J Biomol NMR (2016) 66:209–219
DOI 10.1007/s10858-016-0070-9

ARTICLE

Solution NMR structure of the HLTF HIRAN domain: a conserved module in SWI2/SNF2 DNA damage tolerance proteins

Dmitry M. Korzhnev[1] · Dante Neculai[2] · Sirano Dhe-Paganon[3] · Cheryl H. Arrowsmith[4,5] · Irina Bezsonova[1]

Received: 30 June 2016 / Accepted: 17 October 2016 / Published online: 22 October 2016

Abstract HLTF is a SWI2/SNF2-family ATP-dependent chromatin remodeling enzyme that acts in the error-free branch of DNA damage tolerance (DDT), a cellular mechanism that enables replication of damaged DNA while leaving damage repair for a later time. Human HLTF and a closely related protein SHPRH, as well as their yeast homologue Rad5, are multi-functional enzymes that share E3 ubiquitin-ligase activity required for activation of the error-free DDT. HLTF and Rad5 also function as ATP-dependent dsDNA translocases and possess replication fork reversal activities. Thus, they can convert Y-shaped replication forks into X-shaped Holliday junction structures that allow error-free replication over DNA lesions. The fork reversal activity of HLTF is dependent on 3′-ssDNA-end binding activity of its N-terminal HIRAN domain. Here we present the solution NMR structure of the human HLTF HIRAN domain, an OB-like fold module found in organisms from bacteria (as a stand-alone domain) to plants, fungi and metazoan (in combination with SWI2/SNF2 helicase-like domain). The obtained structure of free HLTF HIRAN is similar to recently reported structures of its DNA bound form, while the NMR analysis also reveals that the DNA binding site of the free domain exhibits conformational heterogeneity. Sequence comparison of N-terminal regions of HLTF, SHPRH and Rad5 aided by knowledge of the HLTF HIRAN structure suggests that the SHPRH N-terminus also includes an uncharacterized structured module, exhibiting weak sequence similarity with HIRAN regions of HLTF and Rad5, and potentially playing a similar functional role.

Keywords DNA replication · DNA damage tolerance · Helicase-like transcription factor · HLTF

✉ Irina Bezsonova
bezsonova@uchc.edu

[1] Department of Molecular Biology and Biophysics, University of Connecticut Health, Farmington, CT 06030, USA

[2] School of Medicine, Zhejiang University, Hangzhou, China

[3] Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA

[4] Structural Genomics Consortium, University of Toronto, Toronto, ON M5G 1L7, Canada

[5] Department of Medical Biophysics, Princess Margaret Cancer Centre, University of Toronto, Toronto, ON M5S 1A8, Canada

Introduction

Human HLTF (Helicase-Like Transcription Factor) and its functional homologues, *S. cerevisiae* Rad5 and human SHPRH (SNF2, histone linker, PHD, RING, helicase), are SWI2/SNF2-family ATP-dependent chromatin remodeling factors that promote error-free DNA damage tolerance (DDT) in eukaryotes (Chang and Cimprich 2009; Unk et al. 2010). To cope with exogenous and endogenous DNA damage that creates DNA replication blocks, cells have evolved DDT mechanisms that enable bypass replication over sites of DNA damage lesions (Chang and Cimprich 2009; Waters et al. 2009; Unk et al. 2010; Sale et al. 2012). The error-prone branch of DDT, which involves DNA lesion bypass by mutagenic translesion synthesis (TLS) DNA polymerases (Waters et al. 2009; Sale et al. 2012), is activated by mono-ubiquitination of the DNA-sliding clamp PCNA by the Rad6/Rad18 E2/E3 enzyme pair (Hoege et al. 2002). Subsequently, the error-free DDT that utilizes a template switching mechanism can be activated by K63-polyubiquitination of PCNA catalyzed by the

RING E3 ligases HLTF, SHPRH or Rad5 and the Ubc13/Mms2 E2 conjugating enzyme (Motegi et al. 2006, 2008; Unk et al. 2006, 2008, 2010; Chang and Cimprich 2009).

One possible mechanism of the error-free DDT is replication fork reversal that involves conversion of a Y-shaped DNA structure into an X-shaped four-way Holliday junction structure with one arm composed of the newly synthesized daughter DNA strands annealed against one another (Chang and Cimprich 2009; Unk et al. 2010). Such a 'chicken foot' intermediate allows switching DNA synthesis blocked by DNA damage from the lesion-containing strand to the undamaged template newly synthesized at the complementary DNA strand. Beyond serving as E3 ligases that polyubiquitinate PCNA and trigger switching to the error-free DDT (Motegi et al. 2006, 2008; Unk et al. 2006, 2008), HLTF and Rad5 function as ATP-dependent dsDNA translocases and can catalyze replication fork reversal into a Holliday-like structure (Blastyak et al. 2007, 2010). The fork reversal activity, however, has not yet been demonstrated for SHPRH.

The core helicase-like domains of HLTF, SHPRH and Rad5 consist of the two recA-like domains and share SWI2/SNF2 architectures, featuring seven conserved helicase motifs (Thoma et al. 2005; Flaus et al. 2006). The helicase motifs are separated by the two ('minor' and 'major') insert regions that may include additional domains embedded into a core domain (Flaus et al. 2006). The 'major' insert regions of HLTF, SHPRH and Rad5 harbor RING-finger domains characteristic of the E3 ubiquitin ligases, while the 'minor' insert region of SHPRH also includes H1.5 and PHD finger domains. Such a discontinuous arrangement complicates structural studies of SWI2/SNF2 DDT enzymes; at this time, the only structure available from the core helicase-like regions is that of PHD domain embedded in the 'minor' insert region of human SHPRH (Machado et al. 2013).

HLTF, also called HIP116, SMARCA3 and other alternative names, was first identified as a transcription factor for plasminogen activator inhibitor 1 (PAI-1) (Ding et al. 1996, 1999) and β-globin genes (Mahajan and Weissman 2002) capable of binding specific DNA target sequences (Sheridan et al. 1995; Ding et al. 1996, 1999; Mahajan and Weissman 2002). In these original studies, the DNA-binding region of HLTF was mapped to its N-terminal part encompassing residues 38–219 (Sheridan et al. 1995; Ding et al. 1996) preceding the core helicase-like domain. Later, this N-terminal region was shown to include a standalone β-strand rich domain conserved in eukaryotic SWI2/SNF2 DNA-dependent ATPases dubbed as HIRAN (HIP116, Rad5p N-terminal domain) (Iyer et al. 2006). This domain was also found as a standalone protein in a range of bacteria (Iyer et al. 2006). Based on its contextual information network, the HIRAN domain was predicted to recognize DNA features at stalled replication forks such as ssDNA stretches or DNA lesions (Iyer et al. 2006). Although the HIRAN domain was also found in *S. cerevisiae* Rad5, this module is reportedly lacking in the HLTF's human homologue SHPRH (Unk et al. 2008, 2010).

Recently, we and others have demonstrated that the HLTF HIRAN domain is involved in 3′ ssDNA end recognition at replication forks and is required for fork reversal, but not for other activities of HLTF (Achar et al. 2015; Hishiki et al. 2015; Kile et al. 2015). These works have also reported structures of the HLTF HIRAN domain in complexes with DNA substrates revealing structural basis of the 3′ DNA end recognition. Here we complement these studies by reporting a high-resolution structure of the free HIRAN domain from human HLTF determined by solution NMR spectroscopy. Based on the obtained structure and sequence alignment of the N-terminal regions of HLTF, SHPRH and Rad5 we predict that the N-terminal region of human SHPRH preceding the core helicase-like domain also includes yet uncharacterized structured module, exhibiting weak sequence similarity with HIRAN regions of HLTF and Rad5. These findings highlight a potential conserved role of the N-terminal regions of SWI2/SNF2 chromatin remodeling enzymes in mediating error-free DNA damage tolerance.

Materials and Methods

Protein sample preparation

$^{15}N/^{13}C$ labeled human HLTF HIRAN (residues 50–171) was expressed in BL21 (DE3) *E. coli* cells transformed with pET28a-LIC based plasmid encoding the domain (Addgene plasmid # 28152). Cells were grown in minimal M9 medium supplemented with $^{15}NH_4Cl$ and ^{13}C-glucose as sole nitrogen and carbon sources, respectively. The protein was affinity purified on a Ni–NTA agarose beads (GE Healthcare), followed by cleavage of a hexahistidine tag by Thrombin (MP Biomedicals) and a second purification step on a size-exclusion Superdex 75 HiLoad 16/60 chromatography column (GE Healthcare). The final NMR sample of the free HIRAN domain contained 1.0 mM protein, 20 mM sodium phosphate pH 6.8, 100 mM NaCl, 2 mM DTT and 10 % v/v D_2O.

NMR spectroscopy and protein structure calculation

The backbone and side-chain ^{15}N, ^{13}C and ^{1}H NMR resonance assignments of the HLTF HIRAN domain were obtained from a set of 2D $^{1}H–^{15}N$ HSQC, $^{1}H–^{13}C$ HSQC and 3D HNCA, HNCACB, HNCO, HBHA(CO)NH,

HC(C)H-TOCSY, (H)CCH-TOCSY and ^{15}N- and ^{13}C-edited NOESY-HSQC (150 ms mixing time) spectra (Kay 1995, 1997; Kanelis et al. 2001) collected at 25 °C on the Agilent VNMRS 800 MHz spectrometer equipped with a cold probe. NMR spectra were processed with NMRPipe (Delaglio et al. 1995) and analyzed with Sparky software (Goddard and Kneller), resulting in 86 % assignment of all expected NMR chemical shifts, including 98 % of backbone, 82 % of aliphatic and 53 % of aromatic side-chain. Of note, PDB NMR structure validation report indicates one outlier chemical shift for $^{1}H^{\alpha}$ of Arg 157 (1.92 ppm) likely due to a ring current shift involving the nearby aromatic side chain of Phe 60. NMR resonance assignments of the HLTF HIRAN domain were deposited to BioMagResBank with the accession number *30098*.

Solution NMR structure of the HLTF HIRAN domain was determined based on 2343 NOE-derived ^{1}H–^{1}H distance restraints obtained from 3D ^{15}N- and ^{13}C-edited NOESY-HSQC spectra (Kay 1995, 1997; Kanelis et al. 2001), and 192 dihedral φ and ψ angle restraints derived from the backbone ^{1}H, ^{15}N, and ^{13}C chemical shifts using TALOS+ program (Shen et al. 2009). The NOE peaks were assigned automatically and the ensemble of 100 structures was calculated using CYANA software (Guntert 2004; Guntert and Buchner 2015). The 20 structures with the lowest energy were further refined in explicit solvent using CNS (Brunger et al. 1998; Brunger 2007). The restraints used for structure calculation and structure refinement statistics are summarized in Table 1.

Table 1 Summary of NMR-based restrains used for structure calculation of the human HLTF HIRAN domain and structure refinement statistics

Summary of restraints	
NOE distance restraints	
Short range ($\|i - j\| \leq 1$)	1186
Medium range ($1 < \|i - j\| < 5$)	387
Long range ($\|i - j\| \geq 5$)	770
Total	2343
Dihedral angles (φ and ψ)	192
Hydrogen bonds	40
Deviation from experimental restraints	
NOE (Å)	0.016 ± 0.002
Dihedral restraints (°)	1.141 ± 0.154
Deviation from idealized geometry	
Bonds (Å)	0.014 ± 0.000
Angles (°)	0.925 ± 0.023
Ramachandran plot statistics	
Favored regions	91 ± 2 %
Additionally allowed regions	8 ± 2 %
Outliers	1 ± 1 %
RMSD from mean structure (Å)	
All residues	
Backbone atoms	0.87 ± 0.07
Heavy atoms	1.40 ± 0.10
Residues in ordered regions[a]	
Backbone atoms	0.78 ± 0.05
Heavy atoms	1.22 ± 0.14
RPF quality scores Huang et al. (2005, 2006)	
Recall	0.95
Precision	0.89
F-score	0.92

[a] The following secondary structure elements were included as ordered regions in the RMSD calculation: α-helices 52–56, 113–124, 157–167; β-strands 57–67, 81–87, 98–102, 107–110, 129–134, 144–153

Protein domain architecture analysis

The domain architectures of 766 prokaryotic and eukaryotic proteins that contain annotated HIRAN domain were analyzed using Simple Modular Architecture Research Tool (SMART) (Schultz et al. 1998; Letunic et al. 2006, 2015).

The atomic coordinates along with NMR chemical shift assignments and NMR restraints were deposited to Protein Data Bank (5K5F) and Biological Magnetic Resonance Bank (30098).

Results and Discussion

NMR structure of the free HLTF HIRAN domain

To determine solution NMR structure of the free form of ^{15}N/^{13}C labeled human HLTF HIRAN domain we performed the backbone and side-chain resonance assignments of the domain using a set of standard triple-resonance NMR experiments (Kay 1995, 1997; Kanelis et al. 2001). Figure 1 shows ^{1}H–^{15}N HSQC spectrum of the domain with resonance assignments for the backbone amide groups annotated. Solution structure of the HIRAN domain was calculated based on NOE derived ^{1}H-^{1}H distance restrains and chemical shift derived backbone torsion angle restraints, as described in Materials and Methods. Figure 2a illustrates the resulting structural ensemble (backbone heavy atoms only), which is in good agreement with the input restraints (Table 1). The resulting ensemble is well defined with mean pairwise RMSD among 20 best models of 0.87 ± 0.07 Å for the backbone and 1.40 ± 0.10 Å for all heavy atoms.

The isolated HLTF HIRAN domain adopts a mixed α/β fold with (α0)–β1–β2–β3–β4–α1–β5–β6–α2 topology (Fig. 2b, c). The six β-strands are organized into a 'greek key' with strands arranged from longest (β1) to shortest

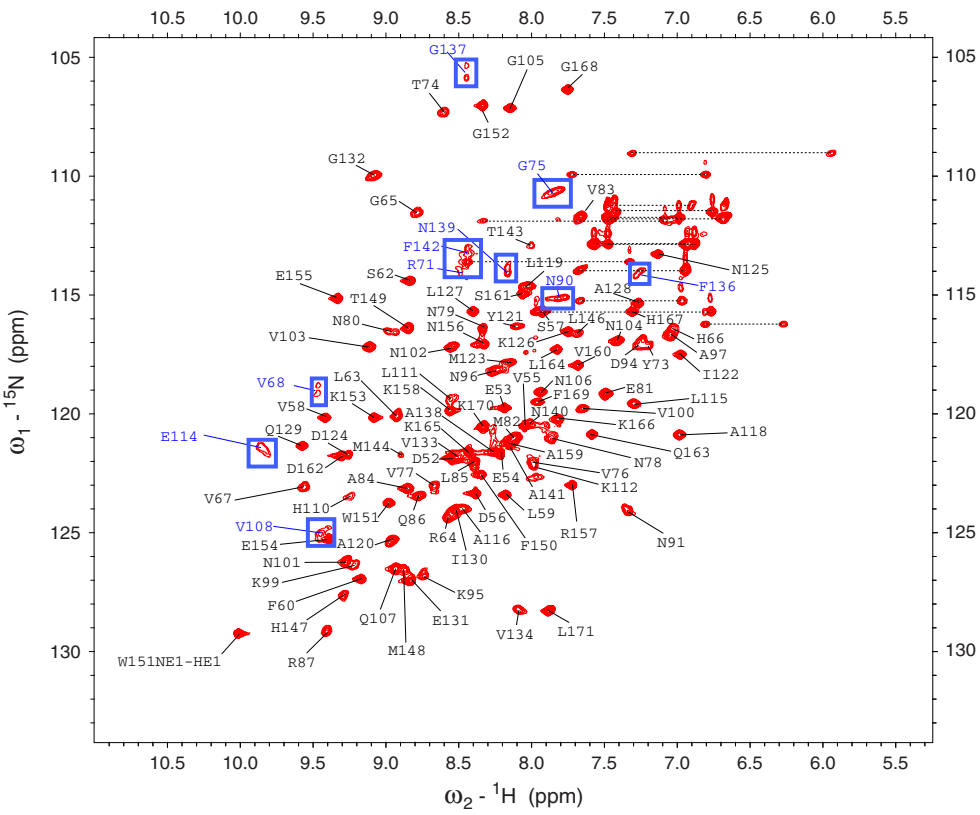


Fig. 1 Backbone NMR chemical shift assignment of the HLTF HIRAN domain. $^{1}H–^{15}N$ HSQC spectrum of the human HIRAN domain with resonance assignments annotated. Asn and Gln side chain peaks are indicated by *dotted lines*. Due to slow to intermediate conformational exchange several amide groups display multiple broaden peaks enclosed in *blue boxes*

(β4) (Fig. 2b), and form a diagonally truncated β1–β6–β5–β2–β3–β4 β-barrel with a wide opening on one side covered with the two α-helices (α1 and α2) creating an OB-like (Oligonucleotide/Oligosaccharide-Binding) fold (Murzin 1993; Arcus 2002; Agrawal and Kishan 2003; Kerr et al. 2003; Theobald et al. 2003; Fig. 2c). The resulting NMR structure also contains a helical turn α0 formed by the N-terminal residues 52–56 of the domain. Figure 2d shows surface representation of HLTF HIRAN colored by electrostatic potential (from –5 to 5 kcal mol^{-1} e^{-1}, red to blue), revealing that the domain has two oppositely charges faces. The positively charged face is involved in recognition of DNA-substrates, as described several recent works (Achar et al. 2015; Hishiki et al. 2015; Kile et al. 2015).

Recently, we and others have shown that HLTF HIRAN specifically recognizes 3′ ssDNA end found at replication forks, and determined structures of the HIRAN domain bound to single-stranded ssDNA substrates (Hishiki et al. 2015; Kile et al. 2015) and a dinucleotide (Achar et al. 2015). Figure 2e shows superposition of NMR structure of the free HIRAN domain reported here and the previously reported X-ray crystal structure of its complex with $(dT)_{10}$ oligonucleotide (Kile et al. 2015). The two structures agree well with one another with the backbone RMSD between the free and bound forms of 1.8 Å, suggesting that the HLTF HIRAN domain does not undergo significant structural changes upon binding a DNA substrate. Similar level of agreement has been observed between the free HIRAN domain and its complexes with other ssDNA and oligonucleotide substrates (Achar et al. 2015; Hishiki et al. 2015; Kile et al. 2015).

Previous studies have shown that residues of the HLTF HIRAN domain involved in 3′ ssDNA end recognition are located in β1–β2 loop (V68, R71, Y72, Y73), β2–β3 loop (N91, Y93, D94), β4 strand and β4–α1 loop (H110, K113), and β5–β6 loop (F142) (Flaus et al. 2006; Achar et al. 2015; Hishiki et al. 2015; Kile et al. 2015). Comparison of solution NMR structure of the free HIRAN domain with the structure of its complex with ssDNA (Fig. 2e) suggests that this site does not undergo significant change in overall conformational upon substrate binding. However, an inspection of $^{1}H–^{15}N$ HSQC spectrum of the free HLTF HIRAN domain (Fig. 1) revealed decreased peak intensities for a number of residues from the DNA binding region (Hishiki et al. 2015; Kile et al. 2015). Instead of a single peak expected for each backbone amide group, these residues display a set of multiple broadened resonances in slow to intermediate exchange with one another (enclosed in blue boxes in Fig. 1). The residues experiencing conformational exchange (V68, R71, G75 from β1 to β2 loop; N90 from β2 to β3 loop; V108 from strand β4; E114 from β4 to α1; F136, G137, N139, F142 from β5 to β6 loop) line up precisely at the DNA-binding grove of the HIRAN

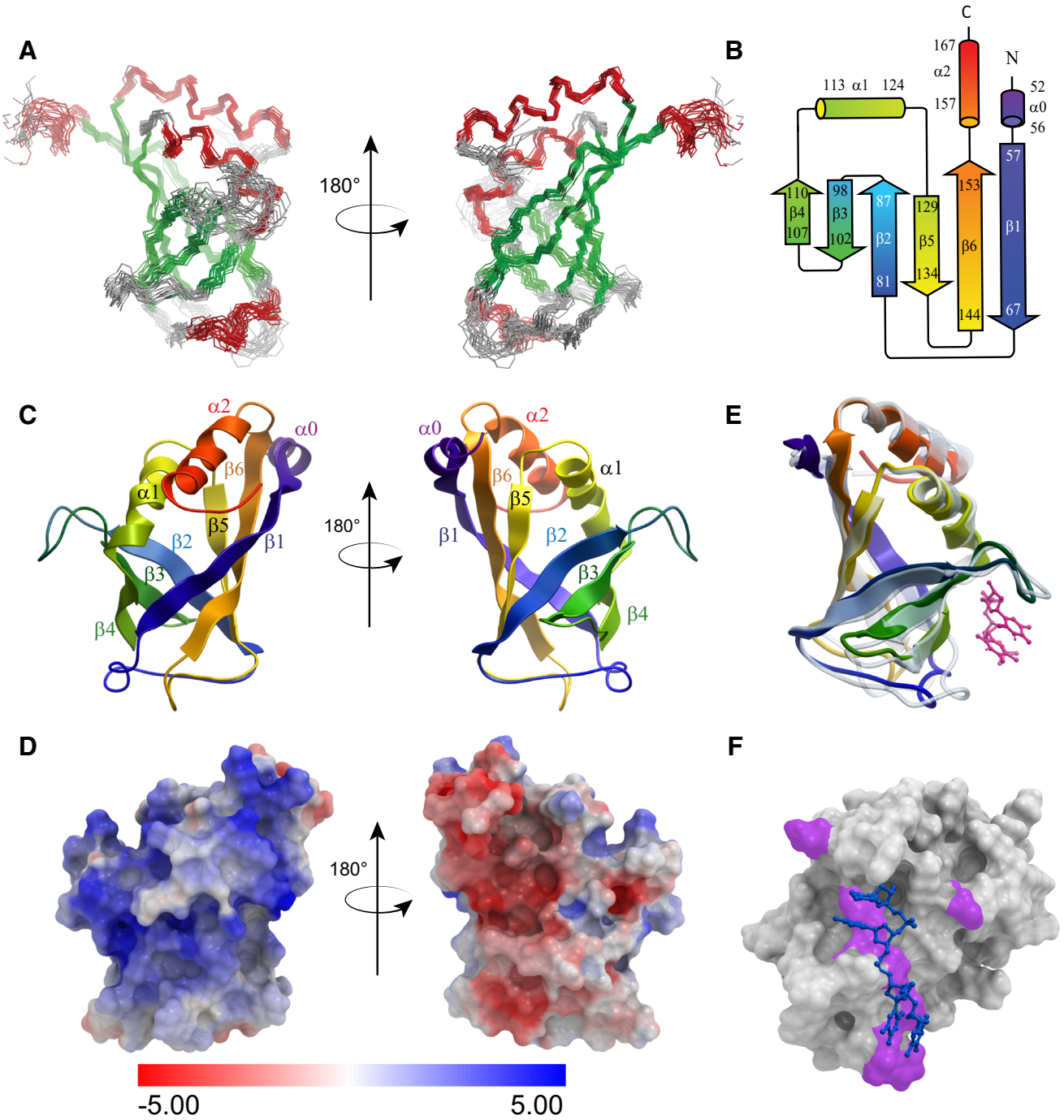


Fig. 2 Solution structure of the HLTF HIRAN domain. **a** Ensemble of 20 lowest energy NMR structures of the HIRAN domain (backbone only); β-strands and α-helices are colored in green and red, respectively. **b** Secondary structure schematic of the HLTF HIRAN domain colored from purple (N-terminus) to red (C-terminus). **c** Ribbon representation of the HIRAN domain structure **d** HIRAN domain surface colored by qualitative electrostatic potential (the same orientation as in panel **c**). ICM was used to calculate the electrostatic potential using the boundary element solution of the Poisson equation. The electrostatic potential was projected onto the molecular surface using a color scale from *red* (-5 kcal mol^{-1} e^{-1}) to *blue* (5 kcal mol^{-1} e^{-1}). **e** Superposition of the NMR structure of the free HIRAN (rainbow) and X-ray crystal structure of its complex with $(dT)_{10}$ oligonucleotide (PDB: 4S0 N; transparent). The two 3′nucleotides are shown in magenta. **f** Residues exhibiting conformational exchange in solution are mapped in magenta on the surface of the HLTF HIRAN domain bound to a single stranded DNA oligonucleotide (PDB: 4S0 N). All molecular structures were rendered using ICM software (*Molsoft*)

domain, as shown in Fig. 2f. This observation clearly suggests that the DNA-binding region of the free HLTF HIRAN domain is dynamic and pre-samples multiple slowly exchanging conformations in solution.

HIRAN is prevalent in SWI2/SNF2 DNA damage tolerance proteins

In eukaryotes, HIRAN domain is almost exclusively found in association with a subfamily of SWI2/SNF2 DNA-dependent ATPases containing RING finger domain embedded in the SWI2/SNF2 helicase-like domain (Flaus et al. 2006; Iyer et al. 2006). HIRAN is typically located N-terminal to helicase-like domain, which in SMART database can be identified as a combination of DEXDc (or SNF2_N) and HELICc domains with the RING-finger inserted in-between. The evolutional analysis of the domain architecture of all HIRAN domain-containing proteins found in SMART database suggests that this domain organization is prevalent in eukaryotes (Fig. 3). Thus, 62 % of HIRAN domains are coupled with SWI2/SNF2 helicase-like and RING-finger domains in green plants, 98 % in fungi and 94 % in animals. This domain organization is also present in HLTF, which is the only human protein described so far containing the HIRAN domain. In this domain architecture, the DEXDc/HELICc helicase-like domain has ATP-dependent DNA translocase activity, while the RING domain confers E3 ubiquitin ligase activity required for K63-polyubiquitination of PCNA, signaling activation of 'error-free' DDT (Waters et al. 2009; Branzei and Psakhye 2016; Callegari and Kelly 2016). Recent reports suggest that the HIRAN domain is a DNA-binding module that specifically recognizes 3′-end of a single stranded DNA found at replication forks, which is required for HLTF's fork reversal activity (Achar et al. 2015; Hishiki et al. 2015; Kile et al. 2015). Therefore, the evolutional analysis indicates that in eukaryotes HIRAN mainly functions as a 3′-end DNA recognition module that targets SWI2/SNF2 DDT enzymes to replication forks and enables fork regression, initiating 'error-free' replicative

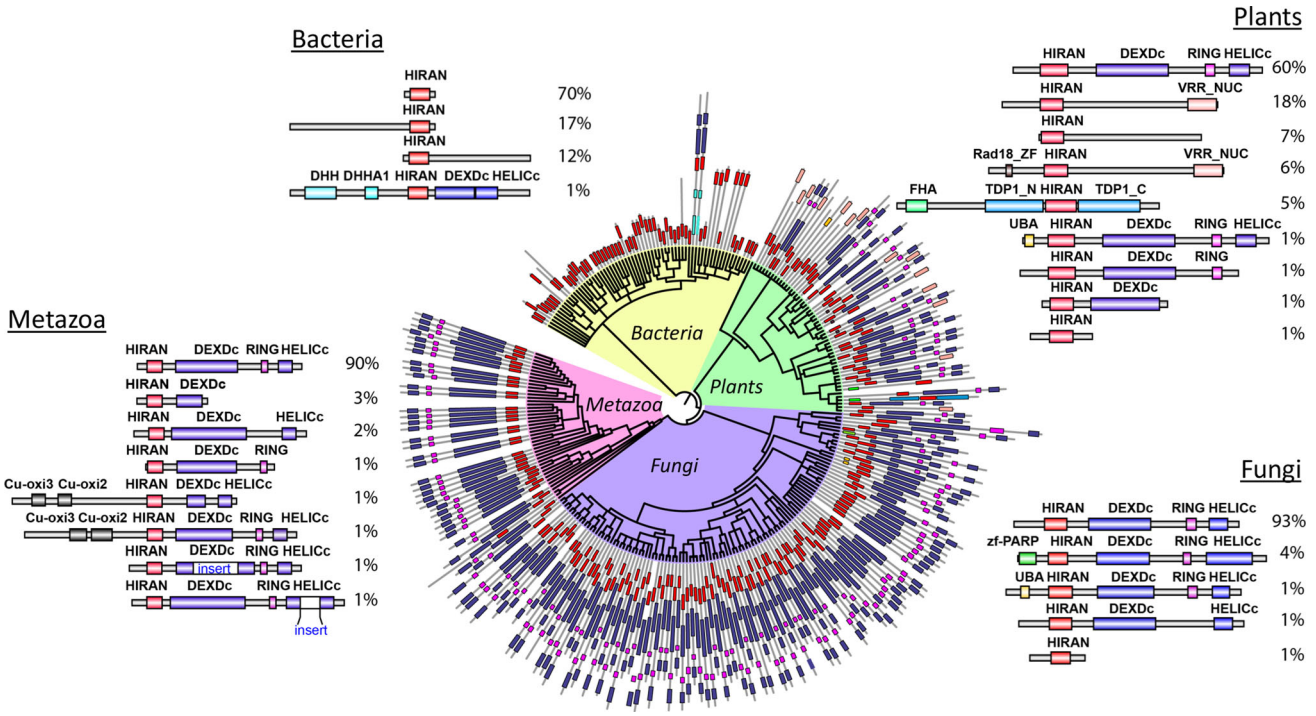


Fig. 3 Evolution of a HIRAN domain. Phylogenetic tree of HIRAN domain-containing proteins generated using SMART database and iTOL (Letunic and Bork 2007, 2011). Representative domain architectures of the HIRAN-containing proteins and their percent distribution in each class of organisms is shown. A total of 766 proteins containing the HIRAN domain were found in SMART database including 278 bacterial, 283 fungal, 88 green plant and 117 metazoan proteins. Domain are labeled as follows: HIRAN (HIP116/RAD5N-terminal domain, SM00910), DEXDc (DEAD-like Superfamily, SM00487) and HELICc (Helicase Superfamily C-terminal, SM00490) that form the SWI2/SNF2 helicase-like domain, VRR_NUC (Virus-type Replication-Repair Nuclease domain, SM00990), DHH (DHH-family phosphoesterase domain, PF01368), DHHA1 (DHH Associated domain 1, PF02272), TDP1 (Tyrosyl-DNA Phosphodiesterase domain, PF06087), FHA (Forkhead associated domain, SM00240), RING (Really Interesting New Gene, or Ring finger domain, SM00184), UBA (Ubiquitin Associated domain, SM00165), Rad18_ZF (Rad18-like CCHC Zinc Finger domain, SM00734), zf-PARP (Poly(ADP-ribose) polymerase type Zn-finger domain, PF00645), Cu-oxi2 (Multicopper oxidase, type 2, PF07731), Cu-oxi3 (Multicopper oxidase, type 3, PF07732)

bypass of DNA lesions by a template switching mechanism (Waters et al. 2009; Branzei and Psakhye 2016; Callegari and Kelly 2016).

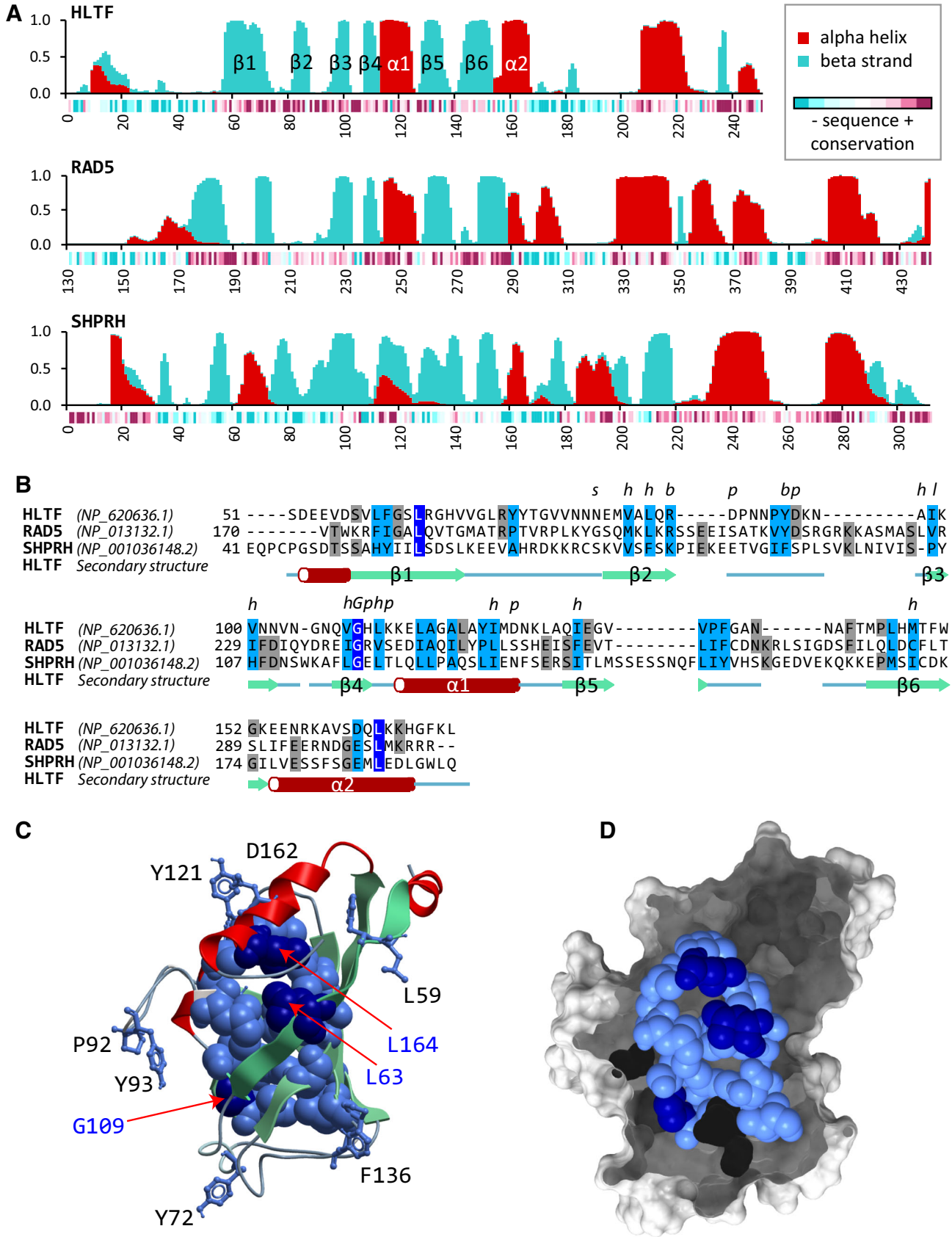
While prevalent in eukaryotes, a combination of HIRAN with SWI2/SNF2 helicase-like and RING-finger domains is not unique (Fig. 3). The phylogenetic tree of HIRAN domain suggests that it is an ancient module that first appears as a standalone protein in prokaryotes, where small single-domain proteins account for 70 % of all HIRAN-containing members, while the remaining 29 % include HIRAN at C- (17 %) or N- (12 %) terminus of a longer polypeptide chain. Note that in about 1 % of the considered bacterial proteins HIRAN is already found in combination with SWI2/SNF2 helicase-like (DEXDc/HELICc) and phosphatase DHH/DHHA1 domains (Fig. 3). While in fungi and metazoans there is little variation in architecture of HIRAN domain-containing proteins, in green plants HIRAN is found in several distinct classes of proteins. Thus, in addition to domains characteristic of SWI2/SNF2 DNA-dependent ATPases (63 %), HIRAN is also found in the context of restriction endonuclease VRR_NUC (24 %) and phosphatase TDP1 (5 %) domains and likely functions as 3′-end recognition module in other DNA-modifying enzymes.

Conservation of the N-terminal domains in HLTF, SHPRH and Rad5

HLTF and a closely related SWI2/SNF2-family enzyme SHPRH were described as the two human homologues of *S. cerevisiae* Rad5 involved in 'error-free' DDT (Chang and Cimprich 2009; Unk et al. 2010). Both HLTF and SHPRH function as RING E3 ubiquitin ligases that, together with Ubc13/Mms2 E2 enzyme, catalyze K63-polyubiquitination of PCNA, promoting 'error-free' bypass of DNA lesions (Motegi et al. 2006, 2008; Unk et al. 2006, 2008, 2010; Chang and Cimprich 2009). SHPRH reportedly lacks the HIRAN domain found in the N-terminal parts of HLTF and Rad5 (Unk et al. 2010). Furthermore, the analysis of SNF2 family of helicase-related proteins suggests that HLTF and Rad5 belong to the same (Rad5/16-like) subfamily, while SHPRH was placed to a separate (SHPRH-like) subfamily (Flaus et al. 2006). Thus, unlike HLTF and Rad5, the 'minor' insert region of SHPRH includes two unique domains, H1.5 (linker histone H1 and H5) and PHD (plant homeodomain), in addition to the RING finger domain embedded in the 'major' insert region of the helicase-like domain. Recent reports established a strict requirement of the HIRAN domain for replication fork reversal by HLTF (Achar et al. 2015; Kile et al. 2015). Consistent with this, the fork reversal activity was reported for HLTF and Rad5 that possess the N-terminal HIRAN domain (Blastyak et al. 2007, 2010), while such activity has not yet been confirmed for SHPRH suggesting a possible divergence of HLTF and SHPRH functions.

Interestingly, SHPRH has a 308 residue N-terminal region with the length comparable to those of N-terminal parts of HLTF (240 residues) and Rad5 (437 residues) preceding the Q-box (first conserved element) of the SWI2/SNF2 helicase-like domain. Figure 4a shows a secondary structure prediction for the N-terminal regions of HLTF, Rad5 and SHPRH performed in RaptorX (Wang et al. 2011). As expected, a structured domain with the HIRAN topology (β1–β2–β3–β4–α1–β5–β6–α2) was predicted in the N-terminal regions of HLTF (57–167) and Rad5 (170–310) separated from the helicase-like domain by a flexible linker. Remarkably, the N-terminal region of SHPRH is also predicted to include a β-strand rich domain of similar size (residues 45–220), although the exact topology of this domain is uncertain.

To our surprise, a multiple sequence alignment shown in Fig. 4b revealed a weak homology of the SHPRH N-terminal domain with the HIRAN domains of HLTF (11 % identity; 26 % similarity) and Rad5 (12 % identity; 33 % similarity; Table 2), which is comparable to the homology between the HIRAN domains of HLTF and Rad5 (14 % identity; 40 % similarity). In the alignment shown in Fig. 4b, the residues that are similar/identical among the N-terminal domains of HLTF, Rad5 and SHPRH are marked in light/dark blue. The multiple sequence alignment revealed that the N-terminal domains of HLTF, Rad5 and SHPRH share key residues characteristic of the HIRAN domain (as defined by Iyer et al. 2006), including the absolutely conserved Gly residue (G109 in HLTF, G240 in Rad5, G117 in SHPRH), as well as a pattern of polar and hydrophobic residues shown on the top of the sequence (Fig. 4b). Furthermore, mapping of the residues that are similar/identical among HLTF, Rad5 and SHPRH onto the HLTF HIRAN structure (light/dark blue balls in Fig. 4c, d) suggests that hydrophobic residues conserved in all three proteins form a hydrophobic core of the HIRAN domain. In addition, the N-terminal domain of SHPRH shares with the HLTF and Rad5 HIRAN domains several key residues involved in 3′ ssDNA end recognition. Thus, the aromatic side-chain of Y93 in HLTF makes stacking interactions with a ring of nucleotide base, while the following polar residue D94 interacts with free 3′-OH group of the deoxyribose (Achar et al. 2015; Kile et al. 2015). In Rad5 the Y213/D214 pair is absolutely conserved, while in SHPRH the corresponding residues are aromatic F92 followed by a polar S93. Therefore, considering the presumed similar functions of HLTF, SHPRH and Rad5 in 'error-free' DDT, (Chang and Cimprich 2009; Unk et al. 2010) one may hypothesize that the identified N-terminal 'HIRAN-like' domain of SHPRH plays the same functional


A
HLTF
RAD5
SHPRH
β1 β2 β3 β4 α1 β5 β6 α2
alpha helix
beta strand
- sequence +
conservation
B
HLTF (NP_620636.1) 51 ----SDEEVDSVLFGSLRGHVVGLRYYTGVVNNNEMVALQR-----DPNNPYDKN---------AIK
RAD5 (NP_013132.1) 170 -------VTWKRFIGALQVTGMATRPTVRPLKYGSQMKLKRSSEEISATKVYDSRGRKKASMASLVR
SHPRH (NP_001036148.2) 41 EQPCPGSDTSSAHYIILSDSLKEEVAHRDKKRCSKVVSFSKPIEKEETVGIFSPLSVKLNIVIS-PY
HLTF (NP_620636.1) 100 VNNVN-GNQVGHLKKELAGALAYIMDNKLAQIEGV--------VPFGAN------NAFTMPLHMTFW
RAD5 (NP_013132.1) 229 IFDIQYDREIGRVSEDIAQILYPLLSSHEISFEVT--------LIFCDNKRLSIGDSFILQLDCFLT
SHPRH (NP_001036148.2) 107 HFDNSWKAFLGELTLQLLPAQSLIENFSERSITLMSSESSNQFLIYVHSKGEDVEKQKKEPMSICDK
HLTF (NP_620636.1) 152 GKEENRKAVSDQLKKHGFKL
RAD5 (NP_013132.1) 289 SLIFEERNDGESLMKRRR--
SHPRH (NP_001036148.2) 174 GILVESSFSGEMLEDLGWLQ
HLTF Secondary structure
C
D162
Y121
L59
L164
L63
P92
Y93
G109
F136
Y72
D

◀**Fig. 4** Conservation of N-terminal domains in HLTF, SHPRH and Rad5. **a** Per-residue secondary structure propensities calculated by RaptorX Property (Wang et al. 2011; Kallberg et al. 2014) for the N-terminal regions of HLTF, SHPRH and Rad5 are shown as stacked bars. Helical and beta-strand propensities for each residue are shown as red and cyan bars, respectively. Amino acid sequence conservation for each protein as determined by ConSurf (Armon et al. 2001; Goldenberg et al. 2009; Ashkenazy et al. 2010) is shown underneath the histograms and color coded from magenta (conserved) to cyan (variable). **b** Multiple sequence alignment of the HIRAN domains of yeast Rad5 (170–306) and human HLTF (51–171), and the N-terminal HIRAN-like domain of human SHPRH (41–194) performed using Clustal Omega (Sievers and Higgins 2014) and then refined manually. The residues that are similar/identical in all three domains are *shaded light/dark blue*. The residues identical in only two of three sequences are shaded in *gray*. A conserved pattern of amino acid residues characteristic of HIRAN fold as defined by Iyer et al. (2006) is shown on the top of the sequence (*h* hydrophobic, *b* big, *p* polar, *l* aliphatic, *G* absolutely conserved Glycine). Secondary structure elements of HLTF HIRAN domain are shown below the alignment. **c** Ribbon representation of the HLTF HIRAN domain with side chains of buried conserved residues shown as space-filling atoms. Side chains of conserved surface-exposed residues are shown in ball and stick representation (color scheme is the same as in **b**). **d** Buried conserved residues displayed beneath the surface of HLTF HIRAN, suggesting that residues conserved among HLTF, Rad5 and SHPRH form a continuous core of the HIRAN domain. All molecular structures were rendered using ICM software (*Molsoft*)

Table 2 Sequence identity/similarity between the N-terminal (HIRAN) domains of yeast Rad5 and its human homologues HLTF and SHPRH

	HLTF	RAD5	SHPRH	Similarity
HLTF		40 %	25.8 %	HLTF
RAD5	13.9 %		32.9 %	RAD5
SHPRH	10.7 %	12.4 %		SHPRH
Identity	HLTF	RAD5	SHPRH	

Percent identity (PID) was calculated as $PID = 100 \times \left(\frac{Identical\ residues}{Length\ of\ shortest\ sequence}\right)$

To calculate % similarity the amino acids were grouped into four categories: aromatic (F, Y, W), aliphatic (V, I, L), positively charged (R, K, H), negatively charged (D, E), polar (N, Q, T, S, E, D) and small (A, G, T, S)

role in a DNA substrate recognition as the HIRAN domain of HLTF. This implies yet unconfirmed ability of SHPRH to mediate replication fork reversal, similar to that of HLTF and Rad5, (Blastyak et al. 2007, 2010), which needs to be tested in subsequent works.

Conclusions

We have determined solution NMR structure of the free form of human HLTF HIRAN domain. Recently, we and others have shown that the HIRAN domain is involved in 3′ ssDNA end recognition and mediates replication fork reversal by HLTF, facilitating 'error-free' replicative bypass of DNA lesions by a template switching mechanism (Achar et al. 2015; Hishiki et al. 2015; Kile et al. 2015). Our phylogenetic analysis suggests that in eukaryotes the HIRAN domain is almost exclusively found in the N-terminal parts of the SWI2/SNF2 ATP-dependent DNA translocases, likely, conferring these enzymes replication fork reversal activity. Interestingly, the HIRAN domain has been previously identified in human HLTF and its yeast homologue Rad5, but not in a functionally related enzyme, human SHPRH. The comparison of primary sequences and secondary structure propensities of the N-terminal parts of HLTF, SHPRH and Rad5 revealed that the SHPRH N-terminus also includes a tentative 'HIRAN-like' domain, which potentially plays a role in DNA substrate recognition similar to that of HIRAN domains of HLTF and Rad5.

Acknowledgments This work was supported by Connecticut Department of Public Health Biomedical Research Grant (DPH-UCHC BIOMED 2013-0203) and Connecticut Innovation Research Grant (13-SCA-UCHC-03). Research in the Bezsonova and Korzhnev labs is supported by NSF MCB (1616184, IB; 1615866, DK). The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Bayer, Boehringer Ingelheim, Genome Canada through the Ontario Genomics Institute [OGI-055], GlaxoSmithKline, Janssen, Lilly Canada, the Novartis Research Foundation, the Ontario Ministry of Economic Development and Innovation, Pfizer, Takeda, and the Wellcome Trust [092809/Z/10/Z].

References

Achar YJ, Balogh D, Neculai D, Juhasz S, Morocz M, Gali H, Dhe-Paganon S, Venclovas C, Haracska L (2015) Human HLTF mediates postreplication repair by its HIRAN domain-dependent replication fork remodelling. Nucleic Acids Res 43(21):10277–10291

Agrawal V, Kishan KV (2003) OB-fold: growing bigger with functional consistency. Curr Protein Pept Sci 4(3):195–206

Arcus V (2002) OB-fold domains: a snapshot of the evolution of sequence, structure and function. Curr Opin Struct Biol 12(6):794–801

Armon A, Graur D, Ben-Tal N (2001) ConSurf: an algorithmic tool for the identification of functional regions in proteins by surface mapping of phylogenetic information. J Mol Biol 307(1):447–463

Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res 38(Web Server issue):W529–W533

Blastyak A, Pinter L, Unk I, Prakash L, Prakash S, Haracska L (2007) Yeast Rad5 protein required for postreplication repair has a DNA helicase activity specific for replication fork regression. Mol Cell 28(1):167–175

Blastyak A, Hajdu I, Unk I, Haracska L (2010) Role of double-stranded DNA translocase activity of human HLTF in replication of damaged DNA. Mol Cell Biol 30(3):684–693

Branzei D, Psakhye I (2016) DNA damage tolerance. Curr Opin Cell Biol 40:137–144

Brunger AT (2007) Version 1.2 of the crystallography and NMR system. Nat Protoc 2(11):2728–2733

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr Sect D Biol Crystallogr 54:905–921

Callegari AJ, Kelly TJ (2016) Coordination of DNA damage tolerance mechanisms with cell cycle progression in fission yeast. Cell Cycle 15(2):261–273

Chang DJ, Cimprich KA (2009) DNA damage tolerance: when it's OK to make mistakes. Nat Chem Biol 5(2):82–90

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPIPE—a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6(3):277–293

Ding H, Descheemaeker K, Marynen P, Nelles L, Carvalho T, Carmo-Fonseca M, Collen D, Belayew A (1996) Characterization of a helicase-like transcription factor involved in the expression of the human plasminogen activator inhibitor-1 gene. DNA Cell Biol 15(6):429–442

Ding H, Benotmane AM, Suske G, Collen D, Belayew A (1999) Functional interactions between Sp1 or Sp3 and the helicase-like transcription factor mediate basal expression from the human plasminogen activator inhibitor-1 gene. J Biol Chem 274(28):19573–19580

Flaus A, Martin DMA, Barton GJ, Owen-Hughes T (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. Nucleic Acids Res 34(10):2887–2905

Goddard TD, Kneller DG, SPARKY 3. University of California, San Francisco

Goldenberg O, Erez E, Nimrod G, Ben-Tal N (2009) The ConSurf-DB: pre-calculated evolutionary conservation profiles of protein structures. Nucleic Acids Res 37(Database issue):D323–D327

Guntert P (2004) Automated NMR structure calculation with CYANA. Methods Mol Biol 278:353–378

Guntert P, Buchner L (2015) Combined automated NOE assignment and structure calculation with CYANA. J Biomol NMR 62(4):453–471

Hishiki A, Hara K, Ikegaya Y, Yokoyama H, Shimizu T, Sato M, Hashimoto H (2015) Structure of a novel DNA-binding domain of helicase-like transcription factor (HLTF) and its functional implication in DNA damage tolerance. J Biol Chem 290(21):13215–13223

Hoege C, Pfander B, Moldovan GL, Pyrowolakis G, Jentsch S (2002) RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature 419(6903):135–141

Huang YJ, Powers R, Montelione GT (2005) Protein NMR recall, precision, and F-measure scores (RPF scores): structure quality assessment measures based on information retrieval statistics. J Am Chem Soc 127(6):1665–1674

Huang YJ, Tejero R, Powers R, Montelione GT (2006) A topology-constrained distance network algorithm for protein structure determination from NOESY data. Proteins 62(3):587–603

Iyer LM, Babu MM, Aravind L (2006) The HIRAN domain and recruitment of chromatin remodeling and repair activities to damaged DNA. Cell Cycle 5(7):775–782

Kallberg M, Margaryan G, Wang S, Ma J, Xu J (2014) RaptorX server: a resource for template-based protein structure modeling. Methods Mol Biol 1137:17–27

Kanelis V, Forman-Kay JD, Kay LE (2001) Multidimensional NMR methods for protein structure determination. IUBMB Life 52(6):291–302

Kay LE (1995) Pulsed field gradient multi-dimensional NMR methods for the study of protein structure and dynamics in solution. Prog Biophys Mol Biol 63(3):277–299

Kay LE (1997) NMR methods for the study of protein structure and dynamics. Biochem Cell Biol 75(1):1–15

Kerr ID, Wadsworth RI, Cubeddu L, Blankenfeldt W, Naismith JH, White MF (2003) Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of Sulfolobus SSB protein. EMBO J 22(11):2561–2570

Kile AC, Chavez DA, Bacal J, Eldirany S, Korzhnev DM, Bezsonova I, Eichman BF, Cimprich KA (2015) HLTF's ancient HIRAN domain binds 3′DNA ends to drive replication fork reversal. Mol Cell 58(6):1090–1100

Letunic I, Bork P (2007) Interactive tree of life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics 23(1):127–128

Letunic I, Bork P (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Res 39(Web Server issue):W475–W478

Letunic I, Copley RR, Pils B, Pinkert S, Schultz J, Bork P (2006) SMART 5: domains in the context of genomes and networks. Nucleic Acids Res 34(Database issue):D257–D260

Letunic I, Doerks T, Bork P (2015) SMART: recent updates, new developments and status in 2015. Nucleic Acids Res 43(1):D257–D260

Machado LE, Pustovalova Y, Kile AC, Pozhidaeva A, Cimprich KA, Almeida FC, Bezsonova I, Korzhnev DM (2013) PHD domain from human SHPRH. J Biomol NMR 56(4):393–399

Mahajan MC, Weissman SM (2002) DNA-dependent adenosine triphosphatase (helicaselike transcription factor) activates beta-globin transcription in K562 cells. Blood 99(1):348–356

Motegi A, Sood R, Moinova H, Markowitz SD, Liu PP, Myung K (2006) Human SHPRH suppresses genomic instability through proliferating cell nuclear antigen polyubiquitination. J Cell Biol 175(5):703–708

Motegi A, Liaw H-J, Lee K-Y, Roest HP, Maas A, Wu X, Moinova H, Markowitz SD, Ding H, Hoeijmakers JHJ, Myung K (2008) Polyubiquitination of proliferating cell nuclear antigen by HLTF and SHPRH prevents genomic instability from stalled replication forks. Proc Natl Acad Sci USA 105(34):12411–12416

Murzin AG (1993) OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. EMBO J 12(3):861–867

Sale JE, Lehmann AR, Woodgate R (2012) Y-family DNA polymerases and their role in tolerance of cellular DNA damage. Nat Rev Mol Cell Biol 13(3):141–152

Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci USA 95(11):5857–5864

Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J Biomol NMR 44(4):213–223

Sheridan PL, Schorpp M, Voz ML, Jones KA (1995) Cloning of an SNF2/SWI2-related protein that binds specifically to the SPH motifs of the SV40 enhancer and to the HIV-1 promoter. J Biol Chem 270(9):4575–4587

Sievers F, Higgins DG (2014) Clustal Omega, accurate alignment of very large numbers of sequences. Methods Mol Biol 1079:105–116

Theobald DL, Mitton-Fry RM, Wuttke DS (2003) Nucleic acid recognition by OB-fold proteins. Annu Rev Biophys Biomol Struct 32:115–133

Thoma NH, Czyzewski BK, Alexeev AA, Mazin AV, Kowalczykowski SC, Pavletich NP (2005) Structure of the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54. Nat Struct Mol Biol 12(4):350–356

Unk I, Hajdu I, Fatyol K, Szakal B, Blastyak A, Bermudez V, Hurwitz J, Prakash L, Prakash S, Haracska L (2006) Human SHPRH is a ubiquitin ligase for Mms2-Ubc13-dependent polyubiquitylation of proliferating cell nuclear antigen. Proc Natl Acad Sci USA 103(48):18107–18112

Unk I, Hajdu I, Fatyol K, Hurwitz J, Yoon J-H, Prakash L, Prakash S, Haracska L (2008) Human HLTF functions as a ubiquitin ligase

for proliferating cell nuclear antigen polyubiquitination. Proc Natl Acad Sci USA 105(10):3768–3773

Unk I, Hajdu I, Blastyak A, Haracska L (2010) Role of yeast Rad5 and its human orthologs, HLTF and SHPRH in DNA damage tolerance. DNA Repair 9(3):257–267

Wang Z, Zhao F, Peng J, Xu J (2011) Protein 8-class secondary structure prediction using conditional neural fields. Proteomics 11(19):3786–3792

Waters LS, Minesinger BK, Wiltrout ME, D'Souza S, Woodruff RV, Walker GC (2009) Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. Microbiol Mol Biol Rev 73(1):134–154