J., Liu, Z. & Sattler, M. (1999). A novel NMR experiment for the sequential assignment of proline residues and proline stretches in $^{13}C/^{15}N$ labeled proteins. *J. Biomol. NMR* **13**, 5381-385.
45. Fesik, S.W. & Zuiderweg, E.R.P. (1988). Heteronuclear three-dimensional NMR spectroscopy. A strategy for the simplification of homonuclear two-dimensional NMR spectra. *J. Magn. Reson.* **78**, 588-593.
46. Kuboniwa, H., Grzesiek, S., Delaglio, F. & Bax, A. (1994). Measurement of HN-Hα J-couplings in calcium-free calmodulin using new 2D and 3D water-flip-back methods. *J. Biomol. NMR.* **4**, 871-878.
47. Hu, J.-S. & Bax, A. (1997). χ_1 angle information from a simple two-dimensional NMR experiment that identifies trans $^{3}J_{NC\gamma}$ couplings in isotopically enriched proteins. *J. Biomol. NMR* **9**, 323-328.
48. Bodenhausen, G. & Ruben, D.J. (1980). Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **69**, 185-189.
49. Neri, D., Szyperski, T., Otting, G., Senn, H. & Wüthrich, K. (1989). Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional ^{13}C labeling. *Biochemistry* **28**, 7510-7516.
50. Brünger, A.T., *et al.,* & Warren, G.L. (1998). Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905-921.
51. Nilges, M. & O'Donoghue, S.I. (1998). Ambiguous NOEs and automated NOESY assignment. *Prog. NMR Spectrosc.* **32**, 107-139.
52. Rice, L.M. & Brünger, A.T. (1994). Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement. *Proteins* **19**, 277-290.
53. Stein, E.G., Rice, L.M. & Brünger, A.T. (1997). Torsion-angle molecular dynamics as a new efficient tool for NMR structure calculation. *J. Magn. Reson.* **124**, 154-164.
54. Folmer, R.H., Hilbers, C.W., Konings, R.N. & Nilges, M. (1997). Floating stereospecific assignment revisited: application to an 18 kDa protein and comparison with J-coupling data. *J. Biomol. NMR* **9**, 245-258.
55. Yang, D., Gardner, K.H. & Kay, L.E. (1998). A sensitive pulse scheme for measuring the backbone dihedral angle ψ based on cross-correlation between $^{13}C\alpha$-$^{1}H\alpha$ dipolar and carbonyl chemical shift anisotropy relaxation interactions. *J. Biomol. NMR* **11**, 213-220.
56. Ottiger, M. & Bax, A. (1997). An empirical correlation between amide deuterium isotope effects on $^{13}C^{\alpha}$ chemical shifts and protein backbone conformation. *J. Am. Chem. Soc.* **119**, 8070-8075.
57. Linge, J.P. & Nilges, M. (1999). Influence of non-bonded parameters on the quality of NMR structures: a new force field for NMR structure calculation. *J. Biomol. NMR* **13**, 51-59.
58. Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R. & 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.
59. Hooft, R.W., Vriend, G., Sander, C. & Abola, E.E. (1996). Errors in protein structures. *Nature* **381**, 272.
60. Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51-55.
61. Nicholls, A., Sharp, K.A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 282.
62. Piotto, M., Saudek, V. & Sklenar, V. (1992). Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**, 661-665.
63. Grzesiek, S. & Bax, A. (1993). The importance of not saturating H_2O in protein NMR. Application to sensitivity enhancement and NOE measurements. *J. Am. Chem. Soc.* **115**, 12593-12594.
64. Dailey, L., Yuan, H. & Basilico, C. (1994). Interaction between a novel F9-specific factor and octamer-binding proteins is required for cell-type-restricted activity of the fibroblast growth factor 4 enhancer. *Mol. Cell. Biol.* **14**, 7758-7769.
65. Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680.

Because *Structure with Folding & Design* operates a 'Continuous Publication System' for Research Papers, this paper has been published on the internet before being printed (accessed from http://biomednet.com/cbiology/str). For further information, see the explanation on the contents page.