Università degli Studi di Trieste

Graduate School in MOLECULAR BIOMEDICINE

PhD Thesis

Structural and Biochemical study of human RecQ4

ADITYA MOJUMDAR

… ciclo – Anno Accademico 2013/2014 …

UNIVERSITÀ' DEGLI STUDI DI TRIESTE

XXVII CICLO DEL DOTTORATO DI RICERCA IN BIOMEDICINA MOLECOLARE

Structural and Biochemical study of human RECQ4

Settore scientifico-disciplinare: BIO/11 Biologia Molecolare

DOTTORANDO / A
ADITYA MOJUMDAR

COORDINATORE
Prof. Guidalberto Manfioletti

SUPERVISORE DI TESI
Dr.ssa Silvia Onesti

ANNO ACCADEMICO 2013 / 2014

To my Parents

মা ও বাবা

Abstract

RecQ helicases belong to a ubiquitous family of DNA unwinding enzymes that are essential to maintain genome stability by acting at the interface between DNA replication, recombination and repair. Humans have five different paralogues of RecQ helicases namely RecQ1, BLM, WRN, RecQ4 and RecQ5. This work focuses on the structural and biochemical study of human RecQ4. Germ-line mutations in the *RECQ4* gene give rise to three distinct human genetic disorders (Rothmund-Thomson, RAPADILINO and Baller-Gerold syndromes). Despite the important roles of RecQ4 in various cellular processes, RecQ4 have never been fully characterized.

In addition to the helicase domain, RecQ4 has a unique N-terminal part that is essential for viability and is constituted by a region homologous to the yeast Sld2 replication initiation factor, followed by a cysteine-rich region, predicted to fold as a Zn knuckle. A part of this work focuses on the structural and biochemical analysis of both the human and *Xenopus* RecQ4 cysteine-rich regions, and shows by NMR spectroscopy that the *Xenopus* fragment does indeed assumes the canonical Zn knuckle fold, whereas the human sequence remains unstructured, consistent with the mutation of one of the Zn ligands. Both the human and *Xenopus* Zn knuckles bind to a variety of nucleic acid substrates, with a preference for RNA. We also investigated the effect of an additional Sld2 homologous region upstream the Zn knuckle. In both the human and *Xenopus* system, the presence of this region strongly enhances binding to nucleic acids. These results reveal novel possible roles of RecQ4 in DNA replication and genome stability.

Recently the catalytic core of RecQ4 has been predicted to include RecQ-like-C-terminal (RQC) domain at the C-terminus of the helicase domain, similar to other RecQ helicases. This domain is composed of a Zn-binding region and a winged helix (WH) domain. Another part of this thesis centers on the structural and biochemical characterization of the catalytic core of RecQ4 including the helicase and RQC domain. The results provide an insight in the Zn binding ligands present in the RQC domain that plays a role in DNA binding and unwinding activity of the protein. Also the presence of the characteristic aromatic residue at the tip of the WH β-hairpin and its role in DNA binding and unwinding has been established. Finally, it provides a low resolution SAXS model of the catalytic core of RecQ4.

Acknowledgements

I want to thank my supervisor Dr. Silvia Onesti for giving me the opportunity to work in her laboratory in the Sincrotrone Trieste Elettra. Thank you for the scientifical advices, for encouraging me in pursuing my research and for your support throughout these years.

I have to thank present and past members of the Structural Biology (Elettra) group, especially Dr. Francesca Marino, Dr. Matteo De March and Dr. Barbara Medagli.

I need to thank also our collaborator Dr. Alessandro Vindigni, Dr. Emanuelle Buratti and Dr. Riccardo Sgarra for giving me the valuable suggestions and lab space in ICGEB to complete this piece of work.

I thank the DSV University of Trieste for enrolling me in the PhD programme of Molecular Biomedicine and funding me all these years.

I dedicate this thesis to my mother, my father, younger sister and thank all my friends especially Vishal, Jashmini, Marta, Valentina and Maryse for providing me the necessary support during all these years.

Contents

List of Tables

List of Figures

1. Introduction

1.1. An introduction to DNA helicases

1.1.1. Properties of nucleic acid helicases

Helicases are molecular motors that couple the energy of nucleoside triphosphate (NTP) hydrolysis to the unwinding and remodeling of double-stranded (ds) DNA and RNA (Singleton, 2007; Spies, 2014). Therefore they play a role in almost every process in cells that involves nucleic acids, including DNA replication and repair, transcription, translation, ribosome synthesis, RNA maturation, splicing and nuclear export processes. A large number of genes have been classified as helicases, especially in eukaryotes (estimates range from 1 to 2% of the entire genome); although some of the proteins including so-called "helicase motifs" may represent DNA/RNA translocases, helicases can still be considered very abundant proteins.

From a biochemical point of view, helicases can be characterized by their rate, directionality, processivity, step size and the mechanism of unwinding that can be active or passive. The **rate** of translocation can vary from a few to several thousand base pairs per second, and is often determined via the ATPase rate. Some helicases are less active until they bind a protein cofactor that highly stimulates their activity. The **directionality** is linked to the fact that DNA is a bipolar molecule with two strands running in opposite direction. Even when these enzymes bind a double stranded DNA, they travel along just one of the two strands, either in the 3' to 5' or in the 5' to 3' direction. Thus, the way the protein is loaded on the substrate is critical for the direction of movement of the enzyme. The **processivity** of helicases reflects the ability of these enzymes to catalyze multiple cycles of reaction before releasing the product, as opposed to a distributive enzyme. The **step size** is the number of base pairs unwound per molecule of ATP hydrolyzed. Finally, **active** helicases participate in the active unwinding of the duplex ahead of the fork, while **passive** helicases wait for thermal fraying of the strands at the fork and then they trap the strands in the unwound state (Lohman and Bjornson, 1996; Singleton et al., 2007).

1.1.2. Classification

The first classification framework for structure-function analysis was provided by Gorbalenya and Koonin in 1993. They pointed out that subsets of distantly related helicase proteins share

short, conserved amino acid sequence motifs and proposed a general classification of helicases into two major groups or superfamilies (SF1 and SF2) with three additional smaller families, differentiated both by the number of distinct motifs that are identified within each group and by differences in the consensus sequences for motifs that are shared by more than one group (Gorbalenya and Koonin, 1993). The original classification was done on seven conserved motifs but it has become clear now that these motifs are not all comparable between the different families. Moreover, new motifs have been identified, characteristic of a single family, leading to an updated classification comprising 6 superfamilies (Singleton et al., 2007)

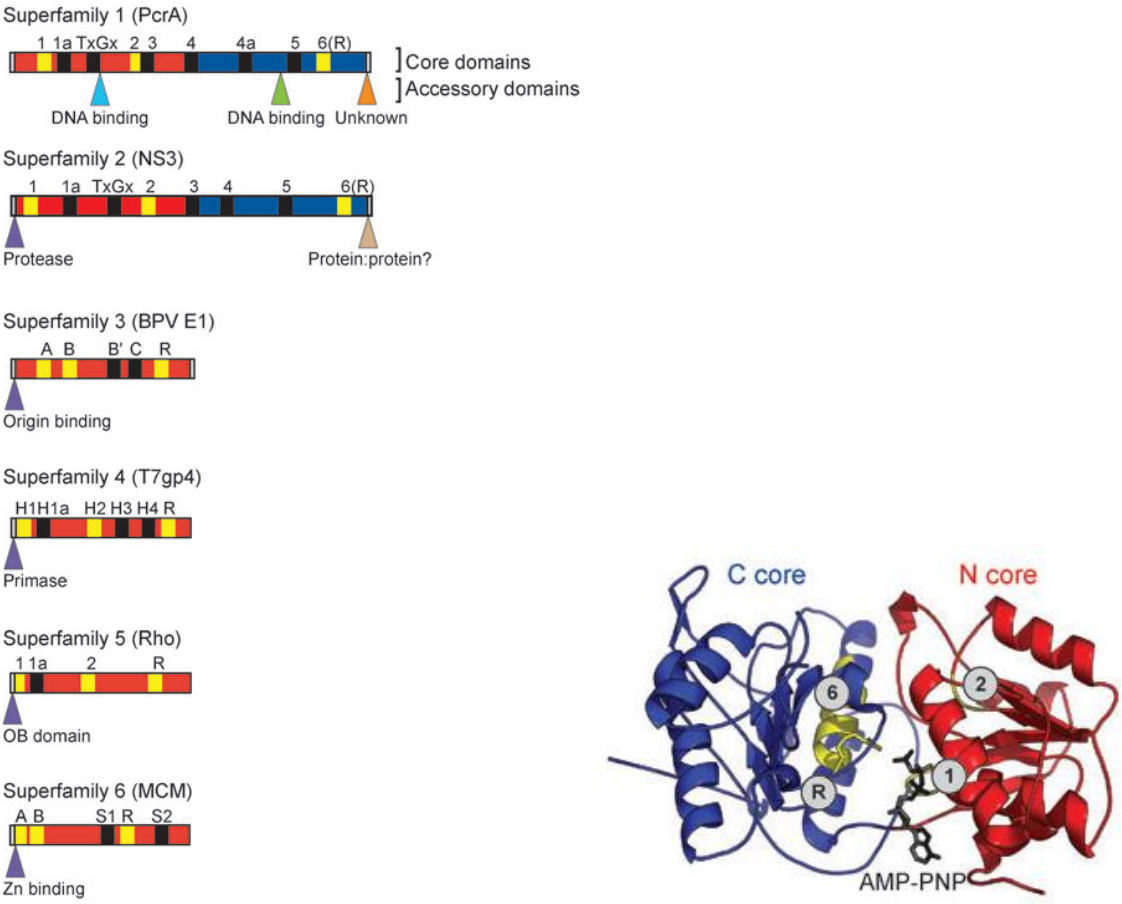

Figure 1.1. The helicase family (left) Helicases classification into 6 superfamilies according to their primary sequence. The name of a representative for each family is shown. Universal conserved motifs are colored in yellow. (right) The SF1 and SF2 core helicase is composed by two RecA-like domains, colored blue and red. In a cleft between the two domains binds the ATP analog AMP-PNP (black). Motif 1 and 2 are related to Walker A and B motifs present at the N-core domain and motif 6 containing the arginine finger residue is present at the C-core domain. The conformational change in the structure of the enzyme

due to coupling of ATP binding and hydrolysis is highlighted in yellow. Pictures adapted from Singleton et al., 2007.

All the 6 superfamilies contain the characteristic signature motifs for "core domains" that form either a RecA or AAA+-like ATPase core with the nucleotide-binding sites at the interface of the monomers (Figure 1.1). This structural unit converts the chemical energy to mechanical energy by coupling NTP binding and hydrolysis to protein conformational changes. Common features of the core domains include conserved residues involved in the NTP binding and hydrolysis equivalent to the Walker A and B boxes of many ATPases (Walker et al., 1982), and an 'arginine finger' that plays a key role in energy coupling (Scheffzek et al., 1997).

According to their direction of translocation they can be distinguished as type A or type B helicases if they move on the substrate from 3' to 5' or 5' to 3' direction, respectively. A final subdivision is drawn on the basis of the substrate on which these enzymes translocate, named α or β helicases if the substrate is single or double stranded substrate, respectively.

It is fair to mention that these last classifications can be limited by a lack of biochemical characterization. Moreover, complexity in these enzymes is given by the diverse range of accessory domains that add to the core domain either as N- or C-terminal flanking domains, conferring each helicase peculiar functions in cellular pathways. A diverse range of nucleic acid processing machinery has developed during evolution by mixing and matching these core and accessory domains.

1.1.3. Helicases superfamilies: an overview

SF1 and **SF2** are the largest and structurally best characterized classes. The hallmark of SF1 and SF2 helicases is a conserved helicase core consisting of two RecA like domains coupled by a short linker. The helicase core contains seven characteristic motifs (I, Ia and II-VI) (Singleton and Wigley, 2002). In both superfamilies motifs I and II are the highly conserved Walker A and B sequences characteristic of ATPases (Walker et al., 1982). The SF1 and SF2 helicases comprise defined protein families with distinct sequence, structural and mechanistic features. To date, SF1 and SF2 contain examples of both type A and B, most SF1 enzymes appear to be α, whereas SF2 has examples of both α and β type.

From a structural point of view the best characterized examples are, for a SF1A helicase, PcrA (Velankar et al., 1999) helicase from gram-positive bacteria, Rep (Korolev et al., 1997) and UvrD (Lee and Yang, 2006) helicase in gram-negative bacteria; for a SF1B helicase RecD2 from *Deinococcus Radiodurans.* A large body of biochemical and structural data has revealed the mechanism of action of these two enzymes, providing a good starting model for other helicases.

SF2 enzymes act on a wide range of nucleic acid substrates, so that it is more proper to rename them as nucleic acid remodeling proteins rather than just helicases. For example, RNA remodeling is usually catalyzed by a subfamily of helicases containing the DExH/D conserved box in motif II, belonging to the SF2 group. These factors are implicated in translation initiation, ribosome biogenesis, RNA splicing, microRNA function, RNA transport, replication of RNA viruses and a lot more. The well characterized DNA remodeling enzymes fall into two families: the RecQ family of DNA helicases and the Snf2 family of chromatin-remodeling proteins. The structural information for SF2 members is less complete compared to the one for SF1 members. Moreover, considering the great variety of functions in which SF2 members are implicated, it is probable that they differentiate in their mechanisms for translocation and unwinding. The best studied examples are the HEL308 (Buttner et al., 2007), flavivirus NS3 helicase (Luo et al., 2008) and the human RNA helicase DDX19 (both containing the DExH box) (Collins et al., 2009).

Whereas enzymes belonging to SF1 or SF2 work as monomers or dimers and contains tandem ASCE (RecA-like) units, proteins belonging to SF3-6 form toroidal rings, with each monomer providing a single ASCE (RecA-like or AAA+) unit.

SF3 helicases were initially identified in the genomes of small DNA and RNA viruses (Gorbalenya et al., 1990) and are associated with multiple enzymatic activities, such as origin recognition and unwinding (Hickman and Dyda, 2005). They form hexamers or double hexamers and have a 3’ to 5’ translocation directionality. SF-3 proteins share the topology and the conserved motifs of the a/b domain of the AAA+ proteins. The best characterized structural example is the hexamer formed by the papilloma virus E1 helicase bound to a 13-mer DNA and ADP (Enemark and Joshua-Tor, 2006).

Members of **SF4** were first identified in bacteria and bacteriophages and act as replicative helicases. In bacteria they associate with a primase as separate polypeptides whereas in some

bacteriophage systems both activities are fused in a single chain. They are known to have 5' to 3' polarity. Five sequence motifs are conserved in SF4 proteins, two of them are the Walker A and B boxes but the others do not have a counterpart in any other helicase. The best studied example is the gene 4 protein from bacteriophage T7 (gp4) (Singleton et al., 2000; Toth et al., 2003).

Rho helicase represents **SF5** family, and is responsible of termination of transcription in bacteria. It binds a specific RNA sequence and then unwinds the DNA/RNA hybrid. Its hexameric structure has been studied by x-ray crystallography (Skordalakes and Berger, 2003), (Skordalakes and Berger, 2006) and electron microscopy (Gogol et al., 1991). A more recent structure with ssRNA bound to the central channel (Thomson and Berger, 2009), reveals intriguing similarities with the SF3 mechanism of action and explains how the proteins utilizes identical conformational changes to run in the opposite direction (5' to 3').

The **SF6** group includes hexameric motor proteins built around a AAA+ core, such as the main eukaryotic replicative helicase MCM (Maiorano et al., 2006) (mini chromosome maintenance) and the prokaryotic RuvB protein which together with RuvA and RuvC is able to process Holliday junctions, an intermediate of homologous recombination (West SC, 1996). No detailed structural information for an active hexamer is available (Costa and Onesti, 2009).

1.1.4. The SF2 superfamily.

The SF2 enzymes are the largest superfamily and are implicated in diverse range of cellular processes. The intensively studied subfamilies include the RecQ-like family (Bennett and Keck, 2004), the DEAD-box RNA helicases (Cordin et al., 2006) and the Snf2-like chromatin remodeling enzymes (Flaus and Owen-Hughes, 2004; Flaus et al., 2006). Besides the minimal biochemical function of the SF2 catalytic core of ATP-dependent directional translocation on single or double stranded nucleic acid some members may play a role in remodeling nucleic acid and nucleoprotein conformations in an NTP-dependent manner (Linder and Lasko, 2006).

The crystal structure of SF2 HEL308 (Buttner et al., 2007), in complex with a dsDNA carrying a 3' single stranded tail, has revealed important features regarding the initiation of duplex unwinding reaction, such as the presence of a prominent hairpin loop involved in strand separation. Other structures of SF2 members supporting inchworm mechanism are available: the flavivirus NS3 helicase (Luo et al., 2008) and the human RNA helicase DDX19 (both containing

the DExH box) (Collins et al., 2009) or the DEAD-box Vasa from Drosophila (Sengoku et al., 2006) and the DEAD-box ATPase eukaryotic initiation factor 4AIII (Andersen et al., 2006). These structures crystallized either in their apo-forms or in complex with different nucleotides or nucleic substrates.

A rare example of complete structural cycle with many snapshots capturing a SF2 enzyme during ATP binding and hydrolysis coupled to translocation along nucleic acid is the NS3 helicase from Hepatitis C virus (HCV) (Gu and Rice, 2010). Three structures have been crystallized for NS3 HCV helicase: the binary complex enzyme-ssDNA, the tertiary complex enzyme-ssDNA-ADP.BeF3, mimicking the ATP bound state, and the tertiary complex enzyme-ssDNA-ADP.AlF4, mimicking the transition state of ATP hydrolysis.

The binding of the ATP analogues induces large conformational rearrangements in the enzyme and closes the cleft in the core domain, similarly to other ATPases composed of RecA-like domains. Motif II has a key role in changes upon nucleotide binding and hydrolysis. The nucleotide driven conformational changes are reflected also on the bound DNA, which bends and rotates altering the interactions between the helicase and its substrate. A cyclic loss and recovery of these interactions in the different nucleotide bound states, allow the enzyme to move along DNA with a 'ratchet' translocation mechanism. NS3h is therefore able to maintain a unidirectional movement defining a step size of one base pair per ATP hydrolysis cycle.

Some SF2 helicases display an enhanced helicase activity when multiple enzymes are loaded onto the single stranded nucleic acid flanking the double stranded regions that are yet to be unwound (Levin et al., 2004; Byrd and Raney, 2005; Tackett et al., 2005). In these so-called 'functional cooperative' models, the presence of multiple enzymes is needed for preventing backward, non-productive movements or to replace the lead helicase after dissociation from nucleic acid by a trailing helicase.

1.2. RecQ helicases

1.2.1. The RecQ family

RecQ helicases belong to a ubiquitous family of DNA unwinding enzymes that are essential to maintain the genome stability by acting at the interface between DNA replication, recombination and repair. The family was named after the *RecQ* gene of *Escherichia coli*, which is the only

member of the family present in this organism (Nakayama, 1984). RecQ helicases belong to superfamily SF2 and have 3' to 5' directionality in unwinding a variety of DNA structures including B-form DNA duplexes, displacement loops, forked duplexes, DNA junctions and G-quadruplex DNA. They can also promote annealing of complementary single-stranded DNA and branch migration of Holiday junctions. As a consequence they are associated with various cellular functions like stabilization and repair of damaged DNA replication fork, DNA damage checkpoint signalling, telomere maintenance, base excision repair and homologous recombination, thus also known as guardians of the genome (Bohr, 2008; Larsen and Hickson, 2013), although their exact role in the cell and mechanism of action are still poorly understood.

Most bacteria have a single RecQ helicase, of which the well-studied *E. coli* protein is the prototype. Only one RecQ helicase, Sgs1, was known to be present in yeasts until a bioinformatics analysis suggested putative orthologues in both plants and fungi, called Hrq1, which was suggested to be equivalent to RecQ4 (Gangloff et al., 1994; Watt et al., 1995; Barea et al., 2008). However the link between Hrq1and RecQ proteins is partially weakened by a detailed analysis of the sequence, as Hrq1 proteins do not appear to include the ubiquitous RecQ motif '0' (see secion 1.2.4), nor an obvious homologue of the RQC domain (see section 1.2.5). In higher organisms the gene seem to have undergone multiple gene duplication with most vertebrate possessing up to five versions of RecQ proteins, with distinct but partly overlapping functions (Hickson, 2003).

1.2.2. Human RecQ helicases and their role in disease

Human cells have five different RecQ paralogues namely RecQ1, BLM, WRN, RecQ4 and RecQ5 (Figure 1.2). They share a similarity in their helicase domain and some biochemical properties but also have some additional domains and unique features. Genetic mutations in three paralogues (BLM, WRN and RecQ4) result in rare genetic disorders namely **Bloom's syndrome**, **Werner's syndrome** and **Rothmund-Thomson, RAPADILINO and Baller-Gerold syndrome** (Ellis et al., 1995; Yu et al., 1996; Kitao et al., 1999; Siitonen et al., 2003; Van Maldergem et al., 2006). In addition to the common characteristics of genetic instability, developmental defects, skin related problems and cellular senescence, the disorders have specific clinical symptoms unique to each mutation. This suggests that the five RecQ helicases have distinct, although

partially overlapping, roles within the cells. Although disease phenotypes have not been genetically linked to mutations in *RECQ1* and *RECQ5* genes yet, they may be responsible for additional cancer predisposition disorders that are distinct from other RecQ syndromes.

1.2.3. Overall architecture of RecQ helicases.

RecQ enzymes have three conserved domains that are commonly found in most of the members: the core helicase domain, the RecQ C-terminal (RQC) domain, and the helicase and RnaseD-like C-terminal (HRDC) domain (Figure 1.2) (Vindigni et al., 2010). While the helicase domain is conserved throughout the RecQ family, RQC and HRDC domains are missing in some family members. Previously, it was thought that RecQ4 lacked the RQC domain, but a recent bioinformatics analysis done in our lab suggests that it has an RQC domain (Marino et al., 2013). In addition, RecQ enzymes differ greatly in the length of their N- and C-terminal domains that are flanking the catalytic core domain. These flanking domains are involved in different protein interactions, regulation of protein subcellular localization, promotion of enzyme oligomerization and in giving additional enzymatic activities, such as the N-terminal domain exonuclease activity of Werner syndrome protein (Shen et al., 1998).

The crystal structures of a number of RecQ helicases or RecQ fragments have been determined and, together with biochemical analysis, have contributed to the understanding of the role of the single domains.

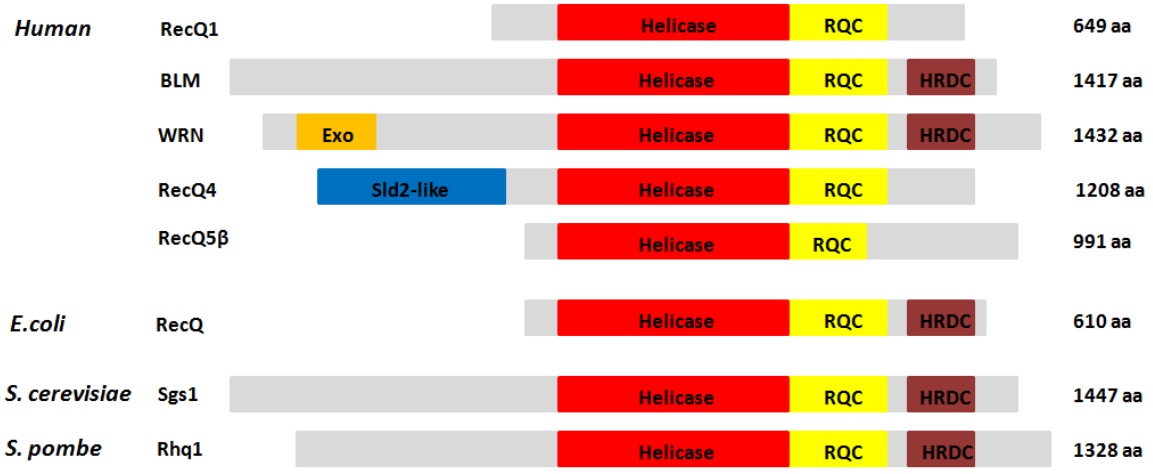

Figure 1.2. RecQ helicases. Schematic representation of RecQ family members from human, *E. coli* and yeast: conserved and additional domains are indicated. In general, less complex organisms have just one family member.

1.2.4. The core helicase domain

As members of SF2 family, the RecQ helicases are defined by the presence of seven highly conserved amino acid sequence motifs (I, Ia, II–VI) in the core helicase domain, which is generally centrally located in all RecQ family members (Figure 1.2). This motif has a size of approximately 400 amino acids, and is required for NTP binding and for coupling of the energy derived from the NTP hydrolysis to the process of nucleic acid unwinding (Gorbalenya and Koonin, 1993; Singleton et al., 2007; Lohman et al., 2008). As for various members of SF1 and SF2 family, crystal structures of RecQ helicases (Bernstein et al., 2003; Pike et al., 2009) have shown that the core helicase motif is formed by two RecA-like domains that are acting as molecular motors of the Helicase (Figure 1.3).

In particular, **Motif I**, also called **Walker A** motif is typically defined by a Gx4GK(S/T) consensus sequence and is essential for ATP binding: a conserved lysine residue makes contacts with the β and γ phosphate of the ATP molecule. The S/T residue plays a role in coordination of divalent metal ion Mg^{2+} which is important for hydrolysis of ATP (Walker et al., 1982; Morozov et al., 1997). This motif differs slightly in RecQ helicases (TGxGKS), but retains the canonical function. The highly conserved aspartate residue in **Motif II or Walker B** with DExHC sequence coordinates the Mg^{2+} ion, while the glutamate acts as a catalytic base (Pause and Sonenberg, 1992; Gorbalenya and Koonin, 1993; Morozov et al., 1997; Bernstein et al., 2003; Pike et al., 2009).

Furthermore, an additional **motif '0'** has been first identified in the *E. coli* RecQ helicase. This sequence element is located N-terminally from motif I and contains a conserved sequence '$LX_3(F/Y/W)GX_3F(R/K)X_2Q$' (Bernstein and Keck, 2003). The crystal structure of the nucleotide bound form of *E. coli* RecQ shows that the adenine moiety of ATPγS is located between the conserved aromatic residue (Tyr23) and Arg27 and is hydrogen-bonded to the conserved glutamine (Gln30). Mutations in this domain show the involvement of motif '0' in nucleotide binding, suggesting that it plays an important function in the core helicase domain. For example, mutation which changes Gln residue to Arg residue in motif '0' of the human *BLM* gene is enough to cause Bloom's syndrome (Ellis et al., 1995). Moreover, budding yeast cells

carrying this mutation in the *Sgs1* gene show the same DNA-damaging agent sensitivity as that of a *Sgs1* deletion strain (Onoda et al., 2000).

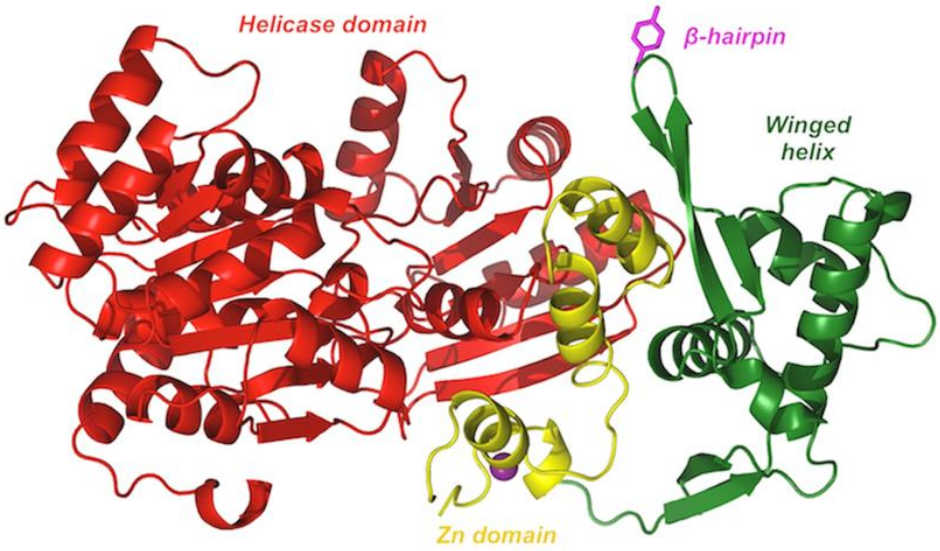

Figure 1.3. The catalytic core. The crystal structure of the catalytic core of human RecQ1 (PDB code 2V1X) including the Helicase domain and the RQC domain. The Helicase domain comprises of two RecA-like folds in red, the RQC domain encompasses the Zn domain in yellow and the winged helix domain in green with the aromatic residue acting as a pin shown in magenta.

1.2.5. The RecQ C-terminal (RQC) domain

This second most conserved domain is unique to the RecQ family of helicases and consists of two domains, a Zn^{2+}-binding domain and Winged-Helix (WH) domain (Figure 1.3). Up to date, the structures of RQC domains have been solved for *E. coli* RecQ, human RECQ1, WRN and BLM helicases (Bernstein et al., 2003; Pike et al., 2009; Kitano et al., 2010; Kim et al., 2013). The structures of the Zn^{2+}-binding modules are highly similar between the bacterial and human enzymes, where a single Zn^{2+} ion is coordinated by four Cys residues located on two anti-parallel α-helices. Considering that the Zn^{2+}-binding site is conserved among RecQ proteins, it is not surprising that functional analyses have shown this region to be important for RecQ function. In particular, it has been demonstrated that mutation at the zinc domain in *E. coli* RecQ helicase abrogates DNA binding of this enzyme and leads to decreased ATPase and helicases activities. In their studies, Liu and colleagues have found this motif to be crucial for the integrity of the whole protein (Liu et al., 2004). Moreover, two disease-causing BLM missense mutations mapping to the highly conserved cystein residues, Cys-1036 and Cys-1055, in the Zn^{2+}-binding domain have been identified (Ellis et al., 1995). Using different biochemical approaches, Guo

and colleagues proposed a role of this motif in DNA binding, protein stability and protein folding of BLM helicase (Guo et al., 2005). Consistently, when equivalent missense mutations found in Bloom's syndrome patients were introduced in the budding yeast BLM homologue, *Sgs1* gene, these mutants could not suppress DNA-damage sensitivity and increased frequency of interchromosomal recombination and sister chromatid exchange (Onoda et al., 2000). Another study on the human RecQ5α and RecQ5β isoforms from Ren and colleagues proposed the RecQ-specific Zn^{2+}-binding motif as an essential DNA-binding module required for RecQ family helicase activity (Ren et al., 2008).

The second RQC-terminal domain folds into a winged helix (WH) (Figure 1.3), an important motif in mammalian proteins that mediates protein-DNA interactions, as well as protein-protein interactions (Clark et al., 1993; Fogh et al., 1994; Littlefield and Nelson, 1999; Gajiwala et al., 2000; Mer et al., 2000). For the RecQ family of helicases, several studies have demonstrated that the WH domain is implicated in dsDNA recognition. Biochemical studies on *E.coli* RecQ and BLM helicases showed that the RQC domain is required for G-quadruplex DNA binding, and similar studies with a WRN fragment comprising the RQC domain demonstrated that this domain is required for the interaction of WRN with Holliday junctions and forked duplex substrates (von Kobbe et al., 2003; Huber et al., 2006). For the WRN protein, Lee and coworkers demonstrated that Lys-1016 residue located in the WH domain is important for proper fork, D-loop, and Holliday junction substrates unwinding and that mutation of this residue significantly reduces the ability of WRN to stimulate flap structure-specific endonuclease 1 (FEN-1) incision activities (Lee et al., 2005). Moreover, this domain has been shown to bind several proteins involved in DNA metabolism, such as human FEN-1, poly(ADP-ribose)polymerase-1 (PARP-1) and telomere repeat binding factor 2 (TRF2) (Opresko et al., 2002; Bachrati and Hickson, 2003; Lee et al., 2005; von Kobbe et al., 2004). Shereda and colleagues have shown for the prototypical *E. coli* RecQ protein that complex formation between RecQ and single-strand binding protein (SSB) is mediated by the RecQ winged-helix domain. In particular, by using nuclear magnetic resonance and mutational analyses, they have identified the SSB-Ct binding pocket on *E.coli* RecQ (Shereda et al., 2009).

The family of RecQ helicases is characterized by a poor degree of primary sequence similarity in their WH domains. Interestingly, the crystal structures of the WH domains of *E. coli* RecQ,

human RecQ1, WRN and BLM, as well as nuclear magnetic resonance (NMR) derived structure of this domain in WRN, show a very well conserved domain architecture (Bernstein et al., 2003; Hu et al., 2005; Pike et al., 2009; Kitano et al., 2010; Kim et al., 2013). Mutagenesis studies on RecQ1 have shown that prominent β-hairpin present in WH domain is essential for DNA unwinding, as the substitution of the Tyr residue at the tip of the loop was sufficient to abolish the unwinding activity of RecQ1 (Pike et al., 2009; Lucic et al., 2011). The corresponding β-hairpin in WRN is similarly capped by a phenylalanine at the same position, supporting a model in which the WH domain splits the DNA duplex using the β-wing as a wedge (Kitano et al., 2010). In contrast with the human RecQ1 and WRN proteins, the *E. coli* RecQ β-hairpin is significantly shorter and lacks corresponding aromatic residues at the tip, indicating that the β-wing is not as important for DNA unwinding in the *E. coli* RecQ enzyme (Pike et al., 2009; Kitano et al., 2010; Lucic et al., 2010). A recent study reports the absence of the aromatic residue at the tip of the β-hairpin in the RQC domain of BLM protein and presence of an insertion between the N-terminal helices exhibiting a looping-out structure that extends at right angles to the β-hairpin. Deletion mutagenesis of this insertion interfered with binding to Holliday junction (Kim *et al.* 2013). A subsequent study confirmed the presence of an asparagine instead of the aromatic residue at the tip of β-hairpin in RQC domain of BLM protein (Swan et al., 2014).

1.2.6. The helicase-and-RNaseD-like-C-terminal (HRDC) domain

The third conserved region of RecQ helicases, the helicase-and-RNaseD-like-C-terminal (HRDC), is present in most RecQ family members (Figure 1.2). This domain is also found in ribonucleases D where has been proposed to be involved in nucleic acid binding (Morozov et al., 1997). Structural and biochemical studies indicate that HRDC domains are the most poorly conserved of the three sequence elements present in RecQ family members. Moreover, some eukaryotic RecQ proteins, such as RecQ1, RecQ4 and all isoforms of RecQ5, lack HRDC domains, whereas several bacterial RecQ proteins have multiple HRDC domains. The first three dimensional structure of an HRDC domain was solved for *S. cervisiae* Sgs1 using heteronuclear multidimensional NMR spectroscopy. NMR studies of the isolated HRDC of Sgs1 suggested that this domain functions as an auxiliary DNA-binding domain, similar to the 1B module of SF1 helicases (Liu et al., 1999). Although all HRDC domains have a similar helical bundle structure,

they have different surface charge distributions and DNA-binding affinities (Vindigni et al., 2010). In particular, Liu and coworkers suggested that a patch of basic residues located in the HRDC of Sgs1 is involved in electrostatic interactions with the phosphate backbone of the DNA, and that this domain is able to bind both ssDNA and partial dsDNA with single stranded overhangs. Similar to the yeast Sgs1, the data from the bacterial RecQ HRDC domain reveal significant structural resemblance to auxiliary DNA binding domains in PcrA and Rep helicases. However, *E. coli* HRDC domain preferentially binds to ssDNA over dsDNA (Bernstein and Keck, 2005). A recent crystallographic study on another bacterial helicase from *Deinococcus radiodurans* showed that an unusual electrostatic surface of the third HRDC domain of DrRecQ (DrRecQ HRDC3) may be important for inter-domain interactions that regulate structure-specific DNA binding and help direct DrRecQ to specific recombination/repair sites (Killoran and Keck, 2008). Among the five human RecQ proteins, only WRN and BLM share the presence of an HRDC domain. In BLM, this domain is a critical determinant for the efficient binding and unwinding of double Holliday junctions (dHJ) (Wu et al., 2005). A study of the solution structure reveals unique features of BLM HRDC that are distinct from the HRDC domain of WRN protein. In particular, BLM HRDC domain contains a patch of acidic residues which makes the domain surface extensively electronegative, suggesting an additional role of this motif apart from DNA binding (Sato et al., 2010). Kim and coworkers have proposed an electrostatic repulsion model in which upon BLM oligomerization, the HRDC domain from each BLM monomer uses electrical repulsion to separate the junction sites of dHJ DNA. In their model, the helicase core domain would then bind and unwind the double stranded region of the dHJ DNA because of its higher binding affinity for dsDNA. Subsequently, the dHJ DNA is divided into non-crossover products (Kim and Choi, 2010). In contrast, the HRDC domain of WRN does not appear to interact with DNA *in vitro* (Liu et al., 1999; von Kobbe et al., 2003; Kitano et al., 2007). However, a WRN fragment containing the HRDC domain and additional residues at the C-terminus binds forked-duplex DNA and Holliday junctions with high affinity (von Kobbe et al., 2003). The crystallized HRDC domain from WRN revealed that this protein possesses an additional N- and C- terminal extension to the standard helical bundle, which is missing in other RecQ helicases. The authors have suggested that WRN HRDC may be adapted to play a distinct function that involves protein-protein interactions, rather than protein-DNA interactions (Kitano

et al., 2007). Consistent with these differences, the heterogeneity in the HRDC domains may correlate with functional differences in WRN and BLM.

1.3. Biochemical characterization of RecQ helicases

Although all members of RecQ helicase family unwind DNA with a 3' to 5' polarity, they diverge noticeably on their activity and substrate specificity; in particular, with respect to most other helicases, they act on a wide range of substrates, shown in figure 1.4, including a number of unusual DNA structures, including forked duplexes, displacement loops (D-loops; an intermediate in homologous recombination reactions), triple helices, 3- or 4-way junctions, and G-quadruplex DNA (Mohaghegh et al., 2001; Popuri et al., 2008; Sun et al., 1998). Beside their helicase activity, some RecQ helicases are also able to promote the single strand annealing reaction of complementary ssDNA strands in an ATP-independent fashion (Garcia et al., 2004; Cheok et al., 2005; Machwe et al., 2005; Sharma et al., 2005; Macris et al., 2006). Additionally, WRN possesses a 3' to 5' exonuclease activity that distinguishes it from the other human RecQ enzymes (Huang et al., 1998; Shen et al., 1998; Suzuki et al., 1999).

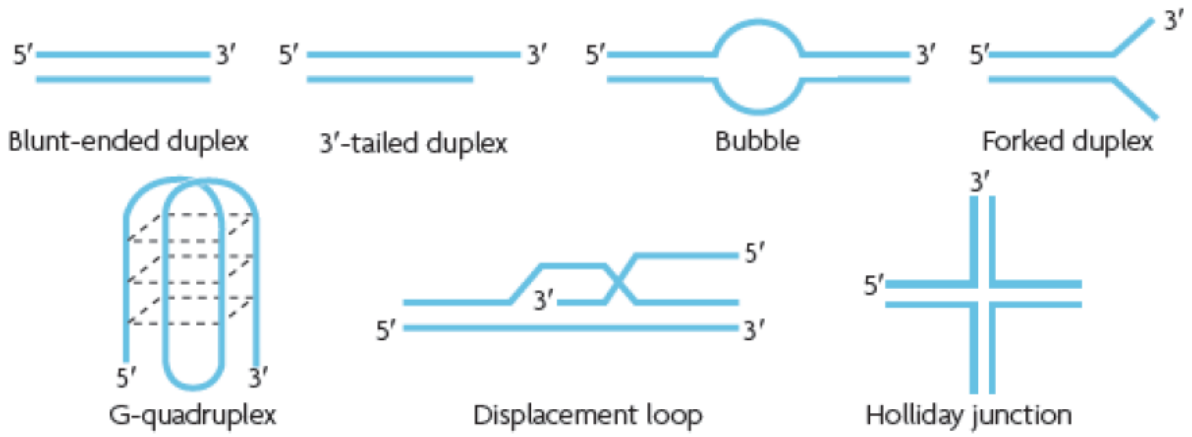

Figure 1.4. RecQ substrates. Typical synthetic substrates of RecQ helicases: some of them resemble intermediates of the homologous recombination process, like D-loops or displacement loops and Holliday junctions; G-quadruplex are formed at telomers, that are rich in guanine sequences; bubbles and forks mimic states of DNA replication. Picture adapted from Chu and Hickson, 2009.

1.3.1. Helicase activity.

The ***E. coli* RecQ protein** was shown to have the characteristic of ATP-dependent unwinding of dsDNA in a 3' to 5' direction *in vitro*. The property of dsDNA unwinding in ATP and Mg^{2+} dependent fashion is apparently conserved among all members of RecQ helicase family (Bennett and Keck, 2004). The *E. coli* RecQ helicase is active on a wide variety of partially and fully duplex DNA substrates (Umezu et al., 1990; Harmon and Kowalczykowski, 1998). RecQ helicase initiates homologous recombination in bacteria via the major RecF pathway by processing a dsDNA molecule to produce a 3'-terminated ssDNA that is then used by the RecA protein for homologous pairing (Harmon and Kowalczykowski, 1998; Handa et al., 2009). In addition to the promotion of homologous recombination, RecQ also disrupts joint molecules such as D-loops and four-way junctions *in vitro,* indicative of a role in suppressing illegitimate recombination by disrupting aberrant recombination intermediates (Harmon and Kowalczykowski, 1998; Kowalczykowski, 2000).

As in the case of human BLM and Drosophila BLM helicases, **yeast Sgs1** is also able to dissolve the double Holliday junctions (dHJs) that arise during homologous recombination to produce non-crossover products (Plank et al., 2006; Wu and Hickson, 2003; Wu et al., 2005). Both Sgs1 and its mammalian ortholog are shown to promote 5'-end resection at double strand break (DSB) sites to form 3'-ssDNA tails which is a crucial early step in homologus recombination (Gravel et al., 2008). Biochemical assays performed with truncated version of yeast Sgs1 protein containing helicase and RQC domains showed a preference for binding and unwinding DNA substrates with 3' ssDNA tails as well as branched substrates, including Holliday junctions (Bennett et al., 1998; Bennett et al., 1999). The interpretation of the biochemical data collected in the absence of N- and C- auxiliary domains is limited by the fact that the truncated version of the protein lacks domain important in mediating DNA interactions, especially via the HRDC domain. Studies on the full-length Sgs1 performed by Cejka and colleagues have described Sgs1 as a helicase that resembles the *E. coli* RecQ more than any of the human family members (Cejka and Kowalczykowski, 2010). In particular, Sgs1 was able to unwind blunt ended duplexes, as well as partial duplexes with 5' ssDNA tails. The same authors have also observed that the Sgs1 helicase is able to unwind large DNA duplexes up to 20 kb, which are of similar length as resection tracks

important for early recombinational events *in vivo* (Gravel et al., 2008; Cejka and Kowalczykowski, 2010). Importantly, the full length Sgs1 helicase showed up to 50 fold stronger DNA binding affinity for 3'ssDNA tail in comparison to its helicase domain, indicating that full-length Sgs1 protein provides additional domains for DNA binding.

A bioinformatic analysis predicted orthologues of RecQ4 in both plant and yeast, called **Hrq1** (Barea et al., 2008). Biochemical analysis of Hrq1 reveals a DEAH box- and ATP-dependent 3'-5' helicase activity on various DNA substrates including fork structure, long duplex DNA, bubbles and possesses a DNA strand annealing activity. Cells lacking Hrq1 suffer spontaneous genomic instability, hypersensitivity to DNA interstrand crosslinks (ICLs) and telomere addition to DNA breaks. Hrq1 supports the nucleotide excision repair of DNA damage caused by the chemotherapeutic agent cisplatin and, in certain genetic contexts, UV light. (Groocock et al., 2012; Kwon et al., 2012; Bochman et al., 2014).

Multicellular eukaryotic RecQ homologues are typically less active in terms of DNA unwinding. They have the ability to unwind DNA structures other than conventional B-form DNA duplexes (Figure 1.6), such as Holliday junctions (Constantinou et al., 2000; Karow et al., 2000), G- quadruplexes (Sun et al., 1998; Sun et al., 1999; Wu and Maizels, 2001), and D-loop like structures (Orren, Theodore et al., 2002; Bachrati, Borts et al., 2006; Popuri, Bachrati et al., 2008). Moreover, BLM and WRN are also active in unwinding gapped DNA substrates and triple helices (Mohaghegh et al., 2001; Brosh et al., 2001).

The ability to unwind G-quadruplex (G4) DNA and Holliday junction is unique to the RecQ helicases Sgs1, BLM and WRN (Sun et al., 1998; Kamath-Loeb et al., 2001; Vaughn et al., 2005). In *E. coli* there is only one way of processing the HJ intermediates through a mechanism of 'resolution' of Holliday junctions, where two strands of like polarity are nicked by the RuvC protein in a way that the products can be ligated without further processing (Connolly et al., 1991; Dunderdale et al., 1991). Consistent with their role in homologous recombination, replication and repair, many RecQ helicases efficiently bind and unwind cruciform DNA substrates resembling Holliday junctions (Harmon and Kowalczykowski, 1998; Garcia et al., 2004; Sharma et al., 2005). It was demonstrated that BLM, WRN and RecQ1 proteins selectively

bind HJs *in vitro*, and are capable to efficiently promote ATP-dependent branch migration of Holliday junctions and D-loops for several kilobases (Constantinou et al., 2000; Karow et al., 2000; Bugreev et al., 2008). In contrast to the D-loops formed by invasion of tailed DNA with the 3'-protruding ends, the loops with 5'-protruding ends and 3'-invading strands cannot be extended by DNA polymerase, and thus they could represent unproductive recombination intermediates during DSB repair. Thus, it is possible that RecQ1 activity may prevent accumulation of these unproductive and potentially toxic intermediates *in vivo* (Bugreev et al., 2008). BLM helicase is able to dissolve double Holliday junction recombinational intermediates through its association with topoisomerase IIIα, RMI1, and RMI2, forming a so-called 'dissolvosome' complex which is essential for genome stability (Wu and Hickson, 2003; Wu et al., 2006; Singh et al., 2008; Xu et al., 2008). Popuri and coworkers have demonstrated that RecQ1 cannot unwind G-quadruplexes or RNA-DNA hybrid structures and cannot substitute for BLM in the regression of a model replication fork. Conversely, RecQ1, but not BLM, is able to resolve immobile Holliday junction structures lacking homologous core, indicating that these enzymes play nonoverlapping functions in cells (Popuri et al., 2008).

The processivity of DNA unwinding by human RecQ helicases is relatively low and can be significantly enhanced by the presence of **accessory factors**, such as ssDNA-binding proteins. For example, human replication protein A (hRPA) binds to the partially unwound single stranded DNA and prevents their reannealing upon helicase dissociation, so that duplexes of more than one-hundred of base pairs in length can be unwound by WRN (Brosh et al., 1999), BLM (Brosh et al., 2000) and RecQ1 (Cui et al., 2004). In addition, hRPA might also promote DNA unwinding by specifically interacting with the selected RecQ helicase through a still poorly defined mechanism. Another single strand binding protein which cooperates with WRN and BLM in DNA unwinding, but exclusively on telomeric duplexes, is the human telomeric protein protection of telomeres 1 (POT1) (Opresko et al., 2005). POT1 promotes unwinding of telomeric forked duplex and D-loop structures (that are poor substrates for the helicase alone) by coating the ssDNA telomeric tail as it is released by the RecQ helicases. This may represent an important step for protection of the telomeric overhangs and for proper reformation of the telomere ends during DNA replication or recombination.

1.3.2. Strand-annealing activity.

All five human RecQ helicases are able to catalyze the annealing of complementary ssDNA molecules in an ATP-independent fashion (Garcia et al. 2004; Cheok et al. 2005; Machwe et al., 2005; Sharma et al. 2005; Macris et al., 2006). Moreover, the strand annealing activity is strongly inhibited by nucleotide binding (Macris et al., 2006; Cheok et al., 2005; Sharma et al., 2005).

Although the biological significance of ssDNA annealing activity of RecQ helicases remains to be determined, one could speculate that the annealing activity is required in replication fork regression via formation of 'chicken foot' structures at the sites of stalled replication *in vitro* and *in vivo* (Ralf et al., 2006; Wu and Hickson, 2006). When the lesions on the leading strand template blocks fork progression, the fork could regress via annealing of the nascent strands. In this process, called template switching, the longer lagging strand provides a template for the prematurely terminated leading strand to be extended. After extension of the leading strand, the regressed fork (four-way junction) can be reversed by branch migration, thus re-establishing an active replication fork (Cheok et al., 2005; Ralf et al., 2006). In addition, the ssDNA annealing activity might be also required to promote branch migration of DNA junctions during homologous recombination. However, the possible roles of the annealing activity of RecQ helicases in human cells are yet to be demonstrated.

1.4. Cellular and physiological roles of RecQ helicases

1.4.1. RecQ helicases and DNA repair.

RecQ helicases interact with proteins involved in DNA repair and participate in various DNA repair pathways, which is consistent with the high affinity of these helicases for DNA structures that resemble DNA repair intermediates. In the cells, complex networks of the DNA repair machinery are activated upon DNA damage. The four main DNA-repair pathways in the cell include base excision repair (BER), which repairs oxidative DNA base modifications such as 8-oxoguanine (8-oxoG), alkylation base damage and ssDNA breaks (SSBs); nucleotide excision repair (NER), which repairs bulky helix-distorting DNA lesions; mismatch repair (MMR), which repairs single nucleotide mismatches and small insertion–deletion mispairs; and double strand

break repair (DSBR), which repairs DSBs either using the homologous recombination (HR) or the nonhomologous end-joining (NHEJ) pathway.

Base Excision Repair (BER) removes oxidative, abasic and other functional base modifications from DNA. Various DNA glycosylases with overlapping DNA lesion specificity are expressed in mammalian cells. These glycosylases recognize base lesions and remove the damaged nitrogenous base leaving the sugar-phosphate backbone intact, creating an apurinic/apyrimidinic (AP) site. Further on, AP endonuclease 1 (APE1) cleaves AP sites, producing a single-strand break intermediate which is utilized by DNA polymerase β (Pol β) for DNA repair synthesis, followed by ligation of the nick by DNA ligase (Wilson and Bohr, 2007).

It has been demonstrated both *in vitro* and *in vivo* that human RecQ proteins interact functionally with many proteins in BER and SSBR pathways. WRN strongly stimulate NEIL1 glycosylase, which removes formamidopyrimidines, 5-hydroxyuracil and 8-oxoG. APE1 is inhibited by WRN, and is upregulated in RTS cells (Das et al., 2007; Ahn et al., 2004; Schurman et al., 2009). WRN and BLM strongly stimulate Pol β and FEN1 endonuclease, and thus enhance strand displacement synthesis and nucleotide incorporation during BER (Brosh et al., 2001; Sharma et al., 2004; Harrigan et al., 2006; Sharma et al., 2004). The strand displacement synthesis of Pol β is enhanced by RecQ4 but not by RecQ5 (Das et al., 2007; Speina et al., 2010). WRN exonucleases also act as an autonomous proofreading enzyme for Pol β during BER, indicating that WRN is an important player in BER (Harrigan et al., 2007). The replicative protein flap endonuclease 1 (FEN1) plays a role in long-patch BER and interacts with most of the human RecQ helicases, this may also promote post-replication DNA repair (Das et al., 2007; Speina et al., 2010; Brosh et al., 2001; Sharma et al., 2004; Sharma et al., 2005; Sharma et al., 2004). Almost all the human RecQ enzymes interact or are modulated by poly(ADP-ribose) polymerase 1 (PARP1) which adds poly(ADP-ribose) moieties to chromatin binding proteins. WRN-deficient cells fail to activate PARP1 in response to oxidative and alkylation DNA damage, indicating a biologically important interaction of WRN and PARP1. However, PARP1 is hyperactivated in RecQ1-depleted cells exposed to oxidative stress and in BLM-, WRN- and RecQ5-depleted cells that are not exposed to exogenous stress. WRN and RecQ1 are inhibited by poly(ADP-ribose) and PARP1 inhibitors alter the cellular localization of RecQ4 (Sousa et al.,

2012; Thomas et al., 2013; von Kobbe et al., 2003; Sharma et al., 2012; Gottipati et al., 2010; Tadokoro et al., 2012; Popp et al., 2013; Berti et al., 2013; Woo et al., 2006).

Nucleotide Excision Repair (NER) removes bulky lesions from DNA, such as UV-induced pyrimidine dimers and carcinogen adducts. Limited evidence suggests that RecQ helicases might play a role in NER. It has been reported that WRN is stimulated by XPG (xeroderma pigmentosum complementation group G) and RecQ4 interacts with XPA (Trego et al., 2011; Fan and Luo, 2008), two proteins involved in NER. However, RecQ helicase-deficient cells are not hypersensitive to UV irradiation and there is no direct evidence that the RecQ helicases alter NER functionally *in vivo*.

Double-Strand Breaks (DSBs) can cause transient or permanent cell-cycle arrest, mutagenesis, chromosomal rearrangements and ultimately cell death or tumorigenesis. Human cells can repair DSBs by nonhomologous end joining (NHEJ), alternative nonhomologous end joining (Alt-NHEJ) or homologous recombination (HR). The error free HR-dependent DSB repair is based on a sister chromatid template that occurs only during the late S and G2 phase of cell cycle. NHEJ and Alt-NHEJ occur throughout the cell cycle and are error prone. RecQ helicases interact with several key components of the DSB repair machinery.

DSBs in nonproliferating human cells are repaired by **NHEJ** and Alt-NHEJ. The proteins involved in several steps of NHEJ are 53BP1, the Ku70/Ku80 heterodimer (Ku), DNA-PKcs and XLF/XRCC4/LIG4. It has been reported that WRN interacts with Ku and is a substrate of DNA-PKcs kinase. WRN exonuclease is stimulated by Ku and XRCC4/LIG4, in contrast DNA-PKcs stimulates WRN helicase but not WRN exonuclease (Rooney et al., 2004; Fattah et al., 2010; Bothmer et al., 2010; Bunting et al., 2010; Cooper et al., 2000; Li and Comai, 2000; Yannone et al., 2001; Karmakar et al., 2002; Kusumoto-Matsuo et al., 2010; Kusumoto et al., 2008). RecQ1 has also been reported to interact with Ku and RecQ1-deficient cells show a reduced Ku-DNA binding activity, thus indicating a role of RecQ1 in NHEJ (Parvathaneni et al., 2013). A recent study reported that RecQ4 interacts with Ku and DNA-PKcs via its N-terminal domain. RecQ4 also stimulates DNA binding of Ku to a blunt end DNA substrate, thus implicating its role in NHEJ (Shamanna et al., 2014).

PARP1 may act as a DNA damage recognition protein during **Alt-NHEJ**, followed by end resection by MRE11, CtIP and EXO1 and ligation by LigIII/XRCC1 (Audebert et al., 2004; Wang et al., 2006; Dinkelmann et al., 2009; Xie et al., 2009; Zang and Jasin, 2011; Wang et al., 2005). WRN, BLM and RecQ1 are shown to interact with EXO1 and MRE11/RAD50/NBS1 (MRN) and an interaction of WRN-LigIIIα has been reported (Aggarwal et al., 2010; Gravel et al., 2008; Nimonkar et al., 2008; Doherty et al., 2005; Cheng et al., 2004; Cheng et al., 2005; Nimonkar et al., 2011; Sallmyr et al., 2008). These preliminary findings suggest a role of RecQ helicases in Alt-NHEJ.

Homologous recombination (HR) is a cellular process that occurs in the late S and G2 phases of cell cycle. It repairs DNA DSBs when a homologous template is available, for example at stalled replication forks during S phase, at DNA lesions during mitosis and at chromosomal pairing during meiosis. Some of the proteins involved during this process are CtIP, MRN, RPA, DNA2, EXO1 and RAD51 (Ciccia and Elledge, 2010). After CtIP and MRN binds to the DSB, CtIP stimulates MRN to recruit two additional nucleases EXO1 and DNA2. RPA plays a role in stimulating BLM DNA unwinding and enforcing DNA2 for 5' to 3' resection. Studies report that MRN recruits and enhances the processivity of EXO1 and BLM increases the affinity of EXO1 for DNA ends. *In vitro* studies suggest that only BLM stimulates DNA end resection of DNA2 while RecQ1, BLM and WRN stimulate EXO1 (Aggarwal et al., 2010; Nimonkar et al., 2008; Doherty et al., 2005; Nimonkar et al., 2011). It has also been reported that BLM and RecQ5 show an antirecombination activity by disrupting the RAD51 nucleoprotein filament essential for homology search and D-loop formation. WRN, RecQ1 and RecQ4 also colocalize with RAD51, however, WRN and RecQ1 cannot disrupt RAD51 filaments (Hu et al., 2007; Wu et al., 2001; Bugreev et al., 2007; Schwendener et al., 2010). WRN has also been reported to interact with several other proteins involved in HR including MRN, BRCA1, RAD52 and RAD54 (Cheng et al., 2004; Cheng et al., 2006; Baynton et al., 2003; Otterlei et al., 2006). RecQ4 is recruited to laser-induced DSBs and RecQ4-deficient cells are sensitive to ionizing radiation. RecQ4 interacts with RAD51 in cells exposed to etoposide (Singh et al., 2010; Kohzaki et al., 2012; Petkovic et al., 2005). WRN and BLM helicases *in vitro* preferentially bind to DNA substrates that resemble recombination intermediates such as Holliday junctions (HJs). Moreover, Bloom in

cooperation with topoisomerase IIIa and RMI1-RMI2 complex can process double HJs to generate non-crossover products (a process named dissolution of HJs), consistent with the fact that BLM defective cells are deficient in SCE (Wu and Hickson, 2003; Yin et al., 2005).

1.4.2. RecQ helicases and DNA replication.

The genomes of eukaryotic cells are replicated during the S phase of the cell cycle, starting from many replication origins distributed along the genome. To initiate DNA replication, many proteins assemble on, and dissociate from, origins (Masai et al., 2010). First, the six subunit origin recognition complex (ORC, Orc1-6) associates with replication origins throughout the cell cycle and the hexameric Mcm2-7 complex, which is a catalytic core DNA helicase, is loaded onto the ORC-bound origins with the aid of Cdc6 and Cdt1 to form the pre-replicative complex (pre-RC) from late M to G1 phases when cyclin-dependent kinase (CDK) is not activated (Figure 1.5). In *S. cerevisiae*, Sld3-Sld7 dimer and Cdc45 form a complex and associate with the pre-RC-formed origins in a mutually dependent manner (Tanaka et al., 2011). The activation of CDK at G1/S boundary leads to the assembly of other replication proteins on origins to form the pre-initiation complex (pre-IC) (Figure 1.5). The Sld2 and Sld3 are the two essential targets of CDK, phosphorylation of the two targets in S-phase leads them to interact with Dpb11 (Tanaka et al., 2007b; Zegerman and Diffley, 2007). These interactions are essential for DNA replication initiation in *S. cerevisiae*. Among the assembled proteins on origins, the heterotetrameric GINS complex, Cdc45, and Mcm2-7 form an active DNA helicase, the Cdc45-Mcm-GINS (CMG) complex, a core component of the replisome progression complex at the replication forks (Moyer et al., 2006). The unwinding of origin DNA by the CMG complex leads to the initiation of the synthesis of DNA strands by three DNA polymerases, α, δ, and ε (Polα, Polδ and Polε). Polα associates tightly with primase and extends a short DNA strand after the primer RNA synthesized by the primase. Subsequently, the five-subunit clamp loader, replication factor C (RFC), binds to the 3'-OH end of newly synthesized DNA and loads Polδ or Polε with a proliferating cellular nuclear antigen (PCNA, homotrimer sliding clamp that increases Polδ and Polε processivity). At the replication forks, Polδ and Polε synthesize mainly the lagging and leading strands, respectively (Kunkel and Burgers, 2008).

Growing evidences are implying that RecQ helicases are implicated in DNA replication. It has been demonstrated that RecQ4 and RecQ1 bind directly to the human replication origins at the G1/S border, after origin licensing and play distinct roles in DNA replication initiation and replication fork progression (Thangavel et al., 2010).

One putative role for these enzymes is to act as a 'roadblock' remover, because replication of a DNA damage-containing template can have strong impacts on replication fork depending on the nature of lesions (Cox et al., 2000; Rothstein et al., 2000). RecQ helicases interact physically and functionally with several key proteins involved in DNA replication, such as replication protein A (RPA), proliferating cell nuclear antigen (PCNA), DNA polymerase δ and FEN1 endonuclease. In particular, the interaction with RPA might be necessary for efficient unwinding of large regions of duplex DNA (Brosh et al., 1999; Brosh et al., 2000). It is also possible that the DNA strand-annealing activity of RecQ helicases is important for their functions at the sites of blocked DNA replication fork since for example, one of the proposed roles of BLM at stalled replication forks, which contain a gap ahead of the leading strand, is to promote fork regression (Ralf et al., 2006). In this model of fork regression, BLM is promoting annealing of the leading and lagging strands to form characteristic ''chicken-foot'' like structures, which could be later resolved to the normal fork-like architecture by a mechanism similar to branch-migration of HJs. Recent *in vivo* analysis of the newly synthesized DNA identified a role for BLM in efficient restart of replication forks and in the suppression of new origin firing after replication stress (Davies et al., 2007; Rao et al., 2007). WRN, in contrast, is not required for efficient restart but is needed to ensure normal fork progression after genotoxic treatments (Sidorova et al., 2008).

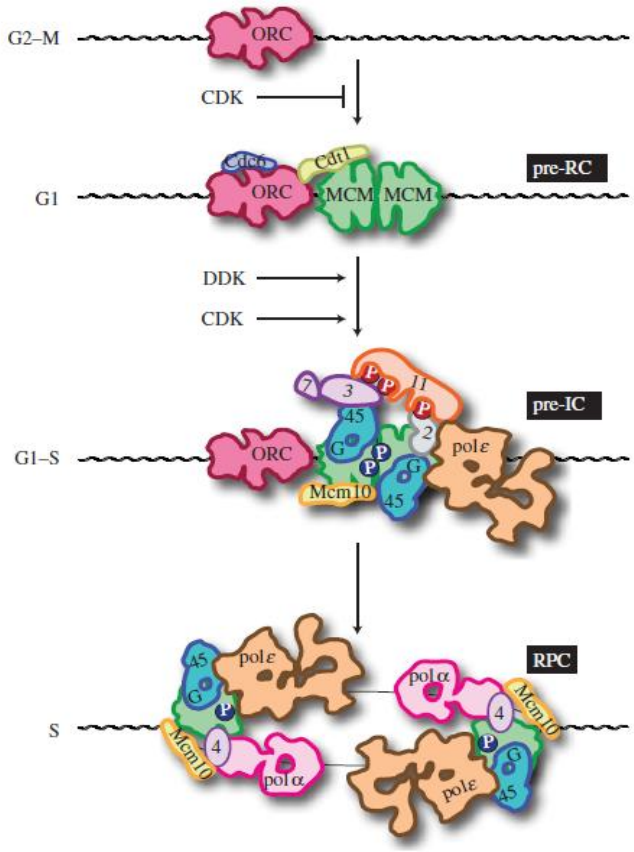

Figure 1.5. Assembly of DNA replication complexes. The individual steps leading to the assembly of bidirectional replisomes in yeast is outlined. Names associated with each of the complexes are shown on the right: pre-RC, pre-replication complex; pre-IC, pre-initiation complex; RPC, replisome progression complex. Cell cycle phases permissive for the individual steps are shown on the left. For simplicity, some of the protein names have been abbreviated: 11, Dpb11; 3, Sld3; 7, Sld7; 2, Sld2; G, GINS; 45, Cdc45; 4, Ctf4. Cyclin-dependent kinase (CDK) phosphorylation sites are shown in red, Dbf4-dependent kinase (DDK) phosphorylation sites are shown in blue. Adapted from Diffley, 2011. See text for explanation.

The involvement of RecQ4 in replication initiation at the replication origins was initially suggested by studies performed with *Xenopus* egg extract system. The N-terminal fragment of RecQ4 shows homology with the yeast replication factor Sld2/Drc1 and interacts with Cut5, which is a Xenopus homolog of *S. cerevisiae* Dpb11 and is required for loading of DNA polymerases onto the chromatin. Genetic complementation experiments in which human wild type, but not ATPase deficient RecQ4 was able to restore replication activity in *Xenopus* depleted extracts indicated that human RecQ4 is very important for initiation of DNA replication (Sangrithi et al., 2005; Matsuno et al., 2006). The function in replication of human RecQ4 is probably preserved in RTS patients, since all the known RecQ4 mutants found in patients maintain the Sld2-like domain at the N-terminus of the protein. Moreover, mutations in N-

terminal portion of RecQ4 are embryonic lethal in mice, supporting an essential role of RecQ4 in DNA replication (Ichikawa et al., 2002).

By chromatin immunoprecipitation experiments, the group demonstrated that RecQ4 is recruited to origins at late G1 after ORC and MCM complex assembly, while RecQ1 and additional RecQ4 are loaded at origins at the onset of S phase when licensed origins begin firing. Both proteins are lost from origins after DNA replication initiation, indicating either disassembly or tracking with the newly formed replisome. Moreover, Xu and collaborators showed that RecQ4 is associated with the replicative MCM2-7 helicase and other factors of replication initiation complex through an interaction with MCM10 (Xu et al., 2009). In agreement with these observations, nascent-DNA experiments and clonogenic assays showed that RecQ1-depleted, and even more strikingly, RecQ4-depleted cells have reduced amounts of nascent, newly synthesized DNA as well as compromised cellular proliferation. In addition, DNA fiber assays performed on RecQ1-depleted cells showed shorter replication tracks in comparison to control cells, suggesting that RecQ1 might play an additional role in replication fork progression in unperturbed cells (Thangavel et al., 2010). It has also been reported that in vitro RecQ1 helicase activity remodels chicken foot-like DNA substrates into replication fork-like molecules but fails to regress forks. Single-molecule DNA fiber analysis revealed that RecQ1-deficient cells exposed to low dose CPT fail to restart replication (Berti et al., 2013). Collectively, these studies suggest that RecQ4 appears to function early, in replication initiation, when pre-replication complex assembly takes place and active replisomes are assembled, while RecQ1 is also required for efficient replication initiation and play an additional role during replication elongation and fork restart.

Data from Kanagaraj and co-workers suggest that also RecQ5β is involved in DNA replication and fork repair by promoting the regression of stalled replication forks. The authors demonstrate that RecQ5β facilitates template-switching to catalyze the lagging strand unwinding and strand exchange on RPA-coated forked structures. In particular, they identify a short region located between the amino acids 561–651 to be important for the processing of forked structures. In addition, they show that RecQ5β interacts with the polymerase processivity factor, PCNA as well as that this helicase localizes in to DNA replication factories in S phase nuclei (Kanagaraj et al., 2006). Moreover, the overexpression of RecQ5 releases cells from the cell cycle arrest upon

thymidine treatment as these cells display fewer RPA foci and less γH2AX activation compared to the cells with endogenous levels of RecQ5. These data suggest that RecQ5 stabilizes the replication fork allowing replication to overcome the effects of thymidine and complete the cell cycle (Blundred et al., 2010).

1.4.3. RecQ helicases and transcription.

RecQ helicases might also be inlvoved in transcription. Biochemical and cellular evidences indicate that WRN modulates RNA polymerase II transcription (Balajee et al., 1999) and RecQ5 was shown to interact with RNA polymerase II (Pol II), and inhibits Pol II transcription both at the initiation and elongation steps, suggesting a possible involvement of this helicase in transcription or transcription-associated recombination processes (Aygun et al., 2008). In agreement with this, Islam and co-workers found that RecQ5 can bind both the initiation (Pol IIa) and elongation (Pol IIo) forms of the polymerase through two independent domains (Phatnani and Greenleaf, 2006; Islam et al., 2010). The authors suggest that the inhibition of transcription by RecQ5 occurs at the initiation, rather than in the elongation process, to reduce transcription-associated replication impairment and recombination. Moreover, they propose that RecQ5 promotes genome stabilization through an additional mechanism using its intrinsic helicase activity to directly decrease HR events by dissociating D-loops and RAD51 filaments. Another study independently showed that RecQ5β enzyme interacts with phosphorylated C-terminal repeat domain (CTD) of the largest subunit of PolII. Moreover, RecQ5 associates with RNA polymeraseII-transcribed genes, and its density on these genes correlates with the density of CTD phosphorylation, which is associated with the productive elongation phase of transcription (Kanagaraj et al., 2010).

1.4.4. RecQ helicases and telomere stability.

Telomeres are protein-DNA structures at chromosome ends that preserve genome stability, survival and proliferation at the cellular level and prevent degenerative diseases and cancer at the organismal level. These specialized nucleoprotein complexes that cap the ends of linear eukaryotic chromosomes are very important in order to distinguish chromosome ends from DSB as well as to facilitate the replication of chromosome ends. Telomeric DNA contains stretches of

tandemly repeated short G-rich sequences (TTAGGG in humans) which together with the multiprotein shelterin comlex form a specialized telomeric D-loop (T-loop) structure (Griffith et al., 1999; O'Sullivan and Karlseder, 2010; Calado and Young, 2012). The shelterin complex includes TRF1 and TRF2, proteins that bind to the double-stranded telomeric DNA and POT1, a protein that binds to the single-stranded telomeric DNA overhang (O'Sullivan and Karlseder, 2010). The closed configuration of the T-loop represents a protective cap that defines the natural end of the chromosome and masks the telomeres from the DNA damage response (DDR) machinery.

Studies suggest interaction of WRN, BLM and RecQ4 with shelterin proteins. The shelterin proteins generally stimulate unwinding of telomeric D-loops and forks by RecQ helicases. WRN helicase and endonuclease activities cooperate to dissociate D-loops (Opresko et al., 2004; Ghosh et al., 2012; Opresko et al., 2002; Lillard-Wetherell et al., 2004; Opresko et al., 2005). WRN and RecQ4 have been reported to localize to some telomeres during S phase and replication stress enhances telomeric localization of WRN (Liu et al., 2010). A recent study suggests a role of RecQ4 in the repair of thymine glycol lesions to promote efficient telomeric maintenance (Ferrarelli et al., 2013). In contrast, BLM localizes to a subset of telomeres in either late S or G2/M phase (Barefield and Karlseder, 2012). Telomere fragility is observed in RecQ4- and BLM-deficient cells, whereas WRN and BLM unwind G4s and preliminary data indicates that RecQ4 does not (Mohaghegh et al., 2001; Rossi et al., 2010).

RecQ helicases also participate in a second process of telomere maintenance called ‘alternative lengthening of telomeres’ (ALT), which is telomerase-independent and recombination mediated. In *S. cerevisiae*, Sgs1 acts in the resolution of recombination intermediates, when critically short telomeres undergo recombination to restore the telomeric length. A fraction of telomeric DNA from human cell lines maintaining the telomeres by ALT co-localizes with WRN (Johnson et al., 2001). In addition, the BLM helicase was also connected to recombination-driven amplification of telomeres in ALT-positive immortalized cells (Stavropoulos et al., 2002). Moreover, BLM acts in the ALT mechanism in complex with TOPIIIα and TRF2, possibly through the resolution of intermediate DNA structures that arise during telomere recombination (Temime-Smaali et al., 2008).

1.4.5. Post translational modifications of RecQ helicases

Post translational modifications of proteins (such as phosphorylation, sumoylation, ubiquitination, glycosylation etc) are events that finely regulate their destiny or their functions inside the cell.

BLM undergoes both constitutive and conditional phosphorylation. These changes are likely to play a role in the targeting and activation of BLM in response to DNA damage, replication stress, and mitotic cycling (Böhm and Bernstein, 2014). Phosphorylated Bloom loses its association with the nuclear matrix and it is found mainly in the nucleoplasmic fraction. Therefore, phosphorylation might facilitate the arrival of Bloom to the sites of DNA damage (Beamish et al., 2002; Dutertre et al., 2000). The modification of BLM lysine residues by SUMO (sumoylation) after prolonged, DSB-inducing fork arrest may switch BLM from DSB avoidance to a DSB repair mode (Eladad et al., 2005; Ouyang et al., 2009). This switch may be mediated via decreased association of sumoylated BLM with RPA, together with increased binding to RAD51 (Ouyang et al., 2013). Finally, ubiquitylation of BLM appears to play a role in both the response to replication stress and nuclear partitioning (Tikoo et al., 2013).

Phosphorylation of WRN by ATR on serine residues promotes the retention of WRN at stalled forks to preserve their ability to restart replication. Prolonged fork arrest results in MUS81-dependent cleavage (Franchitto et al., 2008; Murfuni et al., 2012) to generate DSB ends that trigger ATM-dependent WRN phosphorylation. These alternative phosphorylated states may facilitate the switch from a replicative to DSB repair mode, with dissociation of WRN from DSB sites (Ammazzalorso et al., 2010; Pichierri et al., 2012). WRN was shown to bind free poly(ADP-ribose) (PAR) via a newly identified PAR-binding motif in the WRN exonuclease domain (Popp et al., 2013). This may enable WRN in downregulating the DNA damage response. WRN is acetylated by p300 acetyltransferase in a DNA damage dependent manner and the acetylation helps WRN to localize in the nucleoplasmic fraction (Blander et al., 2002). Acetylation might also assist WRN involved in BER (Muftuoglu et al., 2008a).

The histone acetyltransferase p300 acetylates RecQ4 at up to five different lysine residues contained in a 30-residue stretch of RecQ4 that is required for nuclear localization. RecQ4 acetylation by p300 disrupts nuclear import and leads to the cytoplasmic accumulation of RecQ4 (Dietschy et al., 2009). Phosphorylation of RecQ4 Sld2-like domain by CDK supports its role in

DNA replication modulating the interaction RecQ4-MCM10 (Xu et al., 2009). When cells are exposed to oxidative stress, RecQ4 interacts with PARP1 and gets poly(ADP-ribosyl)ated, implying that this modification might regulate the RecQ4 activity in DNA repair (Woo et al., 2006).

RecQ1 appears to be phosphorylated when the cells are exposed to ionizing radiations (IR). This event was suggested to help RecQ1 associate with chromatin when the cells are exposed to IR, although the phosphorylating kinase and the specific role of phosphorylated RecQ1 remain elusive (Sharma and Brosh, 2007).

1.4.6. Interactions among RecQ helicases

Significant functional interactions between the RecQ helicases have been documented. Although WRN and BLM share similar substrate specificity, BLM inhibits the WRN exonuclease (von Kobbe et al., 2002) and RecQ4 specifically stimulates BLM (Singh et al., 2012). BLM promotes retention of RecQ4 at DSBs *in vivo* and cells defective in both BLM and RecQ4 helicase domain proliferate slowly with elevated SCEs. Cells that are defective in WRN and RecQ5 demonstrate synthetic lethality. *In vivo* RecQ5 dissociates slowly from DSBs in WS and BS fibroblasts. However, RecQ5 and WRN cooperate during repair and restart of synthetic stalled replication fork-like structures (Popuri et al., 2013). These results suggest that RecQ5 and WRN play cooperative and complementary roles.

1.5. RecQ helicases and disease

1.5.1. RecQ helicase syndromes.

As mentioned above, three of the five human RecQ helicases (BLM, WRN, RecQ4) are associated with rare genetic diseases, characterized by a mixture of overlapping and unique features.

Bloom syndrome (BS) was first identified in 1954 by David Bloom who described patients with congenital short stature and skin changes reminiscent of systemic lupus erythematosus (Bloom, 1954; German, 1993). It is a rare autosomal recessive disorder with a high predisposition to cancer development and is caused by defects in the *BLM* gene, located on chromosome 15q26.1 (Ellis et al., 1995; German et al., 2007; German, 1997). Clinical features of individuals with BS

include pre- and post-natal growth retardation, congenital short stature and a skin rash, which develops with sun exposure early in patient's life and may later become chronic with skin hyper- or hypo-pigmentation. Furthermore, male infertility and immunodeficiency, characterized by a decrease in IgA and IgM levels, are also observed in BS patients. Cells from BS patients have a high frequency of somatic mutations and high tendency to develop almost all types of cancers. There is an elevated risk of developing common adult epithelial tumors such as colon, breast and lung cancer; leukemias and lymphomas; sarcomas; and rare pediatric tumors such as Wilms' tumors (German, 1997). The hallmark feature of BS cells is the increased frequency of sister chromatid exchanges (SCEs) and homologous recombination (HR). SCEs arise primarily as part of HR events that occur between sister chromosomes during repair of DNA damage in the S or G2 phases of the cell cycle, and are used as a molecular diagnostic markers for BS (Chaganti et al., 1974; German, 1995). BS cells are sensitive to UV radiation and to the ribonucleotide reductase inhibitor hydroxyurea (HU) and have comparatively high number of DNA gaps, breaks and rearranged chromosomes (Imamura et al., 2001). Some mutations in the *BLM* gene in BS patients lead to a loss of helicase function, suggesting that BLM helicase activity is an important determinant of BLM-associated phenotypes in human somatic cells (German et al., 2007).

Werner syndrome (WS), first reported by Otto Werner in 1904, is an autosomal recessive disorder caused by the specific mutations in *WRN* gene. The key clinical features of this syndrome are associated with a premature aging phenotype, including short stature, early graying and loss of hair, bilateral cataracts and scleroderma-like skin changes (Epstein et al., 1966; Goto, 1997; Werner, 1985; Yu et al., 1996). The short stature of WS patients results from a failure to undergo pubertal growth spurt. WS patients have a high risk to develop premature atherosclerosis, myocardial infarction and stroke, osteoporosis as well as diabetes mellitus (Epstein et al., 1966; Goto, 1997). WS individuals are also cancer-prone, although to a more limited extent than is seen in BS individuals; in particular, they display an elevated incidence of sarcomas (Futami et al., 2008a). Mutations in *WRN* gene lead to loss of the WRN protein and its catalytic activities (Moser et al., 2000). WS cell lines characteristically display replicative senescence and severe telomere shortening. WS cells have been reported to be sensitive to several DNA damaging drugs like 4-Nitroquinoline-1-oxide (4-NQO), mitomycin C, cisplatin,

and camptothecin, an inhibitor of DNA topoisomerase-I. WS cells display increased illegitimate recombination and a high frequency of chromosomal deletions and spontaneous DNA breaks. They also show a higher level of RAD51 foci, which probably indicates the persistence of unprocessed recombination intermediates (Pichierri et al., 2001). Furthermore, WS cells lack the ability to complete recombination reactions.

Rothmund-Thomson syndrome (RTS) is a rare autosomal recessive disorder linked to a mutation in the *RECQ4* gene. It was first described by Rothmund in 1868, as an unusual skin change together with bilateral juvenile cataracts (Rothmund, 1868), subsequent cases were reported by Thomson in 1936 (Thomson, 1936). The characteristic skin changes of RTS typically appear within the first months of life as a sun sensitive rash, swelling and blistering on the face. Besides skin lesions, additional features include sparse or absent hair, short stature with frequent bone and tooth abnormalities and cataracts. Like in above described syndromes, in RTS patients there is an elevated risk of cancer, mainly osteosarcoma (Wang et al., 2001). RTS cells have been reported to be genetically unstable, displaying a high frequency of chromosomal abnormalities such as translocations and trisomes (Vennos et al., 1992) and trigger premature cell senescence (Lu et al., 2014). The immune system appears to be intact, while fertility may be reduced although RTS females have given birth to normal offspring. Life expectancy of RTS individuals in the absence of cancer appears to be normal (Siitonen et al., 2009).

Two other syndromes are associated with a defective *RECQ4* gene: the **RAPADILINO syndrome** and **Baller-Gerold syndrome** (**BGS**). RAPADALINO syndrome is characterized by radial hypoplasia or aplasia, patellar hypoplasia and cleft or high arched palate, diarrhea and dislocated joints, little size and limb malformation. Unlike the RTS, RAPADILINO does not display the same high predisposition for malignancies. Clinical symptoms of BGS are radial ray hypoplasia, skeletal dysplasia, short stature, and craniosynostosis which is a typical feature in BGS (Siitonen et al., 2003; Van Maldergem et al., 2006).

1.5.2. Murine models of RecQ helicases

Murine models have been created for all the 5 human RecQ helicases, in order to better understand the relationships between mutations at the genomic level and phenotypes.

RECQ1 knockout mice produced viable and fertile animals without any particular organismal phenotype. Despite that, cytogenetic studies on mouse embryonic fibroblasts (MEFs) from these mice demonstrated genomic instability, particularly aneuploidy, spontaneous chromosomal breakage, frequent translocation events and hypersensitivity to ionizing radiation (Sharma et al., 2007).

Four **Bloom syndrome** (BS) mice models have been generated. Mice in which part of the *BLM* gene upstream of the helicase domain is removed showed growth defect and micronuclei similar to human BS patients (Chester et al., 1998). Mice in which exons 10–12 were replaced with HPRT (hypoxanthine–guanine phosphoribosyltransferase) showed a slight increase in the frequency of micronuclei (Goss et al., 2002). However, both knockout mice models had a short life span. Given the fact that *BLM* is not essential for viability in humans, Allan Bradley's group generated two more knockout models by replacing exon 2 using an embryonic stem cell (ES) method (Luo et al., 2000). This gave rise to viable homozygous mutant mice that closely recapitulated the cellular phenotype of Bloom syndrome patients and provided better model for the disease.

Two **Werner syndrome** (WS) mice models have been generated so far. The first model, where exons that encode motifs III and IV of the helicase domain were targeted, shared phenotypes similar to those of human WS patients. Similar to WS patients, the knockout mice acquired myocardial fibrosis, T cell lymphoma, and were prone to cancer (Lebel and Leder, 1998). The second model with a mutation that eliminates expression of the C terminus of the helicase domain did not show any organismal phenotype that resembles WS patients. However, WRN deficient mice in a p53-negative background show an acceleration of tumor formation as well as a change in the tumor spectrum compared to *p53* null mice (Lebel et al., 2001).

Three ***RECQ4*** knockout mice models have been reported so far. The first model with a knockout of exon 5-8, covering the N-terminal, proved to be embryonic lethal suggesting a very important function of this N-terminal domain (Ichikawa et al., 2002). The second model, where part of exon 13 which codes for part of the helicase domain was knocked out, showed severe growth

retardation and other organismal phenotypic characteristics resembling RTS patients (Hoki et al., 2003). The third model where exon 9 to 13 were knocked out showed typical RTS clinical features such as hypo-/hyperpigmented skin, skeletal limb defects and palatal patterning defects (Mann et al., 2005). The existence of two mouse models with similar phenotype of RTS patients confirms that mutation in the helicase domain of RecQ4 is the reason behind the RTS and is in line with where mutations are found in patients.

The ***RECQ5*** knockout mice are viable and develop without any clear abnormality (Hu et al., 2005; Hu et al., 2007). Cytogenetical analysis of MEFs derived from these mice revealed an elevated sister chromatid exchanges (SCEs) phenotype in the absence of RecQ5. Interestingly, detailed phenotypic analysis *RECQ5* deficient mice showed an increased incidence of multiple types of cancer upon aging (Hu et al., 2007). Moreover, cells from these animals exhibit elevated frequencies of spontaneous double-strand breaks (DSBs) and accumulation of chromosomal rearrangements in response to replication stress (Hu et al., 2009).

1.5.3. RecQ helicases and cancer

All the human RecQ disorders present cancer predisposition, consistent with the role of these proteins in the key steps of DNA replication and repair, but their cancer profiles are different. Patients with WS are susceptible primarily to thyroid cancer, melanoma, meningioma, soft tissue sarcomas and osteosarcoma (OS). In a study of spectrum of cancers in WS patients OS was found to comprise 7.7% of all neoplasms (Lauper et al., 2013). It has been shown that the MYC oncoprotein directly stimulates transcription of WRN, which may promote MYC driven tumorigenesis through prevention of cellular senescence normally mediated by the WRN protein (Grandori et al., 2003). In contrast, patients with BS are susceptible to all types of cancers seen in the general population but at a much higher frequency and at an earlier age. These include leukemias and lymphomas, epithelial cancers of colon, breast, head and neck, cervix, as well as OS (German, 1997). Among the RecQ4 associated disorders, patients with RTS have a very high and specific risk for OS in addition to non-melanoma skin cancers. Patients with RAPADILINO syndrome are also at risk for cancer, most commonly lymphomas as well as OS (Wang et al., 2001; Siitonen et al., 2009).

RecQ1 has been reported to be highly expressed in glioblastoma tissues and can be a new suitable target for anti cancer therapies aimed to arrest cell proliferation in brain gliomas (Mendoza-Maldonado et al., 2011). Studies have linked a single nucleotide polymorphism present in the RecQ1 gene to a reduced survival in pancreatic cancer patients (Li et al., 2006), also single nucleotide polymorphisms in WS are associated with cancer susceptibility (Wang et al., 2011; Wang et al., 2009b). While exact molecular mechanisms of tumor suppression have yet to be worked out fully, it is clear that deficiency of the WRN, BLM and RecQ4 proteins in humans predisposes to the development of cancer.

1.5.4. RecQ helicases as therapeutic targets

The RecQ helicase proteins may be direct targets for the treatment of cancer or other diseases. The helicase (and in the case of WRN the exonuclease), catalytic activities of the human RecQ helicases provide ready targets for newly identified small-molecule inhibitors (Aggarwal et al., 2011; Nguyen et al., 2013). The direct targeting of tumors with RecQ helicase inhibitors might, in addition, provide a clear, tumor-specific therapeutic advantage if combined with conventional chemotherapy (Brosh, 2013). The targeting of survival pathways specific to one or common to several RecQ helicase could provide a second approach to improve the therapy of patients with tumor-specific RecQ defects. These pathways have been identified in part by the recent gene expression profiling experiments and could be further interrogated by focused RNA inhibition or drug/small-molecule screens (Helleday et al., 2008; Evers et al., 2010; Ashworth et al., 2011; Brough et al., 2011).

1.6. RecQ4 helicase

The human *RECQ4* gene consists of 21 exons and lies on chromosome 8q24.3. The corresponding RecQ4 protein is 1208 amino acids long and has a molecular weight of 150 kDa. RecQ4 expression is tissue and cell-cycle phase specific. In particular, RecQ4 is highly expressed in testis, prostate glands and thymus and peaks in the S-phase of the cell cycle. As discussed above, mutations in the gene leads to three distinct genetic disorders: Rothmund-Thomson syndrome (Lindor et al., 2000; Wang et al., 2001), RAPADILINO (Siitonen et al., 2003), and Baller-Gerold syndrome (BGS) (Van Maldergem et al., 2006).

1.6.1. Cellular studies.

The role of RecQ4 in DNA replication has been suggested since the sequence homology between Sld2 in S. cerevisiae and N-terminus of RecQ4 has been observed (Sangrithi et al. 2005; Matsuno et al. 2006). Like Sld2 in S. cerevisiae, RecQ4 is suggested to be involved in DNA replication and plays a role in constructing replicative complex (Thangavel et al., 2010). The N-terminus of RecQ4 including the Sld2 homology region is essential for replication initiation in *Xenopus* oocyte extract and chromatin binding by DNA polymerase α (Matsuno et al., 2006). Besides the main role of RecQ4 in initiation of DNA replication, this enzyme has also other roles in various cellular pathways such as: response to S-phase arrest (Park et al., 2006), DSB repair (Kumata et al., 2007; Singh et al., 2010), nucleotide excision repair (Fan and Luo, 2008), base excision repair (Schurman et al., 2009), oxidative stress responses (Werner et al., 2006), telomere maintenance (Ghosh et al., 2012) and preserving mitochondrial DNA integrity (De et al., 2012; Croteau et al., 2012).

Park and colleagues showed that human mutant cells lacking RecQ4 are defective in UV-induced S-phase arrest, whereas cells defective in BLM exhibited a normal S-phase arrest following UV irradiation (Park et al., 2006). In keeping with this, inhibition of RecQ4 expression in human 293 cells caused a defect in inducing S-phase (replication) arrest following UV treatment. RecQ4-deficient human cells were also defective in inducing S-phase arrest following hydroxyurea treatment. Together, these results suggest that RecQ4 may have a unique role in replication fork arrest, which may not be shared by other members of RecQ family and might explain the extreme sensitivity to UV and oxidation damage in patients lacking RecQ4.

Kumata and colleagues described a possible involvement of RecQ4 in **DSB repair pathway** in *Xenopus* egg extracts, supported by the fact that after induction of DSBs, RecQ4 is loaded adjacent to Ku heterodimer-binding sites on damaged chromatin (Shamanna et al., 2014). Moreover, RecQ4-deficient fibroblasts are moderately sensitive to gamma-irradiation and accumulate more γH2AX and 53BP1 foci than control fibroblasts. Endogenous RecQ4 also colocalizes with γH2AX at the site of DSBs (Singh et al., 2010). A recent study reported that RecQ4 interacts with Ku heterodimer and DNA-PKcs via its N-terminal domain. RecQ4 also stimulates DNA binding of Ku heterodimer to a blunt end DNA substrate, thus implicating a role

in NHEJ (Shamanna et al., 2014). A possible role of RecQ4 in **NER** has been pointed out by Fan and Luo (2008). They showed that UV treatment of human cells resulted in the colocalization of the nuclear foci formed with RecQ4 and XPA. Consistently, RecQ4 could directly interact with XPA, and this interaction was stimulated by UV irradiation. Schurman and co-workers (2009) support a model in which RecQ4 regulates both directly and indirectly base excision repair (**BER**) capacity. In fact, in cells treated with H_2O_2, RecQ4 co-localizes with APE1 and FEN1, key participants in base excision repair. Additionally, RTS cells display an upregulation of BER pathway genes and fail to respond like normal cells to oxidative stress. Werner et al. (2006) suggest that enhanced oxidant sensitivity in RecQ4 deficient fibroblasts derived from RTS patients could be attributed to abnormal DNA metabolism and proliferation failure, implying a role for RecQ4 in **oxidative stress response**.

Recent studies show a novel role of RecQ4 in maintenance of **telomere integrity**. In fact the helicase localizes to telomeres, associates with shelterin proteins TRF1 and TRF2 and is able to resolve telomeric D-loop structures with the help of TRF1, TRF2 and POT1 (Ghosh et al., 2012). Croteau et al. (2012) demonstrated that RecQ4 is the first 3’ to 5' RecQ helicase to be found in both human and mouse mitochondria, and its loss alters **mitochondrial integrity**. A Mitochondrial Localization Signal (MLS) has been identified within the helicase N-terminus. The MLS causes the localization of RecQ4-p53 complex to the mitochondria. RecQ4-p53 interaction is disrupted after stress, allowing p53 translocation to the nucleus (De et al., 2012). A recent study shows that RecQ4 is critical for skeletal development by modulating p53 activity in vivo (Lu et al., 2015).

1.6.2. Biochemical features

The first report on the biochemical characterization of RecQ4 showed DNA-dependent ATPase activity and strand annealing activity, but did not detect any helicase activity (Macris et al., 2006). Later studies have demonstrated that the recombinant human protein is an active DNA helicase. DNA helicase activity was detected for RecQ4 in reactions containing excess single-stranded DNA to prevent reannealing. Remarkably, it was suggested that in addition to the helicase activity linked with the region containing the helicase domain, a second helicase activity was also associated with the N-terminal region (Xu and Liu, 2009). Ishimi's group discovered

that the purified RecQ4 could displace the annealing single-stranded DNA without adding single-stranded DNA (Suzuki et al., 2009). They suggested that previous failure in detecting helicase activity could be due to the low concentration of ATP used in the assays. RecQ4 was in fact able to displace 17-mer but not 37- or 53-mers, indicative of a weakly processive enzyme. The movement of the helicase was from 3′ to 5′, consistent with other RecQ helicases (Suzuki et al., 2009). The helicase activity of RecQ4 was confirmed later by another report (Rossi et al., 2010), where the helicase activity was shown to be inactivated by a mutation in the conserved helicase domain, confirming that the helicase domain only is responsible for the observed helicase activity.

RecQ4 from *Drosophila* was also shown to possess 3' to 5′ helicase activity as well as single-strand DNA annealing activity (Capp et al., 2009). A point mutation replacing a conserved lysine with asparagine leads to complete loss of helicase activity, but not singlestrand DNA annealing activity. This also indicates the presence of a unique DNA helicase domain in RecQ4.

Despite its importance, the structural and biochemical characterization of the protein is limited. The structure of the first 54 amino acids at the N-terminus of human RecQ4 is known to form a helical bundle resembling a homeodomain (Ohlenschläger et al., 2012). An in silico analysis done in our lab (Marino et al., 2013) provides new insights into the architecture and function of the enzyme by identifying three important features: a second region of homology with Sld2 (Figure 1.6) in addition to the N-terminal 150 amino-acid residues (Sangrithi et al., 2005), a cysteine-rich region classified as "retrovirus Zn finger like" or "Zn knuckle" (Figure 1.7), located between the Sld2 homology region and the helicase domain and a putative RQC domain at the C-terminus of the helicase domain including some of the key residues involved in DNA binding and unwinding (Figure 1.8).

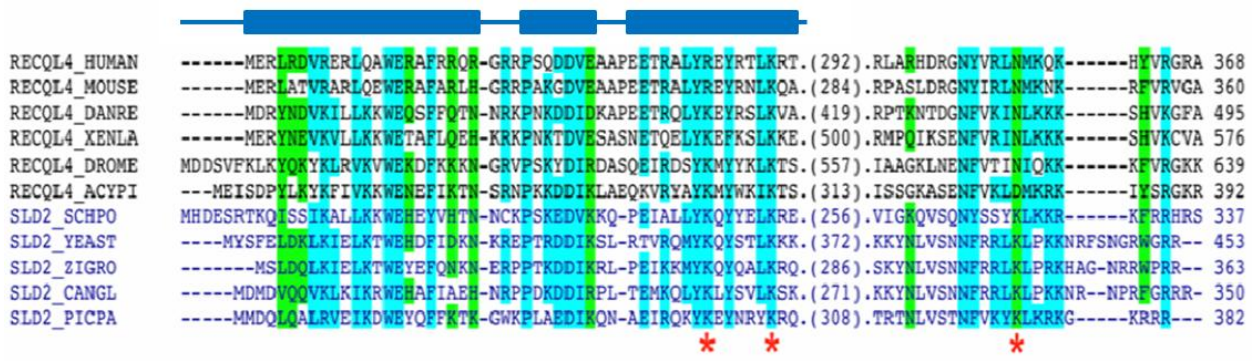

Figure 1.6. The Sld2 homology region. Sequence alignment between the N-terminal domain from a selection of RecQ4 proteins and the fungi Sld2 replication factors, showing that the homology extends to the C-terminus of Sld2. In blue are highlighted the residues that are identical or very similar (such as E/D, R/K/H, V/I/L/M, G/A, Y/F/W, S/T/P, N/Q) and in green residues that show a similar character (large polar/charged, hydrophobic, aromatic, etc.). The α-helices, as revealed by the NMR structure of the first 50 amino-acid fragment of human RecQ4 (PDB code: 2KMU) are shown as rods above the sequence. The three residues that have been implicated in binding to origin DNA for *S. cerevisiae* Sld2 (Bruck et al., 2011) are highlighted by a red star. Picture adapted from Marino, 2012.

The newly identified region of homology to Sld2, shown in figure 1.6, located immediately upstream the Zn knuckle, is highly conserved among species and includes numerous positively charged and aromatic residues possibly involved in DNA binding. In the study, Blast searches using various Sld2 homologues as targets identified the presence of an isolated Sld2 in Fungi, and of a fusion protein in Metazoa, Choanoflagellata and Viridiplantae. This fusion event between the entire Sld2 and a RecQ helicase strongly argues in favour of a role for a RecQ helicase activity in the early processes of eukaryotic DNA replication, as confirmed by a number of *in vivo* studies (Sangrithi et al., 2005; Matsuno et al., 2006; Wu et al., 2008; Xu et al., 2009; Im et al., 2009; Thangavel et al., 2010).

Figure 1.7. The Zn-knuckle. Sequence alignment between the N-terminal region from a selection of RecQ4 proteins, showing the predicted Zn-knuckle. The residues predicted to coordinate a Zn atom are identified by a red cross. In the human sequence the second cysteine is substituted by an asparagine.

Zn knuckles are short cysteine-rich sequences wrapping around a Zn^{2+} ion, often present in multiple copies in nucleocapsid proteins of RNA retroviruses and in eukaryotic gene regulators. Within the RecQ4 paralogues, the Zn knuckle motif is well conserved albeit with variants of the canonical zinc site: for example the human sequence has the second cysteine substituted by an asparagine (CNHC) while the *Xenopus laevis* region contains the canonical Zn^{2+} ligands: CCHC (Figure 1.7). Asparagine and glutamine are occasionally found as Zn^{2+} ligands (Cleasby et al., 1996; Koutmos et al., 2008). This domain is located in between the additional Sld2 homology region and the Helicase core of the protein and thus can share a role with either of the domains. A variety of Zn-knuckles have been shown to bind both ssDNA and ssRNA. Most have a preference for ssRNA binding (Buckman et al., 2003; Loughlin et al., 2012), but the Zn-knuckle has also been found in transcriptional activators and nucleic-acid binding proteins involved in DNA binding (Fields et al., 2008; Armas et al., 2008).

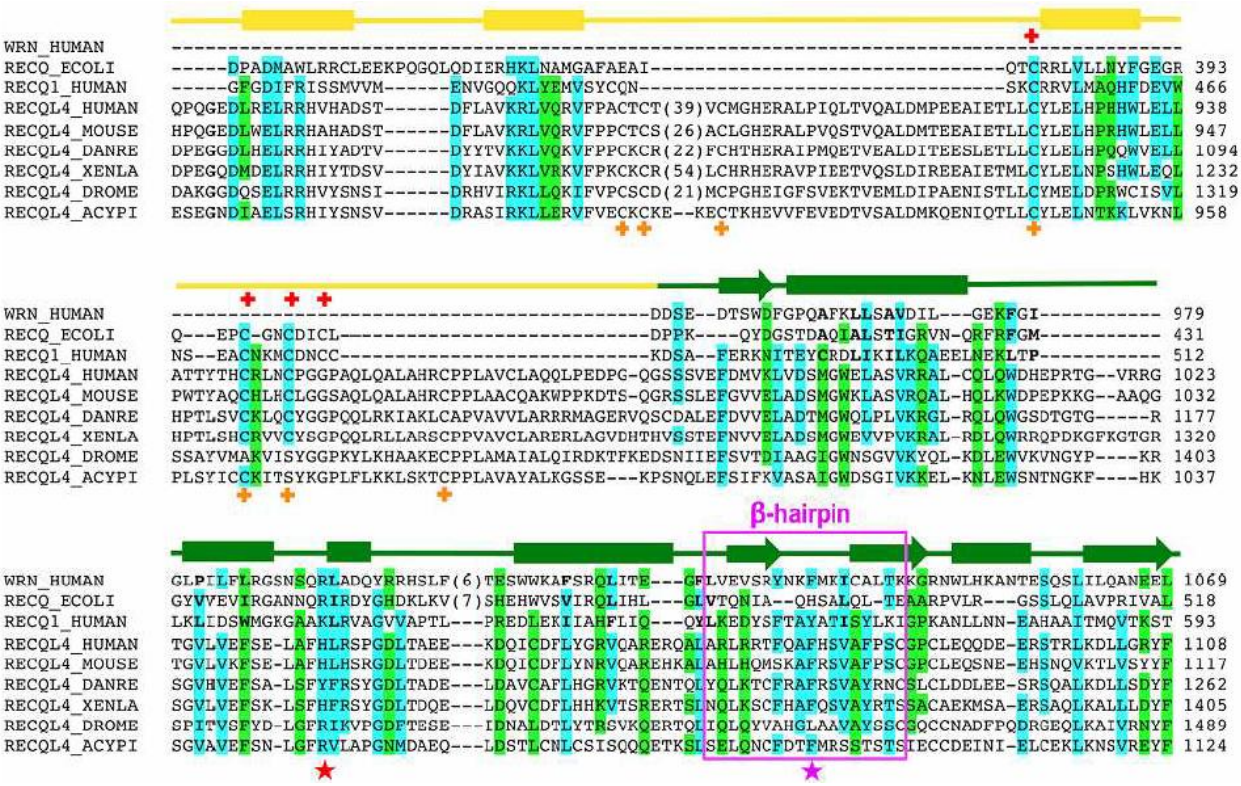

Figure 1.8. RQC domain of RecQ4. The putative RQC domain of RecQ4 is aligned with the equivalent region in human RecQ1 and *E. coli* RecQ, and with the WH domain of Werner. The alignment is structure-based (PDB accession codes: 1OYW, 2V1X, 3AAF). Highlighted are the residues that are conserved within the RecQ4 sequences and at least one other sequence. In blue are highlighted the residues that are identical or very similar (E/D, R/K/H, V/I/L/M, G/A, Y/F/W, S/T/P, N/Q) and in green

residues that show a similar character (large polar/charged, hydrophobic, aromatic, etc.). The secondary structure elements are based on the crystal structure of human RecQ1 (α-helices are shown as rods and β-strands as arrows): in yellow is shown the region corresponding to the Zn domain, in green that corresponding to the winged helix domain. The four cysteine residues coordinating a Zn atom in the two crystal structure are shown by red crosses above the sequences; conserved cysteines which could possibly be involved in Zn binding in RecQ4 are shown by orange crosses below. The non-polar residues that have been described as part of the WH core in Werner (Kitano et al., 2010) are shown in bold. The conserved positively charged residue implicated in DNA binding (Kitano et al., 2010) is identified by a red star. The β-hairpin that in the RecQ1 and WRN proteins has been shown to act as a pin is enclosed in a magenta box, with the key tyrosine/phenylalanine identified by a magenta star. Picture adapted from Marino et al., 2013.

All the RecQ helicases except RecQ4 were known to have RQC domain until recently a bioinformatic analysis predicted the presence of RQC domain in RecQ4 proteins. Structural study of E. coli RecQ (Bernstein et al., 2003) and human RecQ1 (Pike et al., 2009) reports the presence of two motifs: a Zn binding motif and winged helix (WH) motif within the RQC domain. The in silico study reports the presence of conserved residues in RecQ4 and indicates the presence of RQC domain (Figure 1.8). While the helices of the Zn domain are well conserved (Figure 1.8), two long insertions are present before and after the third helix. A number of conserved cysteines are found in the insertions, structurally located in positions compatible with a role in Zn coordination. Most of the key residues are conserved in the WH domain including the hydrophobic core and the aromatic residue at the β-hairpin. In RecQ helicases the aromatic residue at the β-hairpin is shown to be involved in DNA binding and unwinding, by acting as a pin to disrupt Watson-Crick base pairing (Pike et al., 2009; Kitano et al., 2010). In RecQ4 the conserved aromatic reisdue at the β-hairpin is a phenylalanine in most sequences, likely having a similar functional role (Figure 1.8).

1.7. Present contribution

RecQ helicases are diverse enzymes involved in all aspects of DNA metabolism, and their presence is essential to maintain genome integrity, being at the crossroad between DNA replication, repair and recombination. Since the discovery and the description of the syndromes linked to three of the human genes coding for RecQ proteins, many studies have been carried out, especially at the cellular level. Over the last few years efforts from a number of laboratories have contributed to a better understanding of the biochemical and cellular activities of these proteins. Many cellular functions have thus been proposed, ranging from stabilisation and repair of damaged replication forks, telomere maintenance, homologous recombination, base excision repair, and DNA damage checkpoint, but their exact and distinct roles in the cell are still somehow unclear.

Among these proteins RecQ4 is one of the less characterized, despite its role in genetic diseases and carcinogenesis. In contrast with the conservation of the SF2 helicase motifs, no DNA helicase activity was shown by the protein in the initial biochemical reports. But in subsequent studies it has been demonstrated that the recombinant protein displays a weak helicase activity. Although some laboratories have reported the expression of recombinant protein, the amounts obtained are low, and not suitable for a comprehensive biochemical and structural biology approach. The very weak helicase activity towards canonical fork substrates may suggest that the exact nucleic acid substrate may be missing. No large-scale protein purification had been reported, no detailed biophysical characterisation regarded the stoichiometry, no conclusive dissection of the domains responsible for the various activities and no structural information was available.

My PhD project therefore focussed on studying the architecture and mechanism of action of human RECQ4 by expressing and purifying in large amounts various fragments and domains and use them for biochemical and structural studies. The cloning and expression strategy was based on the previous bioinformatic analysis (Marino et al., 2013) and was meant to provide an experimental validation to the existence of the newly identified domains.

The results of this project are described into three sections: 3.1, 3.2 and 3.3.

Section 3.1 summarizes the cloning, small scale test expression and optimization of purification protocols to obtain homogenous pure recombinant fragments of human RecQ4 for structural and biochemical characterization. Out of the many constructs obtained I focussed on two: one encompassing both the Zn knuckle domain and the upstream region and one comprising both the helicase core and the putative RQC domain.

Section 3.2 summarizes the work done on Zn knuckle domain and the upstream regions. Due to the fact that the human Zn knuckle was unusual in lacking one of the canonial Zn binding residues, I carried out a parallel investigation on both the human and frog RecQ4 Zn knuckles. Both have been structurally and biochemically characterized using CD and NMR spectroscopy and EMSA techniques respectively. The three-dimensional structure of the canonical *X. laevis* Zn knuckle was determined by NMR, as part of collaboration with the NMR group of G. Musco (Milan). We also analyzed the effect of the newly identified Sld2 homology region on the biochemical properties of the Zn domains in terms of nucleic acid binding (Mojumdar et al., manuscript submitted).

Section 3.3 summarizes the work done on the helicase and RQC domain of human RecQ4. After optimizing the protocol to obtain large amount of purified protein, a detailed biochemical characterization was carried out in terms of helicase assays, ATPase assays, DNA binding and metal analysis. Numerous site-directed mutants were produced, including the Walker A mutant and the key residues within the putative RQC domain. The analysis of their biochemical properties fully validates the bioinformatic analysis (Mojumdar et al. manuscript in preparation). A molecular envelope of the helicase and RQC domain of the protein was obtained using Small Angle X-ray Scattering (SAXS), both in the presence and absence of DNA.

2. Materials and Methods

2.1. Construct design and cloning

The methods used for cloning the desired fragments into the vectors were Ligation Independent cloning (LIC) and Restriction free (RF) cloning. Ligation independent cloning (LIC) makes use of the 3'→5' activity of T4 DNA polymerase to create very specific 10-15 base single overhangs in the expression vector (Gileadi et al., 2008). PCR products with complementary overhangs are created by building appropriate extensions into the primers and treating them with T4 DNA polymerase. The annealing of the insert and the vector is performed in the absence of ligase by simple mixing of the DNA fragments (section 2.1.4.2).

Restriction Free (RF) cloning methodology facilitates the directional cloning of PCR products without restriction enzyme digestion or ligation reactions (van den Ent and Lowe, 2006; Unger et al., 2010). The RF cloning method allows a precise insertion of a DNA fragment into any desired position within a circular plasmid (section 2.1.4.1).

2.1.1 Expression vectors used for cloning of RecQ4 fragments

One commercially available bacterial expression vector and five other bacterial expression vectors kindly supplied by EMBL and SGC Oxford (Dr. Opher Gileadi) were used to produce the DNA constructs for the expression of recombinant proteins in *E. coli*. The main features of these vectors are summarized in Table 2.1. See Appendix A3 for vectors maps.

Vector	Promoter	Antibiotic resistance	Tag	Protease	Size (bp)	Cloning method	Supplier
pETM-30	T7	Kanamycin	6H-GST, N-ter	TEV	6346	Restriction enzymes	EMBL
pNIC-CTHF	T7	Kanamycin	6H-FLAG, C-ter	TEV	7260	LIC	SGC
pNIC-Bsa4	T7	Kanamycin	6H, N-ter	TEV	7284	LIC	SGC
pET SUMO/CAT	T7	Kanamycin	6H-sumo, N-ter	Sumo protease	6307	RF	Invitrogen

Table 2.1. List of vectors used for the cloning of RecQ4. The vectors used and their characteristics: promoter, antibiotic resistance, fusion tag and protease used to cleave the tag, size, supplier and the cloning method used. LIC = ligation independent cloning (Gileadi et al., 2008). RF = restriction free cloning (van den Ent and Lowe, 2006).

2.1.2. Polymerase chain reaction (PCR)

DNA sequences were amplified by PCR, using a Touchdown PCR cycle. The 50 µl of PCR reaction mixture contained 10ng of template DNA, 125ng of forward and reverse primer, 0.3 mM of dNTP mix, 1 mM MgSO4, 0.4 µl of Platinum Pfx DNA polymerase (2.5U/µl) and 1X Pfx amplification buffer.

The Touchdown PCR program used was the following:

3 min initial denaturation at 95°C

Then 4 cycles:

30 sec denaturation 95°C

1 min annealing 68°C

1-3 min extension 68°C

Then 4 cycles:

30 sec denaturation 95°C

1 min annealing 60°C

1-3 min extension 68°C

Then 4 cycles:

30 sec denaturation 95°C

1 min annealing 55°C

1-3 min extension 68°C

Then 20 cycles:

30 sec denaturation 95°C

1 min annealing 50°C

1 -3 min extension 68°C

Then a final 8-10 min step for extension 72°C

The plasmid used as template in the PCR reaction was a pGEX6p1, kindly provided by Vilhelm A. Bohr (Rossi et al., 2010), and contains wild type human RecQ4 with a N-terminal GST tag and C-terminal 9-histidine tag. The list of primers used to amplify RecQ4 deletion mutants is found in Table 2.2.

Name of primer	Sequence (5' to 3')	Suitable for vector	Cloning
445f	AGCCTGGACCCCACCGTG	(pNIC-CTHF, pNIC-Bsa4) / (pET-SUMO/CAT)	LIC / RF
475f	CAGCTGGGGCACCAAGCC	(pNIC-CTHF, pNIC-Bsa4) / (pET-SUMO/CAT)	LIC / RF
820f	CCCCAGGGCGAAGACCTG	(pNIC-CTHF, pNIC-Bsa4) / (pET-SUMO/CAT)	LIC / RF
895f	AGGGTCTGCATGGGCCAT	(pNIC-CTHF, pNIC-Bsa4) / (pET-SUMO/CAT)	LIC / RF
819r	CTGCAGGAAGAGGTGGCAGTG	(pNIC-CTHF, pNIC-Bsa4) / (pET-SUMO/CAT)	LIC / RF
852r	GGCTGGGAACACGCGCTGTAC	(pNIC-CTHF, pNIC-Bsa4) / (pET-SUMO/CAT)	LIC / RF
1112r	TTCCTCTTCCTCAAAGTAGCG	(pNIC-CTHF, pNIC-Bsa4) / (pET-SUMO/CAT)	LIC / RF
1208r	GCGGGCCACCTGCAGGAGCTC	(pNIC-CTHF, pNIC-Bsa4) / (pET-SUMO/CAT)	LIC / RF
335f	ATAATCCATGGGCACAGCCCCCCTG	pETM-30	Restriction site: NcoI
394f	ATAATCCATGGGGTGCCACAGTCACAACC	pETM-30	Restriction site: NcoI
427r	ATATAGAATTCTCAAGCATCTGTGTCTTCCTCACT	pETM-30	Restriction site: EcoRI

Table 2.2. List of primers used for the cloning of RecQ4 deletion mutants. The name of the primer indicates the start residue of a construct (for the forward primers: f) and the end residue (for the reverse primers: r). For pNIC-Bsa4 the following sequences complementary to the vectors have been added to the 5' end of forward and reverse primers: forward 5'TACTTCCAATCCATG3' (ATG in frame with the desired coding sequence), reverse 5'TATCCACCTTTACTGTCA3'. For pNIC-CTHF the following sequences have been added to the 5' end: forward 5'TTAAGAAGGAGATATACTATG3', reverse 5'GATTGGAAGTAGAGGTTCTCTGC3'. And for pET-SUMO/CAT vector the sequence complementary to the vector 5'ATTGAGGCTCACAGAGAACAGATTGGTGGT3' is added to the 5' of forward primer and 5'TTTGCGCCGAATAAATACCTAAGCTTGTCT3' is added to the 5' end of reverse primer.

2.1.3. Purification of DNA constructs

The amplified PCR products were analysed by agarose gel electrophoresis (1% agarose in TBE Buffer) and purified using the QIAquick PCR purification kit (Qiagen) according to the manufacturer's instructions. Their concentration was determined by measuring the UV at a wavelength of 280nm in a Nanodrop spectrophotometer (Thermo Scientific).

2.1.4. Cloning into expression vectors

2.1.4.1. Restriction Free (RF) cloning

The RF cloning method includes two steps (Figure 2.1). The first step consists of the amplification of the gene of interest using primers containing complementary sequences (25-30 bp) to the flanking sites of integration in the target vector. The products obtained from the first PCR step are then used as mega primers for the second step of amplification. In the second step, the PCR product (from the first step) and the target vector are mixed and following amplification reaction the gene of interest is integrated into the circular vector at a predefined position.

The first PCR step was performed as described in figure 2.1. The PCR products were purified using the Qiaquick PCR purification kit (Qiagen). This DNA was used as mega primers for the RF PCR and 10 µl of the RF PCR product was digested with 20U of *Dpn*I at 37°C for 1 hour. The digested product was then used to transform 90 µl of DH5α (heat shock transformation, see 2.1.4.1.1).

The *Dpn*I endonuclease is specific for methylated or hemimethylated DNA and is used to digest the parental DNA template only. DNA isolated from almost all *E. coli* strains is methylated and therefore susceptible to *Dpn*I digestion. Digestion with *Dpn*I ensures that none of the colonies obtained in subsequent transformation contain the parental plasmid and therefore reduce the transformation background.

Below are shown the reaction mix and the program used for the RF PCR.

Reaction mix

Reagent	Final Concentration
dNTP mix	0.2mM
5X Phusion Polymerase Buffer (NEB)	1X
Megaprimer (first PCR product)	8ng/µl
Template vector	1ng/µl
Phusion DNA polymerase	0.04U/µl
MgCl2	2mM
Milli Q water	Up to 50µl

Here is the protocol used for a PCR reaction:

30sec initial denaturation at 95°C

Then 30 cycles:

30sec denaturation 95°C

1 min annealing 60°C

5 min extension 72°C

Then 7 min final extension

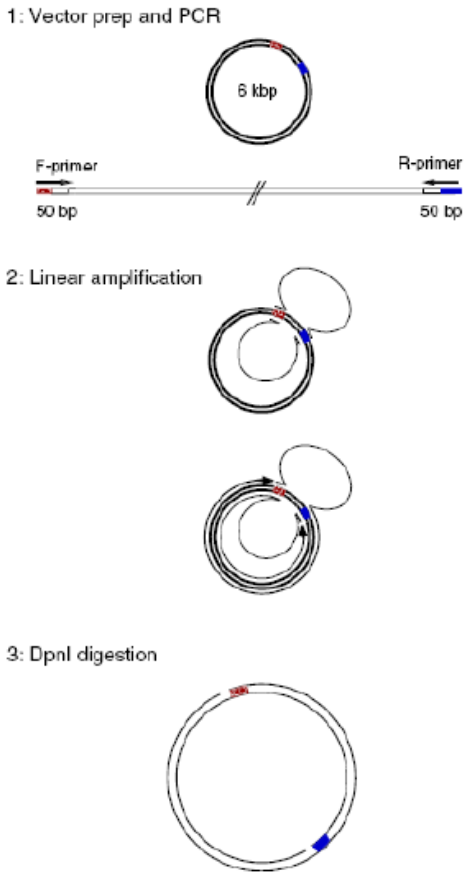

Figure 2.1. Schematic representation of the restriction free (RF) cloning. A circular vector isolated from a dam+ strain has unique priming sites (depicted in red and blue) and is combined with a PCR product encoding the gene of interest flanked by priming sites complementary to those in the vector. 2. The PCR product acts as a primer pair in a linear amplification reaction. Once annealed to the vector, DNA polymerase extends and incorporates the gene into a nicked, circular DNA molecule. 3. The parental vector is digested with DpnI and the double nicked, circular, double-stranded DNA is transformed into a suitable host cell. From van den Ent and Lowe, 2006.

2.1.4.1.1. Transformation of DH5α cells

E. coli DH5α chemically competent cells (Stratagene), prepared according to protocol described in Appendix A2 were used for maintaining the plasmids. For the transformation, the *Dpn*I treated reaction was added to a 90 µl aliquot of DH5α cells. The mixtures were incubated on ice for 30

minutes, heat-shocked at 42°C for 45 seconds and incubated on ice for 2 minutes. LB (Luria Bertani) medium, preheated to 42°C, was added to the mixture tubes to a final volume of 500 µL. The transformation reactions were then incubated for 1 hour at 37°C, shaking at 250 rpm. A volume of 100 µl of each reaction was streaked on LB agar plates containing kanamycin at a final concentration of 50µg/µl.

2.1.4.1.2. Preparation of plasmid minipreps

Single colonies from agar plates were inoculated into 5 ml LB broth medium supplemented with the appropriate antibiotics and the cultures were grown for 16 hours at 37°C with shaking at 200 rpm. The cells were pelleted by centrifugation at 3,500 g for 15 minutes at 4°C. Plasmid DNA was purified from cell pellets using the QIAprep Spin Miniprep kit (Qiagen), according to the manufacturer's protocol, and stored at -20°C. To verify the presence of the insert, vectors were checked by PCR amplification of the DNA construct using the appropriate primers. The PCR program was the same as the one used to amplify the DNA constructs (Touchdown). The amplified PCR products were analysed by agarose gel electrophoresis in 1% agarose in TBE Buffer. The typical yield of plasmid DNA was ~50ng/µl, checked spectrophotometrically measuring the absorbance at a wavelength of 260 nm using a Nanodrop (Thermo Scientific).

2.1.4.2. Ligation Independent Cloning (LIC)

Ligation independent cloning (LIC) (Figure 2.2) is a simple, fast and relatively cheap method to produce expression constructs; in particular we utilized the version that was developed at the Structural Genomic consortium at Oxford (Gileadi et al., 2008).

2.1.4.2.1. Vector preparation

The SGC-made LIC vectors (Gileadi et al., 2008; Appendix A3) all contain the gene encoding for SacB flanked by two *Bsa*I sites (for pNIC-Bsa4, pNIC-Zb and pGTVL2) or two *Bfu*AI sites (for pNIC-CTHF). These cleavage sites are used to linearize the vector, while at the same time removing the SacB gene, allowing negative selection on 5% sucrose. 5 µg of each vector was digested with the corresponding enzyme according to the manufacturer's protocol (New England

Biolabs). Digestion reactions were purified using the Qiagen PCR purification kit and eluted in 50ml elution buffer (EB) provided by the kit.

The digested and purified vector was further treated with T4 DNA polymerase (New England Biolabs) in the presence of dCTP (pNIC-CTHF) or dGTP (pNIC-Bsa4). Because of the 3'→5' exonuclease activity of the polymerase the bases are removed from both 3'-ends until the first guanine (G) or cytosine (C) residue is reached, respectively.

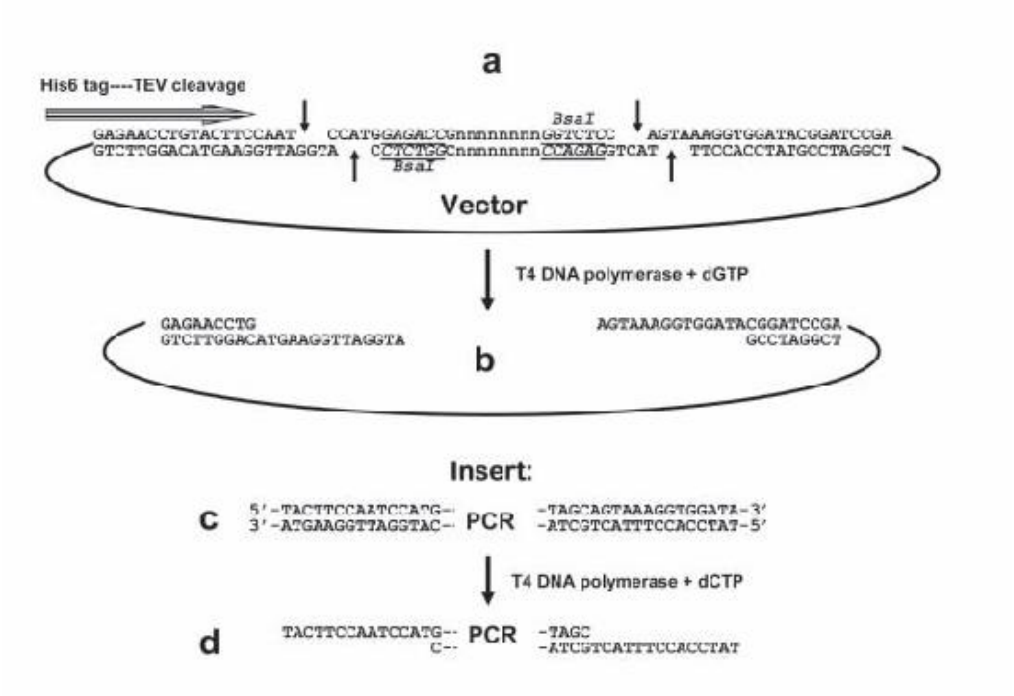

Figure 2.2. Schematic view of LIC cloning (for BsaI treatable vectors) A. The cloning vector is cleaved by the non palindromic enzyme *BsaI*. B The resulting ends are trimmed by T4 DNA polymerase and dGTP. C. PCR fragments are generated with the indicated extensions. D. The PCR fragments are trimmed by T4 polymerase and dCTP, generating single-stranded tails complementary to the vector tails. Adapted from Gileadi *et al*. 2008.

Here is the protocol for theT4 DNA polymerase treatment of vector:

Reagent	**Final concentration**
Digested plasmid	10-50ng/µl
NEB2 buffer	1X
dGTP or dCTP	2.5mM
BSA (NEB)	0.1µg/µl
DTT	0.5mM
T4 DNA Polymerase (NEB)	0.15U/µl
Milli Q water	Up to 100µl

Reactions were incubated at 22°C for 30 minutes then at 75°C for 20 minutes in a thermocycler preheated to 90°C.

2.1.4.2.2. Insert preparation

RECQ4 gene fragments were amplified using a Touchdown PCR program and purified with the Qiagen PCR purification kit, eluted in 30µl of elution buffer (EB). In the next step, the PCR product was incubated with T4 DNA polymerase in the presence of dCTP or dGTP.

Here is the protocol for the T4 treatment of PCR products:

Reagent	**Final concentration**
PCR product	10-50ng/µl
NEB2 buffer	1X
dCTP or dGTP	2.5mM
BSA (NEB)	0.5µg/µl
DTT	0.5mM
NEB T4 DNA polymerase	0.15U/µl
Milli Q water	Up to 10µl

Reactions were incubated at 22°C for 30 minutes then at 75°C for 20 minutes in a thermocycler preheated to 90°C.

2.1.4.2.3. Ligation and transformation

To anneal the treated vector and the treated insert 1µl of treated vector was mixed with 3µl of treated insert (4µL if low DNA concentration). Annealing reactions were incubated 30 minutes at 22°C then transferred to ice. 30µl of competent DH5α cells were added to the ligation mixture, incubated on ice 30 minutes and heat shock transformation performed. Cells transformed were plated onto small petri dishes of LB containing kanamycin and 5% sterilized sucrose.

To further check for the presence of the insert, (section 2.1.4.1.2) single colonies grown overnight at 37°C in LB plus kanamycin, were used for plasmid extraction using the Qiagen Miniprep Kit. Purified cloned vectors were then checked with the specific primers for each DNA fragment using a Touchdown PCR cycle. All the constructs were sequenced to verify the insertion of the correct sequence and to check for PCR mutation or rearrangements.

2.1.5. Site directed mutagenesis

In order to verify the presence of the Zn knuckle and the RQC domain, site-directed mutagenesis was used to mutate critical residues into alanine in the pETM-30 and pET-SUMO/CAT vector containing RecQ4 fragment 335-427 and 445-1112, respectively. All reactions were performed using double-stranded plasmid DNA as a template, and a pair of complementary oligonucleotide primers, incorporating the desired mutation. Mutants generated are listed in Table 2.3.

Name of mutant	Template DNA plasmid	Primer forward (5'→3')	Primer reverse (5'→3')
C403A/N406A	pETM-30 (RecQ4 335-427)	CCAAGGAGTCTGCGTTCCTGG CGGAGCAGTTCGATCACTGGG CAGCCC	GGGCTGCCCAGTGATCGAACTG CTCCGCCAGGAACGCAGACTCC TTGG
H411A/C416A	pETM-30 (RecQ4 335-427)	GAGCAGTTCGATGCGTGGGCA GCCCAGGCGCCCCGGCCAGCA AGTGAG	CTCACTTGCTGGCCGGGGCGCCT GGGCTGCCCACGCATCGAACTG CTC
N406C	pETM-30 (RecQ4 335-427)	GAGTCTTGTTTCCTGTGTGAGC AGTTCGATCAC	GTGATCGAACTGCTCACACAGG AAACAAGACTC
C853A/C855A	pET-SUMO/CAT (445-1112)	CGCGTGTTCCCAGCCGCGACC GCGACCTGCACCAGGCCG	CGGCCTGGTGCAGGTCGCGGTC GCGGCTGGGAACACGCG
C897A	pET-SUMO/CAT (445-1112)	GACCCAGAAGGGTCGCGATGG GCCATGAGCGG	CCGCTCATGGCCCATCGCGACC CTTCTGGGTC
C925A	pET-SUMO/CAT (445-1112)	CATCGAGACTTTGCTGGCGTA CCTGGAGCTGCACC	GGTGCAGCTCCAGGTACGCCAG CAAAGTCTCGATG
C945A	pET-SUMO/CAT (445-1112)	GACCACCTATACCCATGCGCG TCTGAACTGCCCTG	CAGGGCAGTTCAGACGCGCATG GGTATAGGTGGTC
C963A	pET-SUMO/CAT (445-1112)	CCTGGCCCACAGGGCGCCCCC TTTGGCTG	CAGCCAAAGGGGGCGCCCTGTG GGCCAGG
F1077A	pET-SUMO/CAT (445-1112)	CGCAGAACCTTCCAGGCCGCG CACAGCGTAGCC	GGCTACGCTGTGCGCGGCCTGG AAGGTTCTGCG
K508A	pET-SUMO/CAT (445-1112)	CTGCCTACAGGTGCCGGCGCG TCCCTGTGCTACCAGCTC	GAGCTGGTAGCACAGGGACGCG CCGGCACCTGTAGGCAG

Table 2.3. List of primers used for site directed mutagenesis. In red are showed the mutations introduced.

Mutated vectors were generated in a 50µl reaction containing 5-50 ng of template DNA, 125 ng forward primer, 125 ng reverse primer, 10µl Pfu Turbo polymerase (2.5U/µl) and its 10X buffer and dNTP mix using the following PCR cycles:

30sec initial denaturation at 95°C

Then 18 cycles:

30 sec denaturation 95°C

1 min annealing 55°C

14 min extension 68°C (2 min/kb plasmid)

∞ at 4°C

Following temperature cycling, the reactions were digested with 1 µl of the *Dpn*I (stock 20,000U/ml) restriction enzyme at 37ºC, for 1.5 hr, so as to digest the methylated parental DNA. A 4 µl aliquot from the reaction mixture was then used to transform 45 µl DH5α competent cells (Stratagene) and the cell mixture was streaked on LB agar plates, containing kanamycin. After overnight incubation at 37ºC, preparation of plasmid DNA from isolated colonies proceeded as previously.

2.2. Proteins expression and purification

Vectors containing RecQ4 DNA fragments were mainly transformed (heat shock transformation, previously described) into Rosetta 2(DE3) strain. Other *E. coli* strain cells were tested to increase the chances of obtaining soluble protein, such as BL21 (DE3) (Stratagene), B834 (Novagen) and ArcticExpress Competent Cells (Agilent).

Rosetta 2(DE3) strain is derived from BL21 strain and is designed to enhance the IPTG induced expression of eukaryotic proteins that contain codons rarely used in E.coli. This strain provides the tRNAs for seven rare codons in a plasmid (pRARE) containing the chloramphenicol resistance.

BL21 (DE3) is also based on BL21 strain with a T7 RNA polymerase that is induced by IPTG induction. BL21 is a basic E.coli strain used for protein expression; it lacks both *lon* protease and *ompT* protease that can degrade the expressed proteins.

B834 is the parental strain of BL21 and is a methionine auxotroph, usually used to label proteins with selenomethionine. It also is deficient in *lon* and *ompT* proteases.

ArcticExpresscompetent cells provide an in vivo approach to increase the yield of soluble protein. Expressing protein at low temperature is a commonly used method for increasing the yield of soluble protein but E.coli chaperonins are relatively less active at such temperature. To overcome this obstacle ArcticExpress cells are engineered to co-express the cold-adapted chaperonins Cpn10 and Cpn60 from the psychrophilic bacterium, *Oleispira antarctica.* The recombinant proteins are expressed using T7 promoter and the tRNAs coding the rare codons.

2.2.1. Small scale test expression and solubility assays

Before attempting large-scale expression of any newly synthesized DNA construct, the expression of the target protein was verified by the analysis of whole-cell extracts, as described below. Once expression of the target protein was verified, determination of the protein solubility and optimization of growth conditions were carried out using small scale batch purification.

Multiple colonies of *E. coli* cells harbouring the plasmid being tested were used to inoculate 1 mL or 10 mL or 100 mL of LB or TB media (containing appropriate antibiotics), in 15 ml or 50 mL falcon tubes or 250 mL flasks respectively, to allow a sufficient aeration during growth.

The cultures were grown at 37ºC, shaking at 200 rpm until reaching an optical density (OD) of 0.6-0.7 at 600 nm for the cells in LB broth, or 2-2.5 OD for those in TB broth. Protein expression was induced by addition of IPTG at different concentration ranging from 0.1 to 1 mM, at different temperature and times (37°C for 5 hours, 25°C overnight/16 hours, 18°C overnight/16 hours). Autoinduction media (see Appendix A1) were also tested at 17°C for 16 hours. Aliquots were removed from the culture prior to addition of IPTG (pre-induction sample) and prior to harvesting (post-induction sample). The aliquots were pelleted by centrifugation at 16,000 xg for 2 min at 4ºC, the supernatant was removed and each cell pellet was resuspended in 20 µl 2x SDS Sample Buffer and heated at 95ºC for 10 min to fully lyse the cells and denature the proteins. The pre-induced and post-induced SDS-whole cell lysates were analysed by denaturing SDS-polyacrylamide gel electrophoresis (Section 2.3.3), followed by Instant blue staining to verify expression of the target recombinant proteins by comparing the pre- and post-induction whole-cell extracts from each culture.

Cell pellets were thawed on ice and resuspended in 0.5 ml Nickel Buffer A: 50 mM Hepes pH 7.5, 500 mM NaCl, 10 mM imidazole, 10% glycerol, 1mM TCEP (tris(2-

carboxyethyl)phosphine) supplemented with 1 mg/ml lysozyme and 1 mM PMSF protease inhibitor and the suspension was incubated on a rotating wheel in the cold room at 4ºC for 1 hour. The lysates were centrifuged for 30 min at 26,000 xg to remove cellular debris and the supernatant (soluble fraction) was transferred to a fresh 1.5 ml microcentrifuge tube.

Ni-NTA fastflow resin (Qiagen) was used for the analysis of His-tagged target proteins. The resins were prepared according to the manufacturer's instructions: pre-washed 3 times in water and washed twice in Nickel Buffer A prior to use. A 50 µl aliquot of resin was added to the supernatant and the tube was incubated with gentle rotation at 4°C for 45 min. The suspension was centrifuged at 500 xg for 5 min to sediment the resin. The supernatant was removed and the resin was washed with 1 ml Nickel Buffer A by resuspending with a pipette, centrifuging and discarding the supernatant, for a total of three washes. The third wash was performed with Nickel Buffer A supplemented with 50 mM imidazole, to reduce unspecific binding of proteins to the resin. 25 µl of 2x SDS Sample Buffer were added to the sedimented resin and the suspension was heated at 95ºC for 10 min to elute the bound proteins from the resin. The sample was centrifuged at 16,000 xg for 1 minute and 15-20 µl of the resin was analysed by SDS-PAGE with Instant blue staining. Once the optimal conditions for the each construct were found, large scale expression experiments (i.e. 6-10 liters culture) were performed.

2.2.2. Large scale expression and purification

N-terminal region

GST-tagged wild type and mutant proteins were expressed in Rosetta 2(DE3) cells in Terrific Broth (TB, Sigma) at 18° overnight, following induction with 0.2 mM IPTG in the presence of 0.1 mM ZnSO4. The cells were harvested at 5000 rpm for 30 minutes at 4ºC and pellets were frozen at -80ºC. The cell pellet from 1 L of culture was thawed and resuspended in 40 mL (for cells grown in TB) of affinity binding buffer (Nickel Buffer A), supplemented with 2 mM of the protease inhibitor AEBSF (4-(2-aminoethyl) benzenesulfonyl fluoride, Sigma) and/or one tablet of protease inhibitors cocktail from Roche. The resuspended solution was incubated on ice for 20 minutes, stirring with a magnetic bar, in order to allow the lysis. The suspension was then sonicated on ice 15 minutes with short pulses of 15 seconds each followed by 45 seconds pause, at 40% amplitude (Soniprep 150). The cell lysates were clarified by centrifugation for 45

minutes at 26,000 g and the supernatant filtered through a 0.22 mm cut-off filter before applying to affinity columns. Proteins were purified by Nickel-NTA fastflow resin, washed with high salt concentration to eliminate DNA contamination and run on a size exclusion chromatography (Superdex-200) in 250 mM NaCl, 20 mM Tris pH 7.5, 5% glycerol and 5 mM β-mercaptoethanol.

Catalytic core (HelRQC)

The wild type and mutant proteins were expressed in Rosetta 2(DE3) cells using auto-induction method at 17°C for 48 hours and harvested by centrifugation. The cells were harvested at 5000 rpm for 30 minutes at 4ºC and pellets were frozen at -80ºC. The cell pellet from 1 L of culture was thawed and resuspended in 40 mL (for cells grown in TB) of affinity binding buffer (Nickel Buffer A), supplemented with 2 mM of the protease inhibitor AEBSF (4-(2-aminoethyl) benzenesulfonyl fluoride, Sigma) and/or one tablet of protease inhibitors cocktail from Roche. The resuspended solution was incubated on ice for 20 minutes, stirring with a magnetic bar, in order to allow the lysis. The suspension was then sonicated on ice 15 minutes with short pulses of 15 seconds each followed by 45 seconds pause, at 40% amplitude (Soniprep 150). The cell lysates were clarified by centrifugation for 45 minutes at 26,000 g and the supernatant filtered through a 0.22 mm cut-off filter before applying to affinity columns. The clarified lysate was loaded onto HisTrap FF 5ml column pre-equilibrated in Nickel Buffer A, using ÄKTApurifier instrument (GE Healthcare). The column was washed with 10 column volumes (CV) of Nickel Buffer A with 25 mM Imidazole and 2M NaCl (to remove the DNA contamination), and the protein was eluted with a gradient of 1ml/min flow to reach 250 mM Imidazole in 30mins. Peak fractions containing protein of interest were pooled; the concentration of Imidazole was brought back to 10mM by diluting the sample with Buffer A and kept under overnight cleavage with SUMO protease at a w/w ratio of 1:500 at 4°C. Cleaved reaction was passed through a HisTrap FF 5ml column pre-equilibrated in Nickel Buffer A. The flow-through was collected and further diluted 10 times by Heparin Buffer A (50 mM Hepes pH 7.5, 1 mM TCEP and 5% glycerol) to reduce the NaCl concentration to 50mM. This sample was loaded onto HiTrap Heparin 5ml column equilibrated in Heparin Buffer A with 50mM NaCl. The column was washed with 10 column volumes (CV) of Heparin Buffer A with 50mM NaCl, and the protein was eluted with a

gradient of 1ml/min flow to reach 1M NaCl in 30mins. Peak fractions containing protein of interest were pooled and concentrated using Amicon® Ultra-4 centrifugal filter unit with 30 kDa cutoff (Millipore) to further purify it using size exclusion chromatography (Superdex-200) in 20 mM Tris pH 7.5, 250 mM NaCl, 5% glycerol and 5 mM β-mercaptoethanol. Protein concentration was determined by using Nanodrop Instrument (Thermo Scientific), and protein purity was analyzed as described in paragraphs 2.3.1, 2.3.2 and 2.3.3.

2.2.3. Removal of tags from proteins

Proteins expressed from pET-SUMO/CAT expression vectors were subjected to removal of SUMO tag by SUMO Protease. SUMO Protease is a highly active cysteinyl protease which is a recombinant fragment of Ulp1 (Ubl-specific protease 1) from *Saccharomyces cerevisiae* (Mossessova and Lima 2000). SUMO Protease cleaves in a highly specific manner, recognizing the tertiary structure of the ubiquitin-like (UBL) protein, SUMO (Mossessova and Lima 2000) rather than an amino acid sequence. The optimal temperature for cleavage is 30°C; however, the enzyme is active over wide ranges of temperature and pH (pH 7.0-9.0). Samples were treated for 4 hours with SUMO protease at 1:500 w/w ratio at 4°C in the following buffer: 50 mM Tris-HCl (pH 8.0), β-mercaptoethanol, 250 mM NaCl. Following digestion, SUMO Protease and His-SUMO tag were removed from the cleavage reaction by Nickel affinity chromatography using the polyhistidine tag at the N-terminus of the protease. Cleavage was checked by SDS-PAGE gel.

2.3. Biochemical characterisation of protein

2.3.1. Determination of protein concentration

Protein concentrations were determined spectrophotometrically by measuring the absorbance at 280 nm and applying the Lambert-Beer equation using the theoretical extinction coefficient for each protein. Extinction coefficients were determined using the ProtParam tool at the Expasy server (http://ca.expasy.org/tools/protparam.html).

2.3.2. Determination of protein size by size-exclusion chromatography

For the determination of the size and molecular weight of the protein, a standard calibration curve was plotted as follows: for each calibration standard, the logarithm of its known molecular weight log (MW) was plotted against its normalized elution volume (Ve/Vo). Using the least squares method, the following linear equation was calculated:

$$log(MW) = m(Ve/Vo) + b$$

where *m* is the slope and *b* is the intercept.

The equation was then used for the determination of the log (MW) of the protein of interest based on its elution volume (Ve) obtained by size-exclusion chromatography with the Superdex-200 or Superdex-75 column.

2.3.3. SDS-PAGE

Protein samples were analysed by electrophoresis under denaturing, reducing conditions using discontinuous SDS-polyacrylamide gels (4% for the stacking gel and 10%, 12% or 15% for the separating gel). Samples were prepared by addition of 4x SDS sample buffer and heating at 95ºC for 5-10 min. Electrophoresis was carried out at a constant voltage (200V) for 45-55 min in 1x Tris-glycine buffer. Gels were stained with Instant Blue, then stored in water.

2.3.4. Metal binding analysis

The known concentration of purified wild type and mutant proteins were sent for inductively coupled plasma - atomic emission spectroscopy (ICP-AES) to detect the presence of Zn and Fe metals. ICP-AES uses inductively coupled plasma to excite the metal elements which later emits radiation of characteristic wavelength. The intensity of the radiation is the measure of the concentration of the element in the sample.

2.3.5. Thermofluor assays

To determine the protein stability, thermofluor assay was carried out using CFX96 Real Time System C1000 Touch Thermal Cycler instrument (Bio-rad). Protein stability measurements were performed in buffer containing 20 mM Tris pH 7.5, 250 mM NaCl, 5% glycerol and 5 mM β-mercaptoethanol. The reaction mixture (25 µl) contained 0.5 µM of protein and 20x SYPRO

Orange (Invitrogen) in the above mentioned buffer. Heat denaturing curves were observed within the temperature range of 20 °C to 80 °C, with 1.8 °C/min rate of temperature change, collecting data every 10sec. Obtained data was analysed using GraphPad-Prism software.

2.3.6. Preparation of oligonucleotides used for nucleic acid binding assays

All oligonucleotides were chemically synthesized and purified by reverse-phase high pressure liquid chromatography (RP-HPLC) (Sigma-Aldrich, Suffolk, UK). Each nucleotide was then resuspended in Tris–EDTA (TE) buffer (10 mM Tris–HCl, pH 7.5, 1 mM EDTA, pH 8.0), supplemented with RNase inhibitor (NEB) in case of RNA substrates. Oligonucleotide sequences used in this work are reported in Table 3.4 and Table 3.8.

For radioactive EMSA, a single oligonucleotide was 5′-end-labeled with [γ-^{32}P] ATP using T4 polynucleotide kinase. The kinase reaction was performed in PNK buffer (70 mM Tris–HCl, pH 7.6, 10 mM $MgCl_2$, 5 mM dithiothreitol) at 37°C for 60 min. For fork and double-stranded probes, the [γ-^{32}P] ATP-labeled oligonucleotides were then annealed to a 1.6-fold excess of the unlabeled complementary strands in annealing buffer (10 mM Tris–HCl, pH 7.5, 50 mM NaCl) by heating at 95°C for 8 min and then cooling slowly to room temperature. The forked substrates comprised 22 paired and 15 unpaired bases. The purification of the duplex substrates was performed using Micro Bio-Spin columns (Bio-Rad).

For fluorescent assays the oligonucleotides were purchased with the label fluorophores 6FAM and BHQ1 at the 5’ and 3’ end, respectively.

2.3.7. Electrophoretic Mobility Shift Assays

The experiments were performed by incubating increasing concentrations of purified recombinant wild type and mutant RecQ4 fragments with [γ-^{32}P] labelled at a fixed final concentration of 10 nM in a 20 µl reaction mixture containing 20 mM Tris-HCl pH 7.5, 2 mM $MgCl_2$, 50 mM NaCl, 5% glycerol, 1 mM DTT, 5 mM $ZnCl_2$ and 0.1 mg/ml BSA. EMSA in the presence of EDTA were carried out in the same conditions at a fixed protein concentration of 500 nM, and increasing amounts of EDTA (0-50 mM). After incubation for 30 min at room temperature, the reaction products were separated on a 6% non denaturing polyacrylamide gel run at 4°C in TBE buffer. [γ-^{32}P] labelled nucleic acid fragments were detected by

autoradiography (Cyclon, GE Healthcare). The quantification of protein-nucleic acid complexes was performed with ImageQuant image analysis software (GE Healthcare). The apparent equilibrium dissociation constants (K_D) were determined using a single-site binding model in Prism (using the GraphPad software) from the mean of at least three independent experiments.

2.3.8. Helicase assay

The helicase activity was measured by using fluorescence resonance energy transfer (FRET) in which when the double stranded substrate is unwound by helicase, the fluorescence of the fluorophore is emitted (Hidenori et al., 2010). The fluorescence helicase assay was performed in 20 mM Tris–HCl pH 7.5, 5 mM MgCl2, 50 mM KCl, 8 mM DTT, 0.1 mg/ml BSA and 5% Glycerol with 5 nM dsDNA substrate (F1:F2 and B1:B2, Table 2.4), 3 mM ATP and 62.5 nM Capture strand in 30 µl of reaction volume. The unwinding reaction was started by incubating increasing concentrations (0 to 160nM) of the purified proteins in the reaction mixture at 37°C for 30 min. The fluorescence intensity was recorded using Infinite F200 PRO TECAN instrument. To have a measurement of 100% unwinding the reaction was incubated at 95°C and measured. The assay was done in triplicate. The percentage of unwinding was calculated and plotted using GraphPad-Prism software.

2.3.9. ATPase assay

Cayman's Malachite Green Assay Kit was used to determine the ATPase activity by measuring the released inorganic phosphate during ATP hydrolysis (Chan et al., 1986; Huang et al., 2002). The method is based on the change in absorbance (*A*620 – A640) of the complex formed between malachite green (MG) molybdate and free orthophosphate under acidic conditions (molybdophosphoric acid complex). The assay was carried out in 50 µl reaction volumes, containing 20 mM Tris–HCl pH 7.5, 5 mM MgCl2, 50 mM KCl, 8 mM DTT, 0.1 mg/ml BSA and 5% Glycerol with 500 nM purified proteins, 3 mM ATP and 500 nM of substrates (F3, F3:F4 and B3:B4, Table 2.4). The ATPase reactions were incubated at 37 °C for 30 min. After incubation, 5 µl of MG acidic solution was added and incubated at room temperature for 10mins followed by addition of 15 µl MG molybdenum reagent and incubation at room temperature for 5 min. The absorption at 620 nm was then measured using Lambda25 UV/VIS Spectrometer

(Perkin Elmer). The concentrations of inorganic phosphate were determined from a standard curve of A_{620} nm *versus* known phosphate concentrations.

2.4. Structural characterisation of protein

2.4.1. CD spectroscopy

The circular dichroism (CD) spectra were recorded at 25°C on a Jasco J-810 Spetropolarimeter at wavelength 190-260 nm, in 0.1 cm quartz cells, band width 1 nm, response 1 sec, data pitch 0.1 nm and scanning speed 20 nm/min. Three scans were done for each sample and averaged. Every spectrum was background (buffer) subtracted.

The concentration of the two synthetic peptides corresponding to human and *Xenopus* RecQ4 Zn knuckles (pep-hZnK and pep-xZnK) was 20 μM each. Both peptides were dissolved in 50 mM phosphate buffer pH 7.4, 0.1 mM TCEP. Where indicated, Zn^{2+}, DNA or RNA were added in an equimolar ratio.

2.4.2. Peptide synthesis

The 25 amino acid long peptides corresponding to the Zn-knuckle at the N-terminus of human and *Xenopus* RecQ4 were synthesized on an Applied Biosystem machine, purified by reverse phase HPLC (RP-HPLC) to a degree of purity was confirmed by MALDI Tof-Tof analysis to be higher than 95%. These peptides were later used for structural characterization using nuclear magnetic resonance (NMR).

2.4.3. Nuclear Magnetic Resonance (NMR) experiment

2.4.3.1. NMR spectroscopy and resonance assignments

NMR experiments were performed at 298 °K on a Bruker Avance 600 MHz spectrometer equipped with inverse triple resonance cryoprobe (TCI) and pulsed field gradients. Data were processed with Topspin 2.0 (Bruker) and analyzed using CCPNmr 2.1.5 (Vranken *et al*. 2005). 1D 1H spectra were acquired on 0.3 mM human pep-hZnK and 0.3 mM *Xenopus* pep-xZnK samples in 20 mM NaH_2PO_4/Na_2HPO_4 pH 6.3, 150 mM NaCl, 4 mM DTT, 0.3 mM DSS, 10% (v/v) D2O, with or without 0.35 mM $ZnCl_2$. 1D 1H spectra were also acquired on 0.36 mM pep-hZnK sample in 50 mM HEPES pH 6.8, 150 mM NaCl, 4 mM DTT, 0.3 mM DSS, 10% D2O,

with or without 0.35 mM $ZnCl_2$, and 0.09 mM pep-hZnK in 50 mM HEPES pH 6.8, 150 mM NaCl, 4 mM DTT, 0.09 mM $ZnCl_2$, 0.3 mM DSS, 10% D2O with or without 0.09 mM ssDNA (A1, Table 2.4).

1H, 15N and 13C backbone resonances of *Xenopus* Zn knuckle and 1H side chains were assigned through 2D 1H-13C HSQC, 1H-15N HSQC, 1H-1H TOCSY (mixing time 60 ms) and 1H-1H NOESY (mixing time 100, 150 and 200 ms) spectra, which were acquired on 1.1 mM pep-xZnK samples in 20 mM NaH_2PO_4/Na_2HPO_4 pH 6.3, 150 mM NaCl, 4 mM DTT, 1.25 mM ZnCl2, 0.3 mM DSS, 10% (v/v) D2O.

2.4.3.2. Structure calculation

The solution structure of the *Xenopus* Zn knuckle (pep-xZnK) was calculated using ARIA 2.3.1 (Rieping et al., 2007) in combination with CNS (Brunger, 2007), based on the experimentally derived restraints. In particular, proton–proton distances were obtained from a 2D 1H-1H NOESY spectrum (mixing time 200 ms). Φ/Ψ restraints were obtained from backbone chemical shifts using TALOS+ (Shen et al., 2009). The 2D 1H-1H NOESY spectrum was manually assigned and calibrated by ARIA. Eight ARIA iterations were performed, 100 structures were computed in the last iteration and ARIA default water refinement was performed on the 15 best structures from the final round. Initial structures were calculated without zinc ion restraints to verify the position and the geometry of the metal ion ligands, so that the residues involved in zinc binding could be identified in an unbiased manner. Several (Nuclear Overhauser effect) NOEs were observed between metal coordinating residues, clearly revealing the tetrahedral coordination of the ligands around the zinc. Once metal ligands were unequivocally identified, the geometry of the Zn^{2+} coordination was constrained in the final ARIA calculations via covalent bonds and angles in the CNS parameters; the tetrahedral angles and distances for Zn^{2+} coordinating residues were maintained also after water refinement. Structural quality was assessed using Procheck-NMR (Laskowski et al., 1996) and molecular images were generated by PyMOL (http://pymol.org/). The family of the 15 lowest energy structures for *Xenopus* RecQ4 Zn knuckle has been deposited in the PDB (Protein Data Bank) with the accession code 2MPJ. Chemical shift and restraints used in the structure calculations have been deposited in BioMagResBank (19986 code).

2.4.4. Small angle X-ray scattering (SAXS)

2.4.4.1. Sample preparation and Data aquisition

The first sample (HelRQC1) was purified and concentrated using Amicon® Ultra-4 centrifugal filter unit with a cut-off of 30 kDa (Millipore) until 1 mg/ml, 1.5 mg/ml and 2 mg/ml concentrations were reached. The three concentrations were analyzed at the Austrian Small Angle X-ray Scattering (AustroSAXS) beamline in ELETTRA (Amenitsch et al., 1997). The flow-through buffer was used as the blank for the measurements together with empty capillary and water. Data collection was done at λ=1.54Å with the sample to detector (Image Plate Mar300) distance set to 1102.43mm in order to resolve the momentum transfer, ($q=4\pi\sin\theta/\lambda$) in the range from 0.011 to 0.46 $Å^{-1}$. Measurements were done using a quartz capillary of 1.5 mm diameter, at 8°C and with 60s/image of exposure time.

Sample two (HelRQC2) was purified by removing the DNA contamination using 3M NaCl, as described in section 2.2.2. At this step, the third purified protein sample (HelRQC2+forkDNA) was incubated with an equimolar amount of forkDNA (F3:F4, 37bp) for 60min at room temperature. Both HelRQC2 and HelRQC2+forkDNA were later concentrated using Amicon® Ultra-4 centrifugal filter unit with a cut-off of 30 kDa (Millipore) until 1 mg/ml, 1.5 mg/ml and 2 mg/ml concentrations were reached. SAXS data were collected as described above to determine the overall dimensions of the particles (radius of gyration, Rg, and maximum size, Dmax) at AustroSAXS beamline. The sample to detector (Image Plate Mar300) distance was set to 921.62mm (q-range between 0.015 and 0.41 $Å^{-1}$) and the exposure time was 60s/image.

2.4.4.2. Data analysis

After calibration with silver behenate, raw data were radially averaged using Fit2D (A. P. Hammersley, ESRF 1987-2004). The scattering contributions of empty, water and buffers were subtracted from the total scattering intensity of all protein curves using IgorPro (Wavemetrics, Lake Oswego, OR). Data were further normalized taking into account the transmission and the detector parameters. All measurements for the same concentration were averaged and the scattering curves were merged when necessary. Guinier's approximation (Guinier and Fournet, 1955) was used to check for particle aggregation and radiation damage in the sample. Structural parameters (R_g, $I_{(0)}$ and Porod's volume) were calculated from the

experimental curves using Datrg and Datporod (Petoukhov et al., 2007). For this calculation, absolute scale was determined using the reference curve collected from Bovine Serum Albumin (BSA) at 5mg/mL. The maximum length of the model (D_{max}) was estimated using GNOM (Svergun, 1992), getting the best trends for the P(r) function according to the expected particle shape. GNOM reads 1D scattering curves collected for a mono-disperse system and estimates the particle distance distribution function P(r). OLIGOMER (Konarev et al., 2003) was used to exclude oligomerization. Several successful cycles ab-initio shape determination using DAMMIF (Franke and Svergun, 2009) were run on the three scattering curves. DAMAVER (Volkov and Svergun, 2003) was used to average the multiple reconstructions by superimposing them using SUPCOMB (Kozin and Svergun, 2001). CRYSOL (Svergun et al., 1995) was used to confirm the good agreement between the experimental data with models and the corresponding X-ray structures, obtaining their theoretical scattering curve and structural parameters.

3. Results and Discussion

3.1 Protein expression and purification

3.1.1 Cloning strategy

Based on the bioinformatic analysis of RecQ4 sequences previously carried out in our lab (Marino et al., 2013) we have designed a number of fragments encompassing the different domains of human RecQ4 (Figure 3.1). The main aim of the work was to identify autonomously folding units that would provide folded, stable, soluble protein suitable for structural biology and biochemistry, focusing particularly on the catalytic region.

These fragments were amplified by PCR and cloned into a variety of expression vectors to produce variants of the protein with different fusion tags (His-tagged, His-Flag-tagged, His-GST-tagged or His-SUMO-tagged proteins). The vectors used and their characteristics are listed in Table 2.1. This strategy increases the chances of obtaining soluble proteins as different tags can have an effect on the solubility of the protein. Also, for each fragment the domain boundaries were fine-tuned, with multiple start and end points (Table 2.2 and Figure 3.1) so as to increase the chances of expressing the correctly folded proteins.

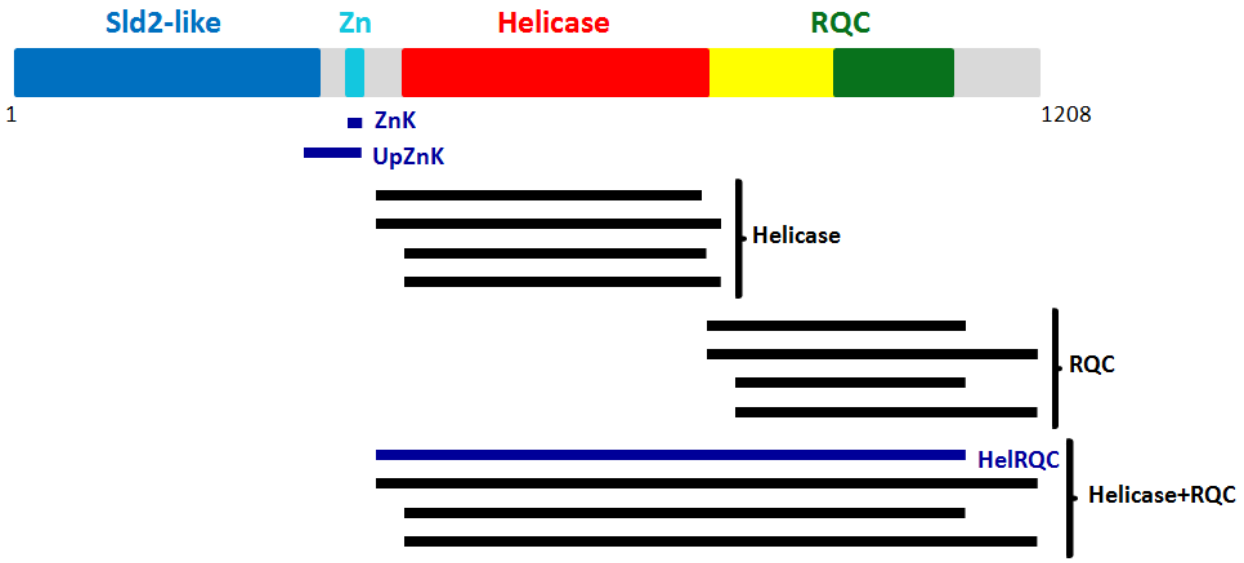

Figure 3.1. Schematic diagram of predicted domain organisation of human RecQ4 (Marino et al., 2013). The RQC domain includes a Zn-binding region (in yellow) and a WH domain (in green). Below are shown the boundaries of the corresponding fragments cloned during this project. The constructs in dark blue are studied in this project.

Two different cloning strategies were employed, based on the vector chosen. For the vectors provided by the Oxford Structural Genomics Consortium (SGC) a ligation-independent cloning strategy devised by the SGC was employed (Gileadi et al., 2008, Section 2.1.4.2). For the other two vectors we used an alternative method called restriction-free cloning as described by Unger, based on two rounds of PCR reactions (Unger et al., 2010, Section 2.1.4.1). These strategies were preferred to the traditional restriction-based cloning method as they make it easier to implement the parallel cloning of numerous fragments in multiple vectors. For each construct, a variety of bacterial strains and expression conditions (varying temperature, IPTG concentration, broth and method of induction) were explored.

Unfortunately, most of the constructs produced insoluble protein. After numerous expression and purification attempts, purification protocols were optimized for three constructs shown in figure 3.1, one including the Zn knuckle only (ZnK), one including the additional Sld2 homologous region and Zn knuckle (UpZnK) and one including the catalytic core of helicase and RQC domain (HelRQC). Two additional constructs encompassing the same regions as ZnK and UpZnK were generated starting from the Xenopus sequence (see Section 3.2). Table 3.1 summarizes the cloning, expression and purification trials of all the constructs in several expression vectors. The expression and purification of the three constructs are briefly summarized in further sections.

Construct name	pETM-30	pNIC-CTHF	pNIC-Bsa4	pET-SUMO/CAT
ZnK (394-427)	**XXX**			
UpZnK (335-427)	**XXX**			
Helicase1 (445-819)		X	X	XX
Helicase2 (445-852)		X	X	XX
Helicase3 (475-819)		X		X
Helicase4 (475-852)		X	X	X
RQC1 (820-1112)			X	X
RQC2 (820-1208)			X	X
RQC3 (895-1112)		X		X
RQC4 (895-1208)		X		X
HelRQC (445-1112)	XX	X	X	**XXX**
HelRQC2 (445-1208)	XX		X	XX
HelRQC3 (475-1112)	XX	X	X	
HelRQC4 (475-1208)	XX	X	X	

Table 3.1. Summary of the behavior of RecQ4 constructs. "X" means that the clones have been produced; "XX" means soluble protein obtained but at lower yields; "XXX" means that the protocol has been optimized and was successful in generating mg amounts of protein. In red are the three constructs which have been the focus of my subsequent work.

3.1.2. Expression and purification of human and *Xenopus* N-terminal regions

In order to evaluate the role of human (hZnK) and *Xenopus* (xZnK) RecQ4 Zn-knuckles (Figure 1.7) in nucleic acid binding, they were cloned in a pETM-30 vector to produce 6His-GST tagged proteins. These constructs were expressed in *E. coli* in TB at 18°C overnight, following induction with 0.2 mM IPTG in the presence of 0.1 mM $ZnSO_4$ and purified by Nickel resin, GSTrap affinity column and injected on a gel filtration column Superdex 200 (Figure 3.2).

To further dissect the role of the additional Sld2 homologous region upstream of the Zn knuckles (Figure 1.6), constructs including both the Zn knuckle and the upstream region for human (hUpZnK, residues 335 to 427) and *Xenopus* (xUpZnK, residues 555 to 638) RecQ4 were designed and cloned in pETM-30 vector containing a 6His-GST tag. These constructs were expressed in *E. coli* in TB at 18°C overnight, following induction with 0.2 mM IPTG in the presence of 0.1 mM $ZnSO_4$ and purified by Nickel resin, GSTrap affinity column and injected on a gel filtration column Superdex 200 (Figure 3.2). The proteins expressed well and could be purified to a reasonable extent; a small amount of proteolysis is due to the intrinsic flexible natures of some of regions involved

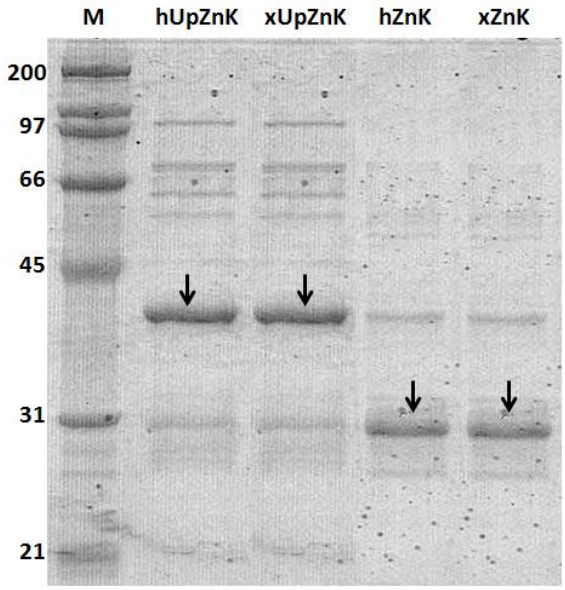

Figure 3.2. SDS-PAGE gel of purified N-terminal variants of RecQ4. Lane M, molecular marker; the protein bands are shown by the arrows.

3.1.3. Expression and purification of the catalytic core of human RecQ4 and its mutants

The human RecQ4 gene encoding the Helicase and the putative RQC domain (445-1112) was cloned into pET-SUMO/CAT vector that once expressed at 17 °C using auto-induction protocol produces the Helicase and the putative RQC domain (HelRQC) with an N-terminal 6His-SUMO tag. The construct was purified using several chromatographic steps. As a first step of purification the cell lysate was applied to HisTrap column followed by an overnight cleavage of 6His-SUMO tag by SUMO protease and another HisTrap to remove the tag and the protease. The sample was further purified using Heparin affinity chromatography and finally a size exclusion chromatography (Figure 3.3). The HelRQC behaved as a monomer as the elution volume was around 14 ml (section 2.3.2). The mutant proteins including the conserved cysteines, the phenylalanine at the β-hairpin and the lysine in the Walker A region (discussed in 3.3.1) were expressed and purified following the same protocol as for the wild type protein and they behaved similar to the wild type protein, the purity of the proteins was estimated to be > 90% by SDS-PAGE (Figure 3.3). The purified samples were kept at -80°C for further characterisation.

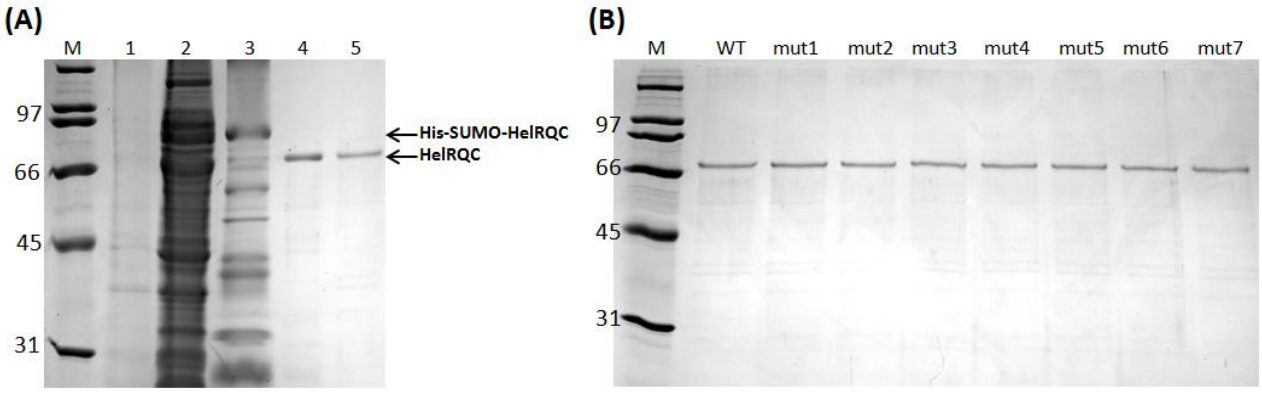

Figure 3.3. Purification of recombinant wild type and mutant RecQ4. (A) SDS-PAGE gel showing the steps of purification of wild type protein. Lane M, molecular marker; lane 1, before induction; lane 2, total cell extract after induction & growth; lane 3, eluted fractions from HisTrap column; lane 4, eluted fraction from Heparin chromatography after cleavage of the 6His-SUMO; lane 5, purified protein from the final step of size exclusion chromatography. (B) SDS-PAGE gel of purified RECQ4 variants. M, molecular marker; WT, wild type; mut1, C853A/C855A; mut2, C897A; mut3, C925A; mut4, C945A; mut5, C963A; mut6, F1077A; mut7, K508A.

3.2 The N-terminal region of human RecQ4

3.2.1. The N-terminus of RecQ4 contains a Zn-knuckle motif

The presence of a cysteine-rich motif [(S/T)**C**(F/Y)x**C**Gxx**H**WAxQ**C**] in the N-terminal domain of RecQ4 strongly suggests the presence of a Zn knuckle (Marino et al., 2013), despite occasional variation in the canonical Zn ligands. As mentioned previously, the human sequence contains a CNHC zinc-binding site, while in *Xenopus* RecQ4 the consensus CCHC is maintained. In order to investigate whether this knuckle is an autonomous structural element and to verify whether the observed changes in the consensus site did affect the structure and/or function of the human protein, we chemically synthesized two 25 amino-acid long peptides corresponding to the human (residues 397 to 421, pep-hZnK) and *Xenopus* (residues 609 to 633, pep-xZnK) Zn knuckles (Table 3.2).

Name	**Sequence**	**Molecular weight**	**pI**	**Extinction coefficient (M^{-1} cm^{-1})**
pep-hZnK (human)	VTTKES**C**FL**N**EQFD**H**WAAQ**C**PRPAS	2863	5.44	5500
pep-xZnK (*Xenopus*)	NRSGDT**C**FR**C**GGMG**H**WASQ**C**PGSVP	2610	7.98	5500

Table 3.2. Synthetic peptides. Human and *X. laevis* RecQ4 Zn-knuckle sequences. In bold red are residues predicted to be involved in zinc coordination.

In order to gain knowledge on the state of folding of the two peptides in absence and presence of Zn^{2+}, both peptides were analyzed by CD spectroscopy. In absence of zinc the two peptides showed a spectrum characteristic of unfolded polypeptides. The addition of zinc ions causes changes to the spectrum of the *Xenopus* Zn knuckle, compatible with the presence of secondary structure, while the human peptide remained unchanged (Figure 3.4A). The results were confirmed by our collaborators (C. Zucchelli and G. Musco, San Raffaele, Milan) who collected NMR 1D 1H spectra for the two peptides in the presence or absence of zinc in phosphate buffer (Figure 3.4B), suggesting that while the *Xenopus* fragment folds upon Zn^{2+} addition, the human peptide remains unstructured. As phosphate salt might sequester zinc, the NMR experiment for the human peptide was repeated in HEPES without observing any folding (Figure 3.5).

To test whether Zn binding and folding may be triggered by the presence of a physiological substrate, we repeated the CD experiments in presence of ssDNA or ssRNA. However, we did

not observe any change in the presence of nucleic acids. As a further check we collected the 1D NMR spectrum in the presence of ssDNA without detecting any folding (Figure 3.5A).

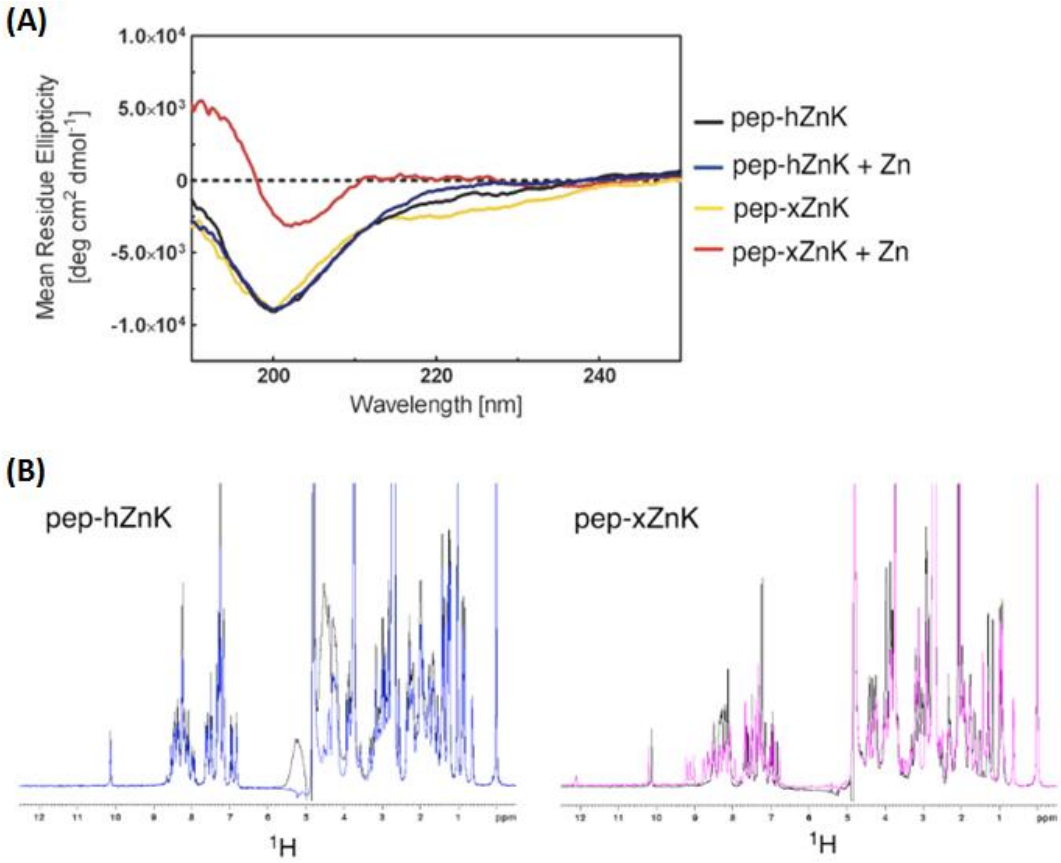

Figure 3.4. Spectroscopic studies of pep-hZnK and pep-xZnK. (A) The CD spectra of human and *Xenopus* Zn knuckles in absence of zinc are characteristic of unfolded proteins (black and yellow, respectively). The addition of zinc in equimolar amount triggers the formation of secondary structures for the *Xenopus* peptide but not for the human one. (B) Left: 1D 1H NMR spectra of the human peptide in absence (black) and presence (blue) of Zn^{2+}. Right: 1D 1H NMR spectra of *Xenopus* peptide in absence (black) and presence (pink) of Zn^{2+}.

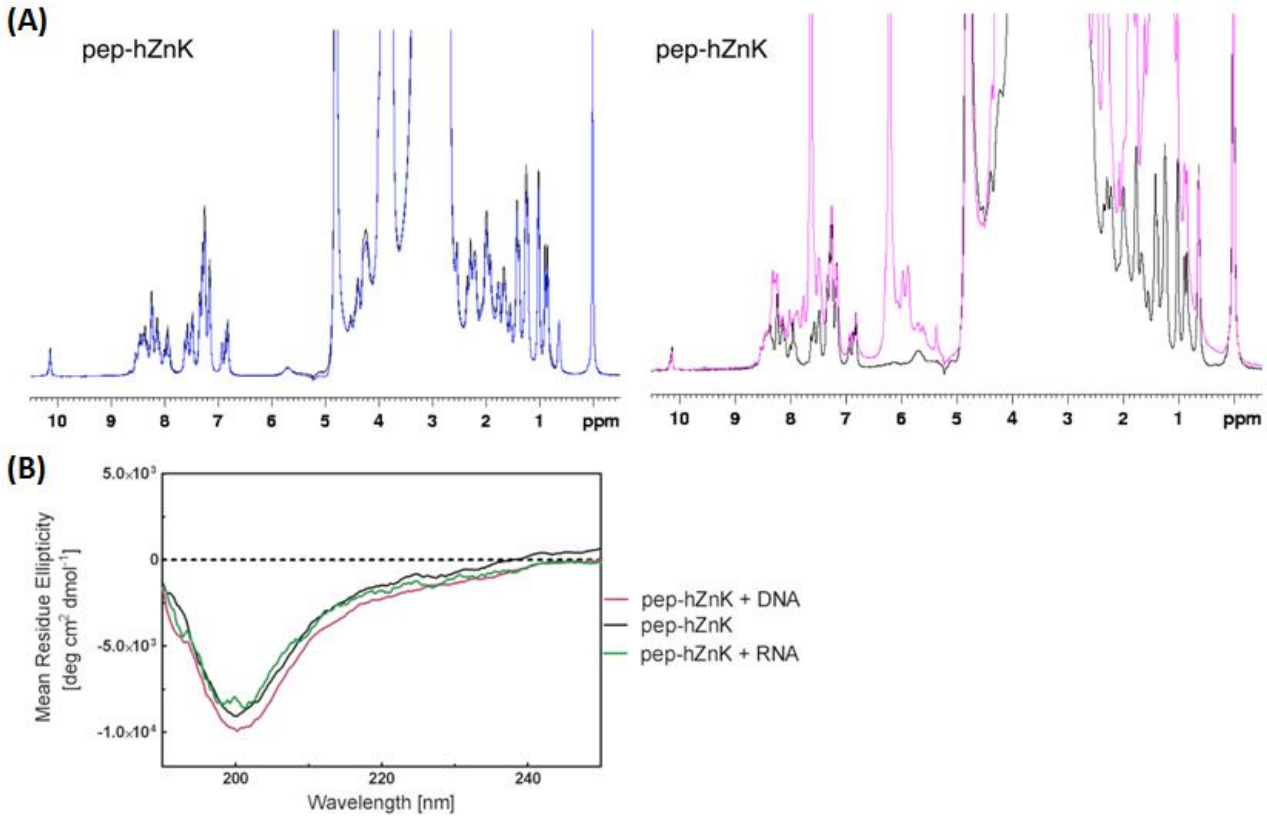

Figure 3.5. Additional spectroscopic studies of pep-hZnK (A) Left: NMR 1D ^{1}H spectrum of pep-hZnK in HEPES buffer without (black) and with (blue) zinc. Right: NMR 1D ^{1}H spectrum of pep-hZnK in presence (pink) or absence (black) of Zn^{2+} and ssDNA. (B) CD spectra of pep-hZnK: no secondary structure appears after the addition of Zn^{2+} (black) or Zn^{2+} plus ssDNA (pink) or ssRNA (green).

In summary, the 25 amino acid peptide comprising the Zn knuckle (397 - 421) corresponding to the N-terminal region of human RecQ4 did not show any folding upon zinc addition in CD or NMR experiments, contrary to its *Xenopus* counterpart. It has been reported that substitutions of zinc binding residues with Asn or Gln are compatible with Zn coordination, albeit with an altered geometry and a lower affinity, without compromising the enzymatic activity (Lesburg et al., 1997; Simpson et al., 2003). However in the case of the human RecQ4 Zn knuckle it appears that, under the experimental conditions assessed, the asparagine substitution makes the peptide unable to assume a stable 3D fold in the presence of Zn, DNA or RNA.

3.2.2. NMR structure of the *Xenopus laevis* RecQ4 Zn-knuckle

All the NMR experiments were done in collaboration with the group of Dr. Giovanna Musco from S. Raffaele Scientific Institute, Milan. The solution structure of *Xenopus laevis* RecQ4 Zn

knuckle (Asn609-Pro633) was determined by bi-dimensional homonuclear $_1$H NMR spectroscopy (Figure 3.6; structural statistics in Table 3.3). Resonance assignment was performed by 2D spectra ($_1$H-TOCSY, $_1$H - $_{15}$N and $_1$H - $_{13}$C HSQC) and the extent of proton assignments amounts to 99%. We found that upon Zn^{2+} addition the Zn knuckle folds into a small globular domain wrapping around one Zn^{2+} ion that is coordinated by three cysteines and one histidine (Cys615, Cys618, His623 and Cys628, Figure 3.6). The domain is well ordered between residues Thr614 and Cys628, with a RMSD of 0,243 Å over the backbone atoms. Typical i-i+3 NOE peaks (Trp624 Ha ~ Gln627 Hba/Hbb; Ala625 Ha ~ Cys628 Hn; Ala625 Ha ~ Cys628 Hba/Hbb) indicate the presence of an α-helical turn comprising residues 624-628. Long range NOE peaks (e.g. Gly620 Hn ~ Cys615 Hn, Gly622 Hn ~ Thr614 Ha) reveal the presence of a short antiparallel β-sheet, formed by residues 613-616 (β1) and 620-622 (β2).

A search for structural homologues of the *Xenopus* RecQ4 Zn knuckle using PDBefold (Krissinel and Henrick, 2004) finds a number of Zn knuckles, including the C-terminal Zn knuckle of the microtubule plus-end tracking protein CLIP-170 and a variety of similar folds in viral nucleocapsides and eukaryotic regulators (with Q-scores ranging from 0.67 to 0.43).

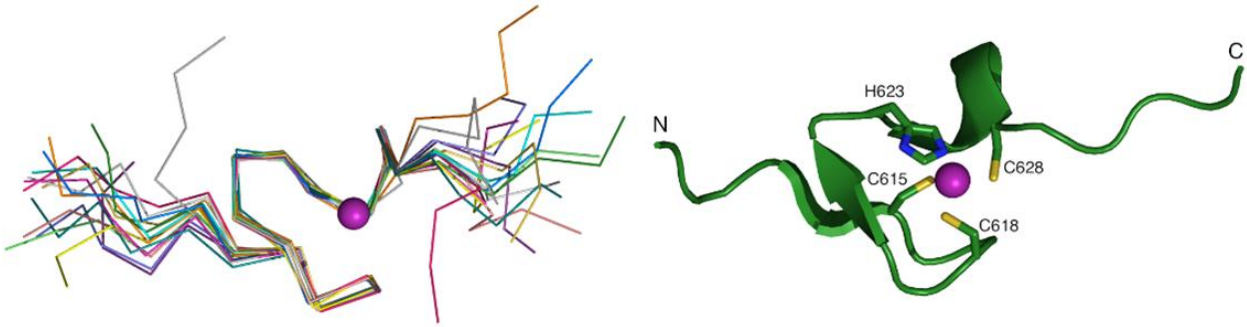

Figure 3.6. NMR structure of pep-xZnK. Left: NMR bundle of the best 15 structures for the Xenopus Zn knuckle peptide. Right: Cartoon representation of one representative NMR structure. Zn^{2+} metal ion is in purple and residues involved in Zn^{2+} coordination are depicted as ball-and-sticks.

Restraints information	**<SA> pep-xZn[a]**
NOE restraints[b]	
All	167
intra [0]	87
sequential [1]	49
short [2-3]	12
medium [4-5]	10
long [>5]	9
Dihedral angles (phi and psi)[c]	14
Hydrogen bonds[d]	2
Zn^{2+} coordination	4
RMSD (Å) [e]	
Ordered heavy atoms	0.564 ± 0.119
Ordered backbone atoms (N, Ca, C')	0.270 ± 0.096
Average RMSD from experimental restraints	
All experimental distances restraints (Å)	0.137 ± 0.010
All dihedral angle restraints (°)	0.542 ± 0.028
Ramachandran quality parameters (%) [d]	
residues in most favoured regions	83,3
residues in allowed regions	16,7
residues in additional allowed regions	0,0
residues in disallowed regions	0,0

Table 3.3. Structural statistics of Xenopus Zn knuckle (pep-xZnK)
a Simulated annealing, statistics refer to the ensemble of 15 structures with the lowest energy
b no distance restraint in any of the ensemble structures was violated more than 0.5
c no dihedral angle restraints in any of the ensemble structures was violated more than 5°
d statistics are given for residues T6-C20

3.2.3. Nucleic acid binding abilities of the Zn-knuckle domain

Many Zn knuckles are capable of binding a wide range of nucleic acid substrates. The purified Zn-knuckles were used to evaluate their nucleic acid binding by electrophoretic mobility shift assays (EMSA). Given the putative role of RecQ4 in DNA replication and a preference of Zn-knuckles towards RNA we used DNA and RNA substrates resembling a replication fork, in addition to linear single and double stranded substrates (Table 3.4).

Name	Sequence 5'-3'
A1	GTAGTGCATGTACACCACACTCTTTTTTTTTTTTTT
A2	TTTTTTTTTTTTTTGAGTGTGGTGTACATGCACTAC
A3	guagugcauguacaccacacucuuuuuuuuuuuuuu
A4	uuuuuuuuuuuuuuugagugugguguacaugcacuac

Table 3.4. List of the oligonucleotides used for the biochemical characterisation of the N-terminal region. DNA oligonucleotides are in capital letters and RNA oligonucleotides are in small letters

The human Zn knuckle, that appears to be unstructured on the basis of our NMR analysis, showed weak binding to ssDNA ($K_D \approx 0.5$ μM). However, it binds either ssRNA or forked substrates (DNA or RNA) with significantly higher affinity (Figure 3.7A and Figure 3.7C). The approximate K_D values were 220 nM for ssRNA (A3), 130 nM for forkRNA (A3:A4) and 270 nM for forkDNA (A1:A2).

The *Xenopus* Zn knuckle that possesses an intact CCHC site, in the same conditions showed no detectable binding to either ssDNA or forked DNA substrates. However, it was able to weakly bind both ss and forkRNA with K_D of 1-3 μM (Figure 3.7B and Figure 3.7D).

For both Zn knuckles, no binding was observed to short double-stranded DNA substrates of 22 base-pairs equivalent to the annealed region of the fork substrates.

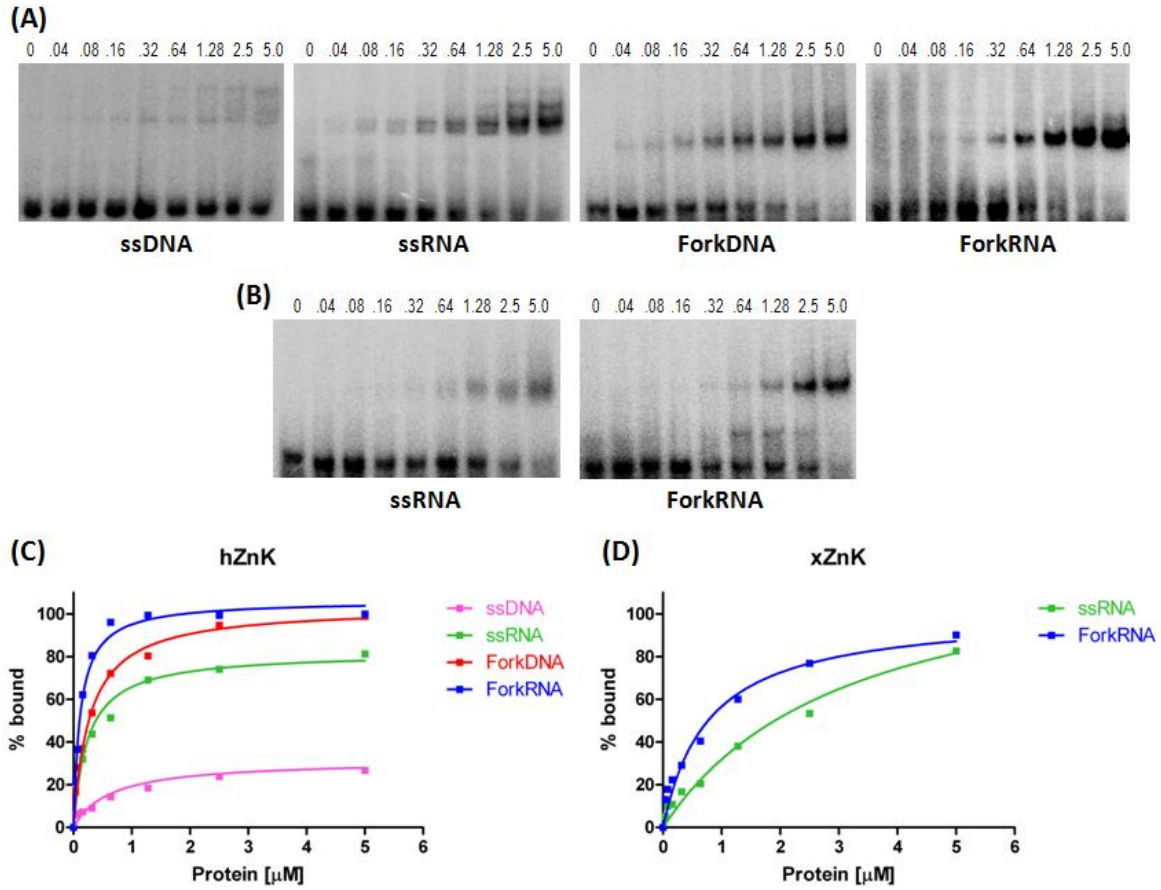

Figure 3.7. Nucleic acid binding by hZnK and xZnK (A) Example of gel shift assays for the hZnK using various nucleic acid substrates. (B) Example of gel shift assays for the xZnK using various nucleic acid substrates. (C) and (D) Quantitative EMSA analysis of various DNA/RNA complexes with the fragments of human and *Xenopus* RecQ4 corresponding to the Zn knuckle alone (hZnK and xZnK). The assays were carried out with increasing concentrations of proteins (0-5μM for the hZnK and xZnK). Each experiment was repeated at least three times to plot the binding curves. Errors were very small: for the sake of clarity error bars are not shown on the plots.

Therefore, we can say that the RecQ4 Zn knuckle preferentially binds forked substrates and that while the *Xenopus* RecQ4 Zn knuckle binds only RNA, its human counterpart is able to bind both DNA and RNA despite the lack of the canonical zinc-binding consensus. It is possible that the asparagine mutation in the human Zn knuckle give rise to a less stable but more versatile fold, able to adapt and modulate its structure according to the substrate bound. Zn knuckles are indeed known to be very adaptable scaffolds, with a mode of binding depending on the nucleic acid they encounter (Zhou et al., 2007; Nam et al., 2011)

3.2.4. Role of the additional Sld2 homology region upstream the Zn-knuckle

A highly conserved region, partly similar to the C-terminus of the yeast Sld2 initiation factors, is present upstream the Zn knuckle and includes many positively charged and aromatic residues (Marino et al., 2013). We therefore decided to investigate the role of this region in modulating the nucleic acid binding properties of the Zn knuckle.

Both the human (hUpZnK) and *Xenopus* (xUpZnK) fragments bind single-stranded and fork substrates with affinities that are consistently higher that the Zn knuckles alone (Figure 3.8). The affinity constants for each complex are reported in Table 3.5. Contrary to what seen with the Zn knuckles alone, in the presence of the upstream region, the *Xenopus* protein binds more tightly all the substrates. Both proteins displayed a preference for RNA substrates. No binding to a 22bp DNA duplex was observed.

	ssDNA	**ssRNA**	**ForkDNA**	**ForkRNA**
hUpZnK	185±33	11±0.7	42±2.8	13±1.1
xUpZnK	25±2.0	7.3±0.4	14±2.3	5.4±0.4

Table 3.5. Apparent equilibrium dissociation constants (K_D, expressed in nM) for hUpZnK and xUpZnK with various nucleic acid substrates

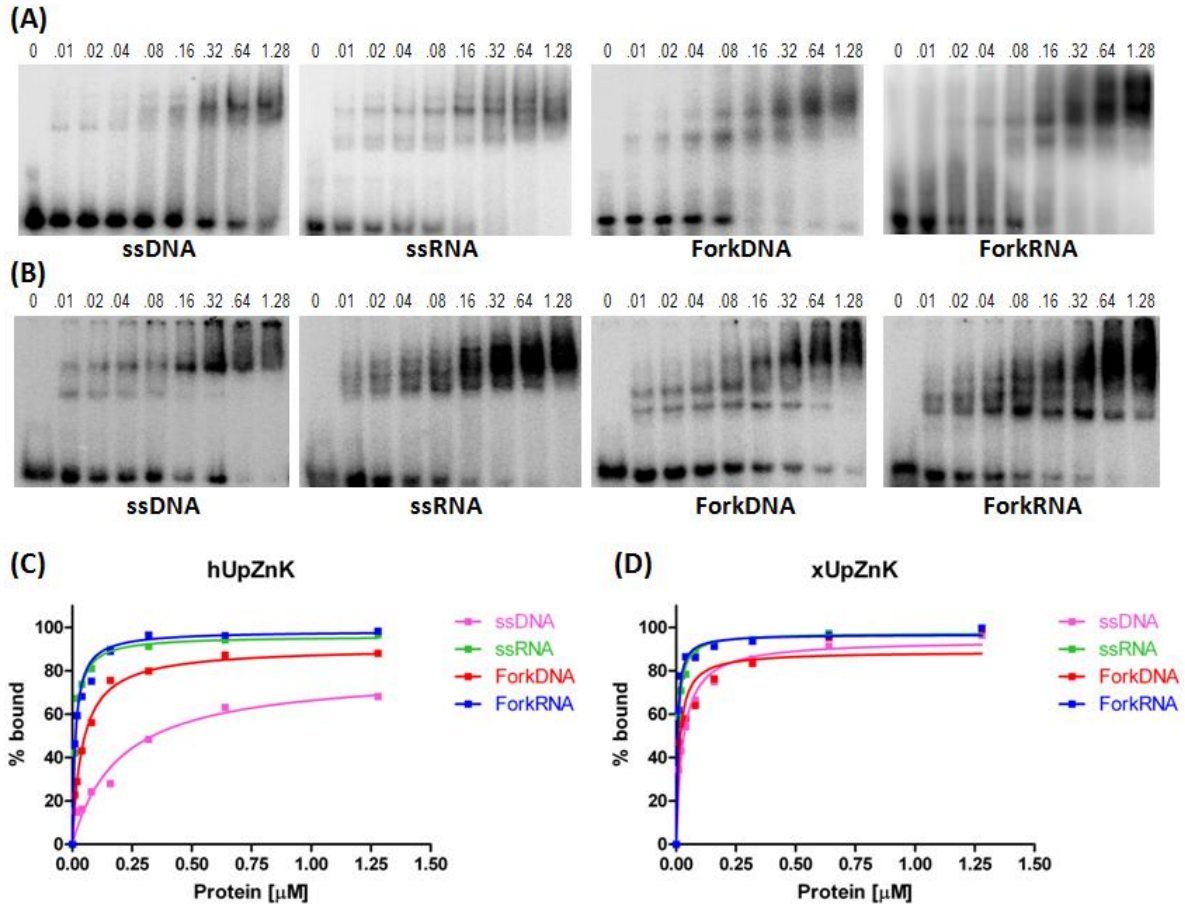

Figure 3.8. Nucleic acid binding by hUpZnK and xUpZnK (A) Example of gel shift assays for the hUpZnK using various nucleic acid substrates. (B) Example of gel shift assays for the xUpZnK using various nucleic acid substrates. (C) and (D) Quantitative EMSA analysis of various DNA/RNA complexes with the fragments of human and *Xenopus* RecQ4 corresponding to the Zn knuckle with the upstream conserved region (hUpZnK and xUpZnK). The assays were carried out with increasing concentrations of proteins (0-1.28μM for the hUpZnK and xUpZnK). Each experiment was repeated at least three times to plot the binding curves. Errors were very small: for the sake of clarity error bars are not shown on the plots.

To further dissect the interaction hUpZnK was subjected to site-directed mutagenesis to target the Zn ligands. Two mutants were designed to disrupt Zn^{2+} binding (m1: Cys403Ala/Asn406Ala; m2: His411Ala/Cys416Ala) and one to restore the canonical CCHC binding site (m3: Asn406Cys). Alanine mutations weakened the binding of the protein to DNA substrates (ssDNA, forkDNA) whereas the affinity towards RNA substrates did not change dramatically (Figure 3.9). The mutation of the Asn to Cys, reconstituting a canonical Zn knuckle pattern, did not improve binding to any of the substrates.

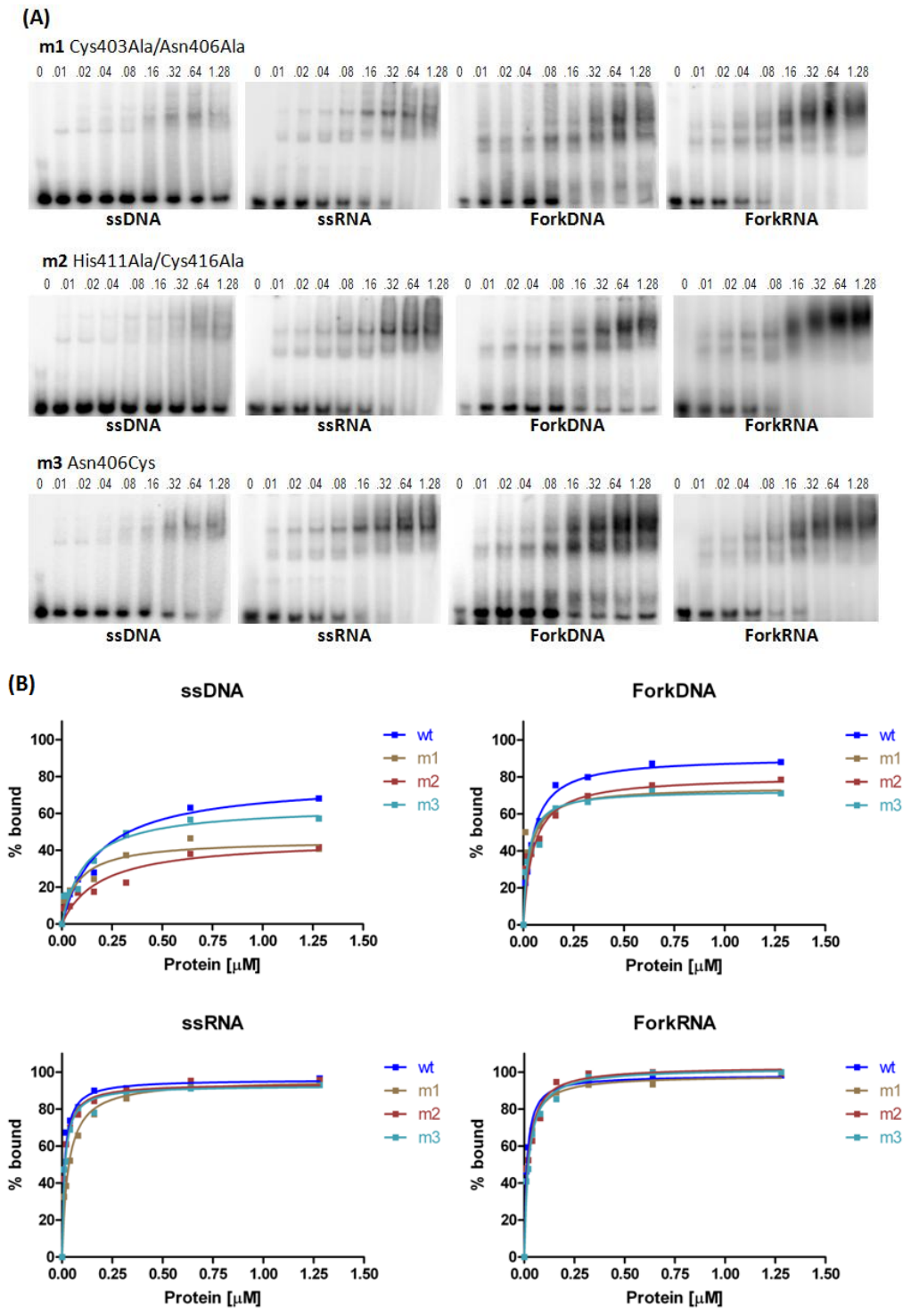

Figure 3.9. Role of the Zn ligands in nucleic acid binding. (A) Example of gel shift assays for the mutants of hUpZnK using various nucleic acid substrates. (B) Binding curves for wild type hUpZnk and mutants (m1: Cys403Ala/Asn406Ala; m2: His411Ala/Cys416Ala; m3: Asn406Cys) with ssDNA, forkDNA, ssRNA and forkRNA. Each experiment was repeated at least three times. The assays were

carried out with increasing concentrations of proteins (0-1.28μM). Each experiment was repeated at least three times to plot the binding curves. Errors were very small: for the sake of clarity error bars are not shown on the plots.

We confirmed these results by carrying out nucleic acid binding experiments in the presence of increasing amounts of EDTA, to chemically disrupt the Zn binding (Figure 3.10). EDTA strongly affects binding to forked DNA substrates, while it does not eliminate the interaction with RNA; it does however change the relative intensity of the shifted bands, suggesting that the Zn knuckle does modulate the interaction with RNA.

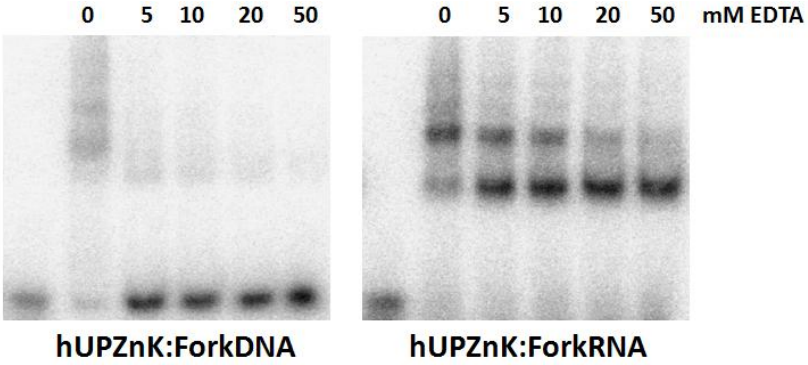

Figure 3.10. Effect of EDTA on nucleic acid binding. Binding of the hUpZnK fragment to forkDNA and forkRNA substrates, in the presence of increasing amount of EDTA. The presence of EDTA strongly affects binding to DNA substrates, while does not impair binding to RNA, although it changes the relative intensity of the shifted bands.

In summary, assessment of the role of additional Sld2 homologous region upstream the Zn knuckle in nucleic acid binding, revealed a much higher affinity for both the proteins towards all the substrates especially forked and RNA substrates compared to that of Zn knuckle alone. The dominant negative mutations of putative residues (Cys403, Asn406, His411 and Cys416) to Alanine decreased the affinity of the protein towards DNA but not RNA; similar results were obtained when EDTA was used to chelate the metal ions in nucleic acid binding reactions. The mutant human protein with a "CCHC" canonical motif in which Asn406 is replaced to Cysteine behaves the same as the other mutant proteins.

The upstream region is a strong determinant of nucleic acid binding; mutations of zinc ligands in the human construct only partially impair nucleic acid interactions. The analysis of the mutations suggests that although the Zn knuckle has a weaker affinity, it works synergistically with the upstream region.

3.2.5. Putative mode of nucleic acid binding by the N-terminal region

The structures of a number of Zn knuckles have been determined in the presence of nucleic acid substrates (Nam et al., 2011; Ohlenschläger et al., 2012; Zhou et al., 2007; D'Souza and Summers, 2004; Schuler et al., 1999; South and Summers, 1993). Although this fold is very versatile, the pattern that emerges is the presence of a nucleic acid binding platform mainly formed by exposed aromatic residues, making π-π stacking interactions with the basis. Indeed the RecQ4 Zn knuckle comprises two conserved aromatic residues (F616 and W624 in the *Xenopus* sequence) that are exposed and equivalent to the residues that sandwich the basis in the known complexes (Figure 3.11).

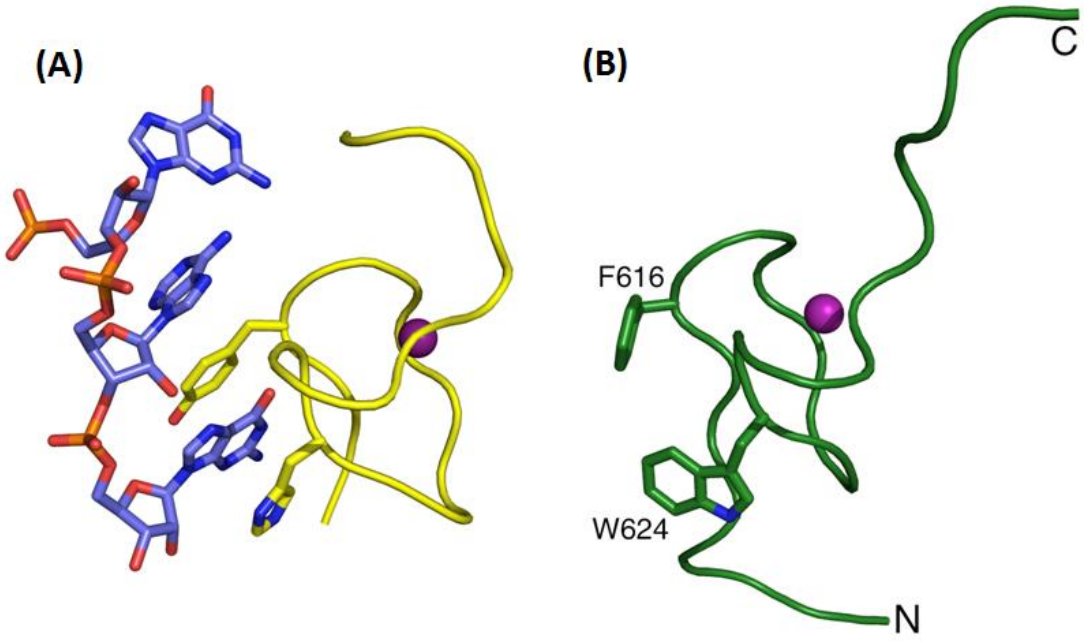

Figure 3.11. Putative mode of nucleic acid binding. (A) A detail of the Zn knuckle-RNA interactions of mouse Lin28 with let-7 microRNA (PDB code: 3TRZ), showing the critical role of aromatic residues stacking with the nucleic acid basis. (B) The Xenopus Zn knuckle is shown in the same orientation; for the sake of clarity, secondary structure elements are not depicted. The two conserved hydrophobic residues F616 and W624 are shown as ball-and-stick.

Less easy is to predict the molecular details of the interactions with the upstream region, in the absence of structural information. Bioinformatic analysis predicts the presence of two short β-strands (in the region homologous to Sld2) followed by a long α-helix. Genetic studies in *S. cerevisiae* identify a positively charged residue (K438) at the C-terminus of Sld2 which is essential for binding to DNA (Bruck et al., 2011). In RecQ4 proteins, this residue is substituted

by a conserved asparagine (N356 in the human orthologue), but a number of positively charged and aromatic residues occur on the same side of the predicted β-strand. The predicted helix, unique to RecQ4 sequences, is also very rich in aromatic and positively charged residues.

3.2.6. Putative role of the Zn knuckle?

At the moment, we can only speculate on the role of the Zn knuckle within the context of RecQ4. All the evidences point to RecQ4 having two distinct functions, one associated with the Sld2-like region and one with the helicase domain. As the Zn knuckle is exactly in the middle, it may assist one or the other. For example yeast Sld2 has been reported to bind to origin DNA (Kantar and Kaplan, 2011), and this binding could be enhanced by the Zn knuckle. At the same time it could act as an additional DNA binding domain within the helicase context.

The most intriguing finding is the ability of these regions to bind RNA substrates, raising a number of questions as to the possible physiological significance of this observation. A growing role for RNAs is emerging in a large number of cellular processes. Non-coding RNAs have been shown to play roles in the initiation of eukaryotic DNA replication in different classes of organisms, such as Y RNAs in higher eukaryotes and G-quadruplex RNAs in the replication of the Epstein-Barr (Krude, 2010). R-loop formation seems to occur more frequently than previously thought and is a potential source of genome instability at the interface between transcription and replication (Aguilera and Garcia-Muse, 2012). Genomic integrity is also dependent on non-coding RNAs: not only they are involved in transcriptional/translational regulation of DNA repair factors, but there is evidence for a more direct role for ncRNAs generated at the site of the DNA damage in establishing repair foci (d'Adda di Fagagna, 2014) and a template role for RNA in the repair mechanism has also been suggested (Storici et al., 2007). The RNA binding capability of this protein suggests that RecQ4 may be involved one or more of these processes. On the other hand, the differences between RNA and DNA binding are not huge and can simply be due to chance.

3.3. The catalytic core of human RECQ4

3.3.1. The catalytic core of RecQ4 contains an RQC domain

The putative RQC domain comprises of a Zn-binding domain and a WH domain (Figure 3.12), as seen in the crystal structures of a number of RecQ proteins such as 2OYW, 2V1X, 3AAF (Bernstein et al., 2003; Pike et al., 2009; Newman et al., 2013). The aromatic residue present at the tip of a β-hairpin in WH domain in most of the RecQ helicases has been reported to be involved in DNA binding and unwinding. An equivalent aromatic residue (Phe1077) is present in RecQ4. Moreover, the Zn binding domain contains a set of four conserved cysteine residues, coordinating a Zn^{2+} atom. This region in RecQ4 contains large insertion regions (Figure 3.12) these insertions makes it difficult to exactly correlate the Zn-binding residues in RecQ4 with the canonical Zn ligands in other RecQ helicases. In particular, the four cysteines coordinating a Zn atom in the atomic structure of *E. coli* RecQ (Bernstein et al., 2003) and human RecQ1 (Pike et al., 2009) are only partially conserved. However, a second pattern of conserved Zn-binding candidates are found in the insertions, structurally located in positions compatible with a role in Zn coordination (Figure3.12), suggesting the possibility of a second metal-binding fold.

To validate the presence of an RQC domain in human RecQ4, predicted by the *in silico* analysis (Marino et al., 2013), several putative functional residues were mutated to alanine. As observed, only two of the Zn ligands seen in RecQ1 and *E. coli* RecQ are absolutely conserved in RecQ4 (Cys925 and Cys945) with other residue only partially conserved. Moreover other four potential conserved Zn ligands are detected in the insertion (Cys853, Cys855, Cys897 and Cys963) (Figure 3.12). The aromatic residue at the β-hairpin of winged helix domain Phe1077, the Walker A Lys508 and the cysteines mentioned above were chosen to be mutated to alanine.

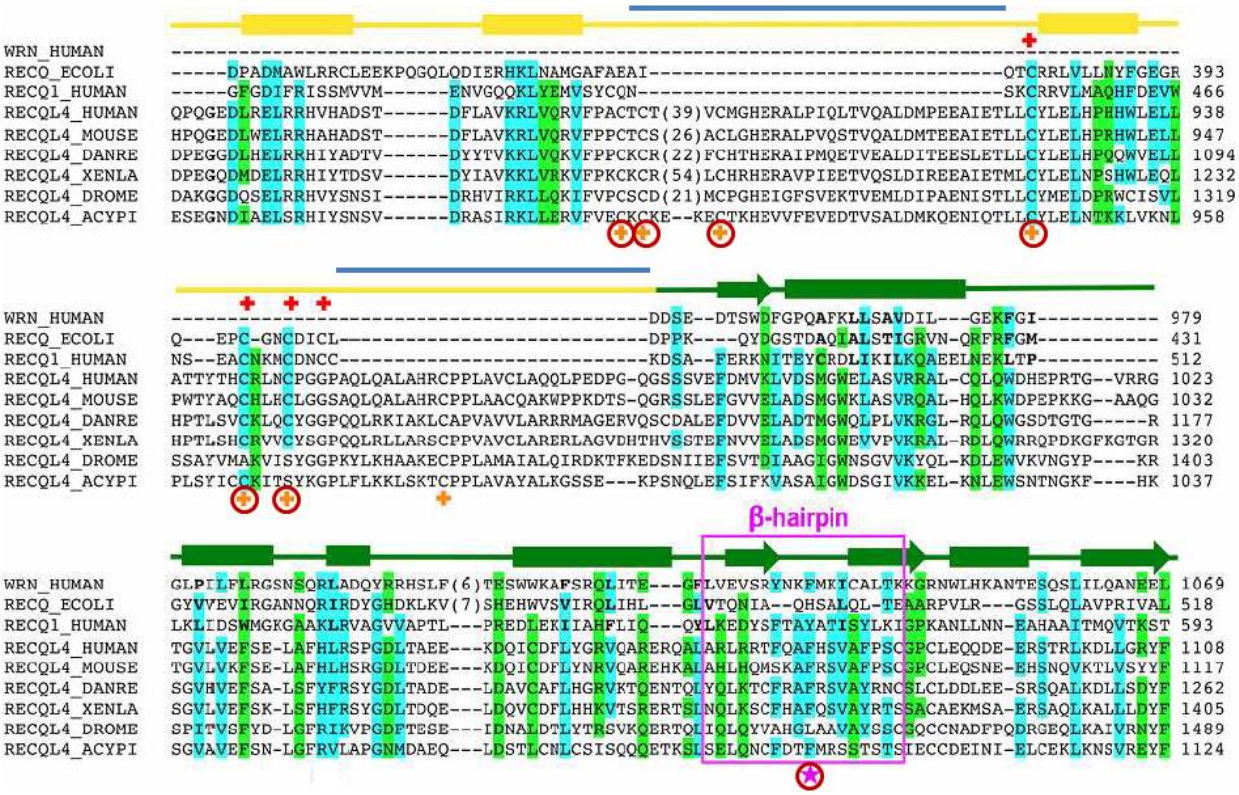

Figure 3.12. The RQC domain of RecQ4. The putative RQC domain of RecQ4 is aligned with the equivalent region in human RecQ1 and *E. coli* RecQ, and with the WH domain of Werner. The alignment is structure-based (PDB accession codes: 1OYW, 2V1X, 3AAF). Highlighted are the residues that are conserved within the RecQ4 sequences and at least one other sequence. In cyan are highlighted the residues that are identical or very similar and in green residues that show a similar character. The secondary structure elements are based on the crystal structure of human RecQ1 (α-helices are shown as rods and β-strands as arrows): in yellow is shown the region corresponding to the Zn domain, in green that corresponding to the winged helix domain. The two insertions in RecQ4 before and after the third helix of RQC domain are shown in blue. The four cysteine residues coordinating a Zn atom in the two crystal structure are shown by red crosses above the sequences; conserved cysteines present in the insertion region which could possibly be involved in Zn binding in RecQ4 are shown by orange crosses below. The non-polar residues that have been described as part of the WH core in Werner (Kitano et al., 2010) are shown in bold. The β-hairpin that in the RecQ1 and WRN proteins has been shown to act as a pin is enclosed in a magenta box, with the key tyrosine/phenylalanine identified by a magenta star. The residues mutated in this work are marked with a red circle. Picture adapted from Marino et al., 2013.

3.3.2. Folding and stability of the mutants

In order to evaluate the folding and stability of the purified proteins heat denaturation experiments were performed using SYPRO Orange (SO) fluorescent dye, using the Thermofluor assay. The fluorescence emission intensity of the SO dye increases when it binds to the hydrophobic patches of the unfolded protein. This property was used to determine the melting temperatures of the purified proteins by monitoring their heat denaturation curves. As shown in

Figure 3.13, the wild type and mutant proteins have the same heat denaturation profiles, and the melting temperatures (Tm) of the proteins were within 41.6±0.3°C (within the error range) as shown in Table 3.6, thus suggesting a comparable stability of the wild type and mutant proteins and indicating that the mutations did not affect the 3D structure of the protein.

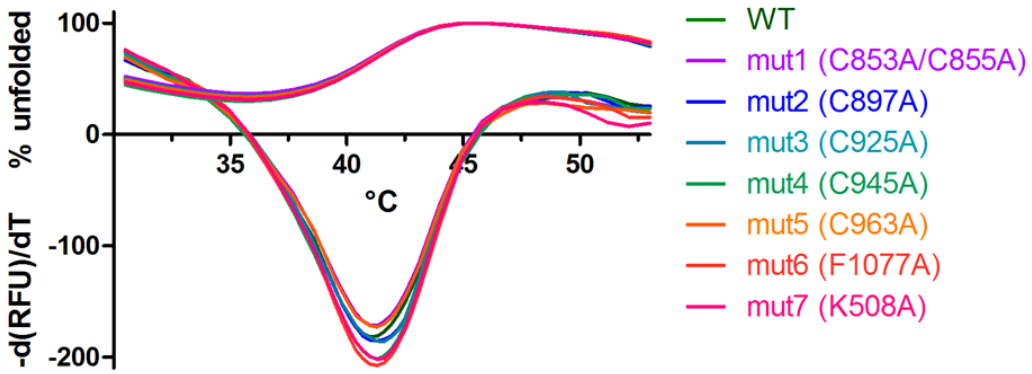

Figure 3.13. Heat denaturation profile of RecQ4 variants. The curves depicting the percentage of unfolded proteins (positive 'y' quadrant, 0 – 100) and the –(dRFU)/dT curves are plotted together (mostly in the negative 'y' quadrant).

Protein	Tm (°C)	ΔTm (°C)
Wild type (WT)	41.7	–
Mut1 (C853A/C855A)	41.4	– 0.3
Mut2 (C897A)	41.6	– 0.1
Mut3 (C925A)	41.4	– 0.3
Mut4 (C945A)	41.4	– 0.3
Mut5 (C963A)	41.8	+ 0.1
Mut6 (F1077A)	41.9	+ 0.2
Mut7 (K508A)	41.8	+ 0.1

Table 3.6. The melting temperature of RecQ4 mutants. These values were obtained from heat denaturation curves of protein. The experiment was carried out in triplicate and errors are ±0.3 °C. ΔTm = Tm(mutant) – Tm(wild type).

After confirming that the mutations did not affect the folding and stability, we tested these recombinant proteins for several biochemical activities such as metal binding, DNA binding activity, ATPase activity and helicase activity.

3.3.3. Metal binding analysis

All the purified proteins were analysed by Inductively Coupled Plasma-Atomic Emission Spectrometry (ICP-AES) to determine the presence of $Zn^{2+}$and to quantify the amount. As cystein residue can also bind Fe^{2+} and form FeS clusters (which are found in a subset of helicases), an analysis of the Fe^{2+} content was also carried out on the wild-type protein, with negative results.

The wild type protein, the β-hairpin mutant (mut6 F1077A) and the Walker A mutant (mut7 K508A) showed the presence of Zn^{2+} with a 1:2 molar ratio (i.e. for every molar of protein two molar Zn^{2+} ion was present), while the Zn binding mutants (mut1 C853A/C855A, mut2 C897A, mut3 C925A, mut4 C945A and mut5 C963A) showed the presence of Zn^{2+} with a 1:1 molar ratio (shown in Table 3.7).

Protein	Protein concentration (µMolar)	Zinc concentration (µMolar)	Protein:Zinc (Molar ratio)
Wild type (WT)	2.5	4.9	1:2
Mut1 (C853A/C855A)	0.3	0.34	1:1
Mut2 (C897A)	2.0	2.2	1:1
Mut3 (C925A)	1.5	1.6	1:1
Mut4 (C945A)	1.5	1.7	1:1
Mut5 (C963A)	2.0	2.1	1:1
Mut6 (F1077A)	1.2	2.5	1:2
Mut7 (K508A)	1.5	2.9	1:2

Table 3.7. The amounts of Zinc present in the RecQ4 variants. The mutations in the Zn binding domain of RecQ4 result in decrease of Zn binding while the mutation at the β-hairpin and Walker-A do not affect the Zn binding property of the protein.

These results suggests that the Zn binding domain within the RQC includes two independent Zn^{2+} binding regions: most likely one region is located in the insertion between the second and third helix within the Zn domain (and includes Cys853, Cys855, Cys897), while the second is located between the third helix and the WH domain, and is equivalent to what seen in RecQ1 and

other RecQ helicases (including Cys925, Cys945, Cys963). Additional mutations, in combination, will be needed to confirm this. As predicted, mutation of the Walker A or β-hairpin residues does not affect the amount of Zn^{2+} bound.

3.3.4. Nucleic acid binding

To test the nucleic acid binding capacity of the mutant proteins, various DNA/RNA substrates were prepared and labelled as described in Section 2.3.6 (Table 3.8) and used for electrophoretic mobility shift assay or fluorescence-based nucleic-acid binding or helicase assay. The reason to explore both is due to the fact that our laboratory is not certified to work with radioactive material. However after doing some preliminary tests the superiority of radioactive based EMSA assays was such that we continued with this technique, thanks to the hospitality of the laboratory of Dr. Riccardo Sgarra (Dipartimento di Scienze della Vita, Universita' di Trieste).

Name	**Sequence 5'-3'**
F1	[6FAM] CTACTACCCCCACCCTCACAACCTTTTTTTTTTTTTT
F2	TTTTTTTTTTTTTTGGTTGTGAGGGTGGGGGTAGTAG [BHQ1]
B1	[6FAM] CTACTACCCCCACCCTCACAACC
B2	GGTTGTGAGGGTGGGGGTAGTAG [BHQ1]
Capture	CTACTACCCCCACCCTCACAACC
F3	GTAGTGCATGTACACCACACTCTTTTTTTTTTTTTT
F4	TTTTTTTTTTTTTTGAGTGTGGTGTACATGCACTAC
B3	GTAGTGCATGTACACCACACTC
B4	GAGTGTGGTGTACATGCACTAC
R1	[6FAM] cuacuacccccacccucacaaccuuuuuuuuuuuuuu
R2	uuuuuuuuuuuuuugguugugaggguggggguaguag [BHQ1]
R3	guagugcauguacaccacacucuuuuuuuuuuuuuu
R4	uuuuuuuuuuuuuuugaguguggguguacaugcacuac

Table 3.8. List of the DNA/RNA oligonucleotides used in the biochemical characterisation of the catalytic core. F1, B1 and R1 are labelled with 6FAM at the 5' end and F2, B2 and R2 are labelled with BHQ1 at the 3' end. DNA oligonucleotides (F1-B4) are in capital letters and RNA oligonucleotides (R1-R4) are in small letters.

Although RecQ helicases have mostly been shown to bind and unwind DNA substrates, the slight preference for RNA displayed by the Zn knuckle prompted us to test also RNA substrates.

Nucleic acid binding activity of the wild type protein on ssDNA (F3), blunt dsDNA (B3:B4), forkDNA (F3:F4) and forkRNA (R3:R4) substrates was evaluated by using electrophoretic mobility shift assay (EMSA). The most favored substrate was the forkDNA with a K_D of around 750nM (Figure 3.14A-B). The forkRNA substrate showed low level of binding.

The assessment of forkDNA binding of the mutant proteins showed that the mutants belonging the RQC domain (the Zn-binding cysteines and aromatic residue at the β-hairpin) lost their ability in DNA binding, confirming the essential role of this region for helicase activity. On the contrary, the WalkerA mutant behaved similarly to the wild type protein, as expected (Figure 3.14C).

The level of binding is not very high. In the follow up we are planning to repeat these experiments in different conditions to optimize the interactions, as well as add a non-hydrolysable ATP analogue to verify whether it stimulates binding. Of course, it may be that the low level of binding is due to the fact that we don't have the physiological substrate (as RecQ have been shown to work with a bewildering range of nucleic acid structures); alternatively the high affinity for substrates of the Zn knuckle and upstream region may act sinergically to deliver the substrate to the enzyme active site. Due to time constraints we didn't manage yet to test all these possibilities.

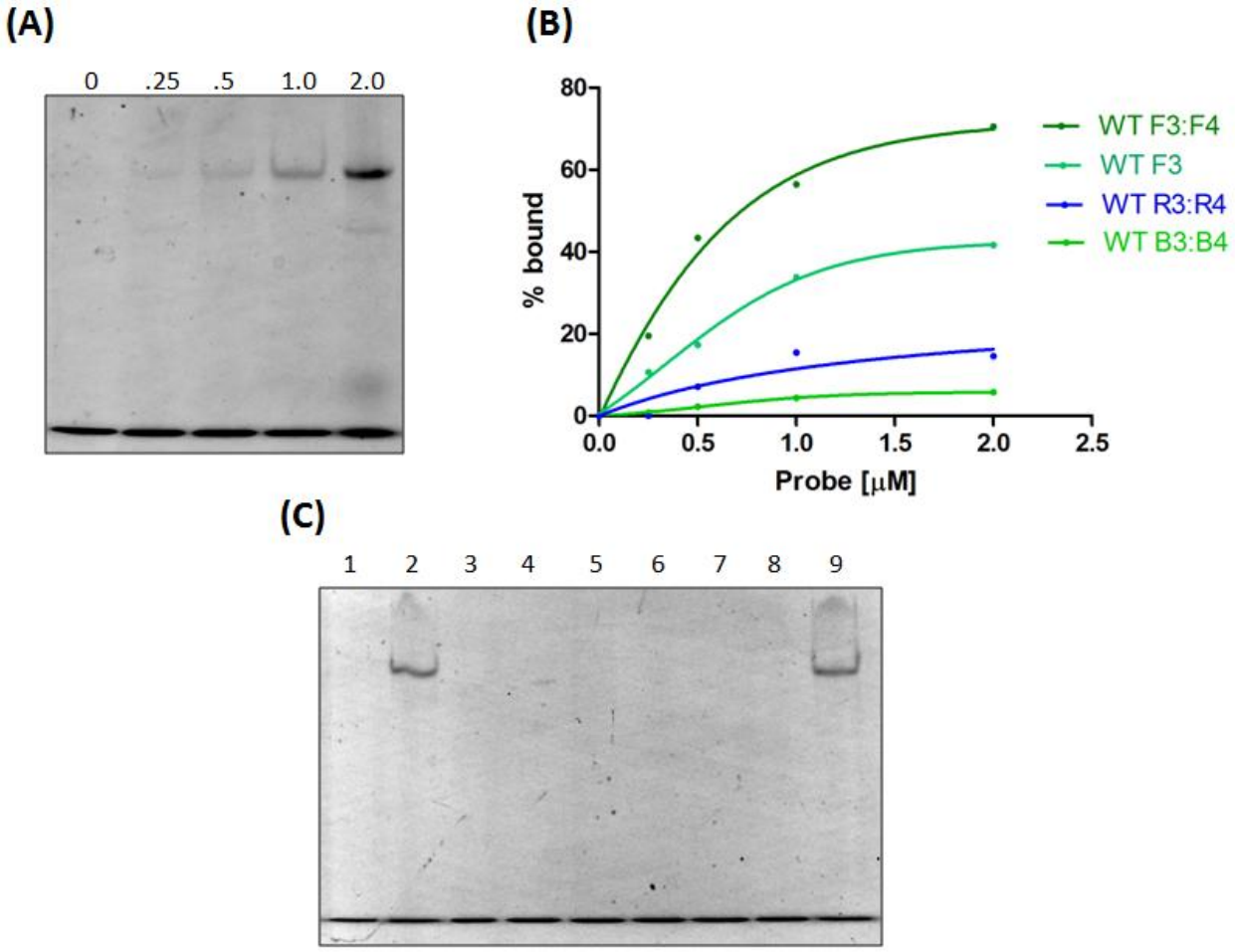

Figure 3.14. Nucleic acid binding of the purified recombinant wild type (WT) and mutant proteins (A) EMSA gel showing the ForkDNA binding activity of wild type HelRQC with increasing concentration of protein (0-2μM). (B) A comparison of DNA binding activity of the WT protein on four different substrates – single stranded DNA (F3), blunt dsDNA (B3:B4), forkDNA (F3:F4) and forkRNA (R3:R4). ForkDNA seems to be the preferred substrate for binding. (C) EMSA gel showing the DNA binding activity of the purified proteins: lane 1, forkDNA (F3:F4) substrate only; lane 2, wild type; lane 3, mut1 C853A/C855A; lane 4, mut2 C897A; lane 5, mut3 C925A; lane 6, mut4 C945A; lane 7, mut5 C963A; lane 8, mut6 F1077A; lane 9, mut7 K508A. All the mutants except mut7 K508A, show a loss of DNA binding activity. Each experiment was repeated at least three times to plot the binding curves. Errors were very small: for the sake of clarity error bars are not shown on the plots.

3.3.5. ATPase activity

ATPase activity of the wild type protein (HelRQC) was evaluated by using ssDNA (F3), blunt dsDNA (B3:B4) and forkDNA (F3:F4) substrates to stimulate the ATP hydrolysis. The activity was stimulated by forkDNA substrate significantly better than by the other substrates (Figure 3.15).

In order to see the effect of the mutations on ATP hydrolysis, ATPase assays were performed on the mutant proteins using forkDNA substrate as a stimulant. All the RQC mutants behaved similarly to the wild type protein, with variation well within the statistical error (Figure 3.15). As a control we tested the WalkerA mutant (mut7 K508A), which is not expected to bind ATP, and indeed showed a loss of ATPase activity. This is a further confirmation that the RQC mutants are properly folded and able to display a catalytic activity. It also demonstrate that the helicase core and RQC are able to work as independent modules.

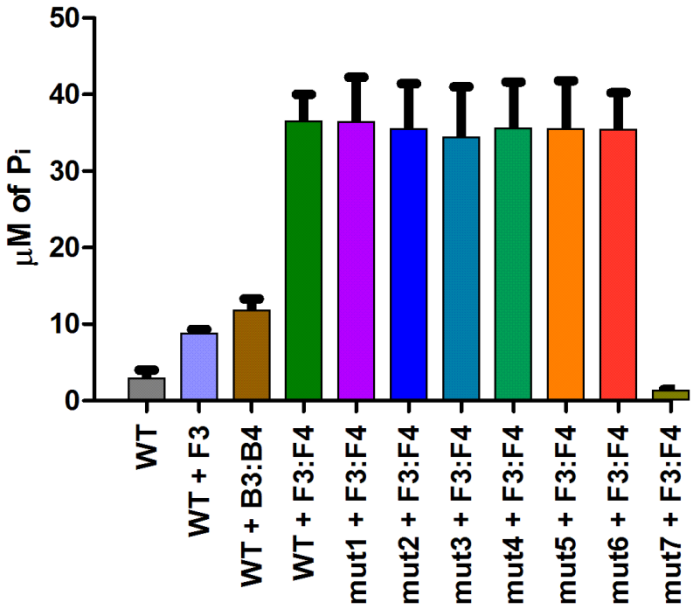

Figure 3.15. ATPase assay of wild type (WT) and mutant proteins. ATP hydrolysis by WT protein is highly stimulated by forkDNA (F3:F4) compared to other substrates (ssDNA F3 and blunt dsDNA B3:B4). Mutation in the Walker-A motif (mut7, K508A) of RecQ4 results in a loss of ATPase activity of the protein. All the experiments are done in triplicate.

3.3.6. Nucleic acid unwinding

Whereas fluorence experiments did not work well for nucleic acid binding, we manage to set up a succesfull protocol for nucleic acid unwinding in our laboratory, using a FRET-based method (Section 2.3.8). A variety of substrates were tested, including blunt dsDNA (B1:B2), forkDNA (F1:F2), forkRNA (R1:R2), forkDNA-RNA (F1:R2, with a DNA 3' overhang) and forkRNA-DNA (R1:F2, with a RNA 3' overhang).

The results shows that the forkDNA is significantly preferred over the other substrates (Figure 3.16A), as already reported for other RecQ helicases (Croteau et al., 2014). Less obviously, a lower level of activity was observed also with the blunt-ended substrate; this had already been observed for the *E. coli* RecQ protein (Umezu et al., 1990) but, to the best of our knowledge, never reported for a eukaryotic enzyme. However the crystal structure of the Werner WH domain with DNA shows the haiprin residue wedging into a blunt substrate (Kitano et al., 2010). Interestingly, the protein shows a lower level of activity with forkRNA and DNA-RNA hybrid substrates (forkDNA-RNA and forkRNA-DNA) compared to that with forkDNA and shows a slight preference for substrate with a DNA 3' overhang (forkDNA-RNA, F1:R2) over the substrate with RNA 3' overhang (forkRNA-DNA, R1:F2). Although, the helicase activity of RecQ4 towards DNA-RNA hybrids is shown for the first time in this work, BLM has been reported earlier to unwind hybrid substrates (Popuri et al. 2008), as well as WRN shows helicase activity towards DNA-RNA hybrids and okazaki fragments (Chakraborty and Grosse, 2010).

We then checked the DNA unwinding activity of the mutant proteins, using forkDNA as a substrate, and all the mutations almost abolished the helicase activity (Figure 3.16B). It was obviously expected that the WalkerA Lys508, which is involved in ATP binding, is crucial for DNA unwinding. The results obtained for the aromatic residue at the end of the WH β-hairpin (Phe1077) does not only confirm its essential role as the "pin" opening up the dsDNA, as seen in RecQ1 and WRN helicases (Pike et al., 2009; Lucic et al., 2011; Kitano et al., 2010), but lends strong support to the accuracy of the bioinformatic analysis (Marino et al., 2013) and validates the presence of a fully functional RQC domain in RecQ4. Even more interesting are the results on the cysteine mutants. Two of the Zn ligands found in *E. coli* RecQ had been mutated to Asn and shown to abolish the helicase activity (Liu et al., 2004), as well as mutation of the Zn-ligands in BLM either results in insoluble protein or biochemically compromised protein (Guo et al., 2005). RecQ4 is unusual in having two Zn^{2+} atoms and remarkably we find that they are both critical for the helicase activity.

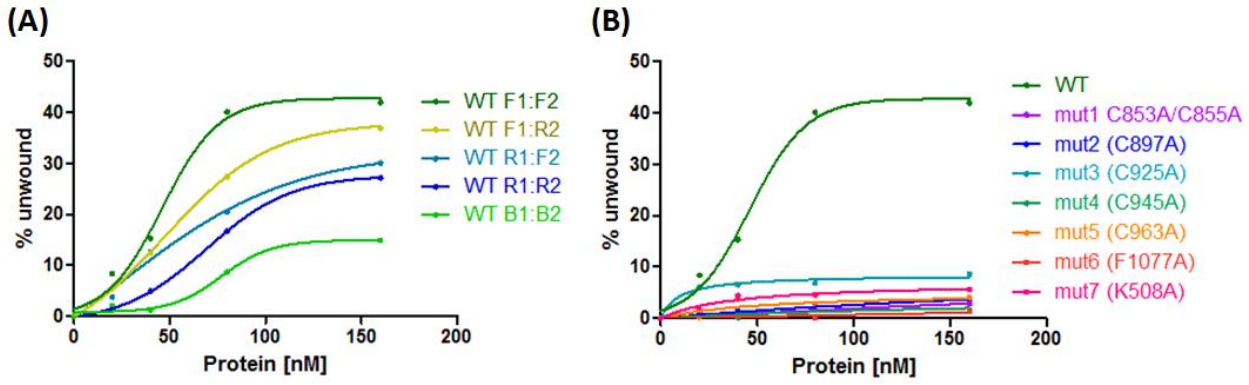

Figure 3.16. DNA unwinding property of RecQ4 and its mutants. (A) Comparison of helicase activity of the WT protein on two different substrates – blunt dsDNA (B1:B2), forkDNA (F1:F2), forkRNA (R1:R2), forkDNA-RNA (F1:R2, with a DNA 3' overhang) and forkRNA-DNA (R1:F2, with a RNA 3' overhang). ForkDNA seems to be the preferred substrate for unwinding. (B) A comparison of helicase assay between the mutant proteins and the wild type. Mutations in the RQC domain and WalkerA motif of RecQ4 almost abolish the DNA helicase activity of the protein. Each experiment was repeated at least three times to plot the binding curves. Errors were very small: for the sake of clarity error bars are not shown on the plots.

3.3.7. Small Angle X-Ray Scattering (SAXS) on the catalytic core of human RecQ4

The protein has been subjected to extensive crystallization trials, but unfortunately no single crystals could be obtained; a possible explanation for this failure is the presence of the two long extensions within the RQC domain (Figure 3.12) which may interfere with the assembly of a large 3D lattice. We are planning to re-clone the protein without these extensions in the hope of obtaining crystals suitable for X-ray diffraction.

In the meantime, to structurally characterize the catalytic core of human RecQ4, Small Angle X-Ray Scattering (SAXS) was used. The most important advantage of SAXS is the use of samples in solution without the need of crystals. In a strongly diluted and monodisperse solution it is possible to estimate with a high degree of precision the structural parameters of the particle: the radius of gyration (Rg) and the maximum length (D_{max}) (Svergun and Koch, 2003). On the other hand, the biggest disadvantage is that only low resolution models could be obtained from randomly oriented particles.

In an attempt to structurally characterize the protein, SAXS data were collected from three samples (HelRQC1, HelRQC2 and HelRQC2+forkDNA), in collaboration with Matteo De

March (Elettra). The data analysis was done as described in section 2.4.4 with the ATSAS suite (Petuokhov et al., 2012).

The structural parameters obtained from the final scattering curve of the first sample (HelRQC1) are summarized in Table 3.9 and compared with those calculated from the RecQ1 crystal structure (PDB 2V1X) and RecQ1 crystal structure in complex with 13bp-dsDNA (PDB 2V1X). As the DNA fork we used had a double-stranded portion significantly longer that the DNA segment co-crystallysed with RecQ1 (21bp rather than 13bp), we attempted to model the extra base pairs.

Sample	Q range ($Å^{-1}$)	$I_{(0)}$	Rg (Å)	Dmax (Å)
HelRQC1	0.011 – 0.46	148.9 ± 0.9	37.77 ± 0.04	150
RecQ1 minus DNA			26.71	95
RecQ1 + 13bp DNA			29.52	115
RecQ1 + 21bp DNA (model)			36.40	145

Table 3.9. SAXS model parameters of HelRQC1. Summary of the parameters estimated from the scattering curves of the first RecQ4 sample (HelRQC1) and the theoretical curves derived from the RecQ1 crystal structure. The parameters include – Scattering angle Q range; Intensity $I_{(0)}$; Radius of gyration Rg and maximum length Dmax.

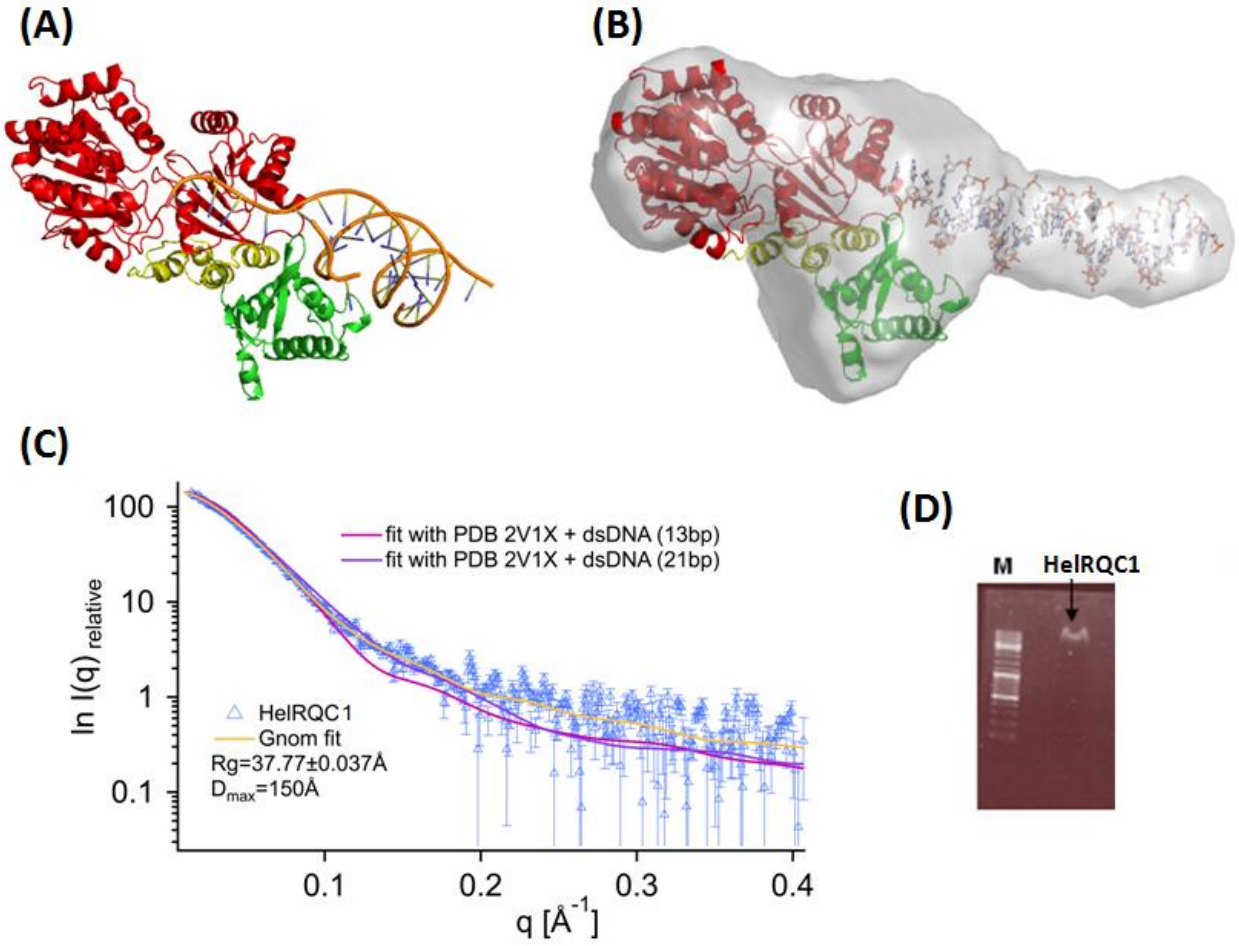

Figure 3.17. SAXS analysis of the catalytic core of human RecQ4 with DNA contamination (HelRQC1). (A) Crystal structure of RecQ1 (2WWY) in the same orientation as fitted in the RecQ4 SAXS model. (B) Low resolution envelope with D_{max}=150Å of HelRQC1 (gray, radius set to 5Å) averaged after 30 DAMMIF runs fitted with the RecQ1 crystal structure (PDB 2V1X) modeled with 21bp-dsDNA. The helicase domain of RecQ1 crystal structure is in red, the Zn-binding domain is in yellow and the winged helix domain is green. (C) Experimental SAXS profile of HelRQC1 with standard deviation (light blue triangle) and its best fit with GNOM (yellow line). Theoretical scattering curves simulated from the crystal structure of RecQ1 with 13bp-dsDNA (PDB 2V1X, magenta line) and from the same structure modeled with 21bp-dsDNA (purple line). (D) Agarose gel of sample HelRQC1 with marker (M) that confirms the presence of DNA. The DNA stays in the well may be because it is tightly bound to the protein.

Sample **HelRQC1** was purified and SAXS data were collected and analyzed (section 2.4.4). As shown in figure 3.17C, the experimental scattering curve of the sample HelRQC1 has a good signal to noise ratio both at low and high scattering angle (q) reaching a resolution around 20-25Å. The linear trend plotted by the Guinier's approximation with Rg = 38Å confirmed the absence of particle aggregation/repulsion. A model of Dmax = 150 Å was obtained by averaging

30 *ab initio* models (obtained from DAMMIF) using DAMAVER. As shown in figure 3.17B, the compact core is in agreement with the three-dimensional crystal structure of the catalytic core of human RecQ1 (PDB 2V1X). The presence of the side blob, modeled with 21bp-dsDNA, was confirmed by testing different Dmax values and after averaging 30 DAMMIF runs with the best Dmax=150Å. It was speculated to be a stretch of 21bp DNA, later confirmed by running the sample on 1% agarose DNA gel (Figure 3.17D). In order to get more information about HelRQC1 in solution and to have a comparison with the available crystal structures of RecQ1, the experimental scattering curve was compared with the theoretical scattering curves simulated from the crystal structure with PDB code 2V1X (RecQ1 with 13bp-dsDNA) and from the same structure modeled with 21bp-dsDNA (Figure 3.17C) showing a best fit with the theoretical scattering curve of RecQ1 modeled with 21bp-dsDNA.

In the cell RecQ4 plays a role at the interface of DNA replication, repair and recombination hence a high affinity towards nucleic acid is expected. The DNA binding property of RecQ4 is a hurdle to cross while purifying the protein to homogeneity. In order to remove the DNA contamination, a new batch of protein was purified (HelRQC2) with an extra step of washing the sample with 3M NaCl (section 2.2.2). To confirm the DNA binding properties of the catalytic core of RecQ4 based on the EMSA results (section 3.3.2) we decided to collect SAXS data on a new sample prepared as HelRQC2 with an equimolar amount of ForkDNA (F3:F4) in order to get the protein-DNA complex (HelRQC2+forkDNA). SAXS data were collected at the AustroSAXS beamline at Elettra as described in section 2.4.4 and reduced using the ATSAS softwares (Petoukhov et al., 2012). The structural parameters obtained from the final scattering curve of HelRQC2 and HelRQC2+forkDNA are summarized in Table 3.10.

Sample	Q range ($Å^{-1}$)	$I_{(0)}$	Rg (Å)	Dmax (Å)
HelRQC2	0.015 – 0.41	2.90 ± 0.01	35.99 ± 0.12	110
HelRQC2+forkDNA	0.015 – 0.41	4.23 ± 0.04	47.52 ± 0.03	160

Table 3.10. SAXS model parameters of HelRQC2 and HelRQC2+forkDNA. Summary of the parameters estimated from the scattering curves of the two RecQ4 samples (HelRQC2 and

HelRQC2+forkDNA). The parameters include – Scattering angle Q range; Intensity $I_{(0)}$; Radius of gyration Rg and maximum length Dmax.

The linear trend from Guinier's approximation of **HelRQC2** experimental data with a resolution around 30-35Å suggested no aggregation with the Rg estimation (Rg=36Å) smaller than estimated for HelRQC1. The best P(r) plot with Dmax=110Å was chosen after checking this value in a range between 110Å and 130Å and it was further used to calculate 30 *ab initio* DAMMIF reconstructions. All of these runs were averaged in order to get the final more compact model (Figure 3.18A) that could be fitted by the 3D crystal structure of human RecQ1 (PDB 2V1X) only after removing the DNA. The sample was run on agarose gel to confirm the absence of DNA (Figure 3.18A) and the theoretical scattering curves both from PDB 2V1X without DNA (green line) and from PDB 2V1X with 13bp-dsDNA (magenta line) were simulated and fitted on the SAXS experimental data of HelRQC2 (Figure 3.18B). None of these simulated curves fit the experimental data and both of them showed a smaller slope at low angle q indicating a smaller Rg. This observation could be explained by the presence of the two insertion regions of 70 and 30 residues in the catalytic core of human RecQ4 respect to RecQ1 structure (Figure 3.12).

SAXS data from a diluted and homogeneus solution (Figure 3.18D) of **HelRQC+forkDNA** showed a resolution below 30Å and gave the higher Rg value of 47 Å with the best Dmax of 160Å. The final model (Figure 3.18C), averaged from 30 DAMMIF runs, with an additional elongation similar to that found in HelRQC1 model is in agreement with the 3D crystal structure with PDB code 2V1X modeled with 21bp-dsDNA. The presence of DNA in the sample was further verified by running the sample in agarose gel (Figure 3.18C).

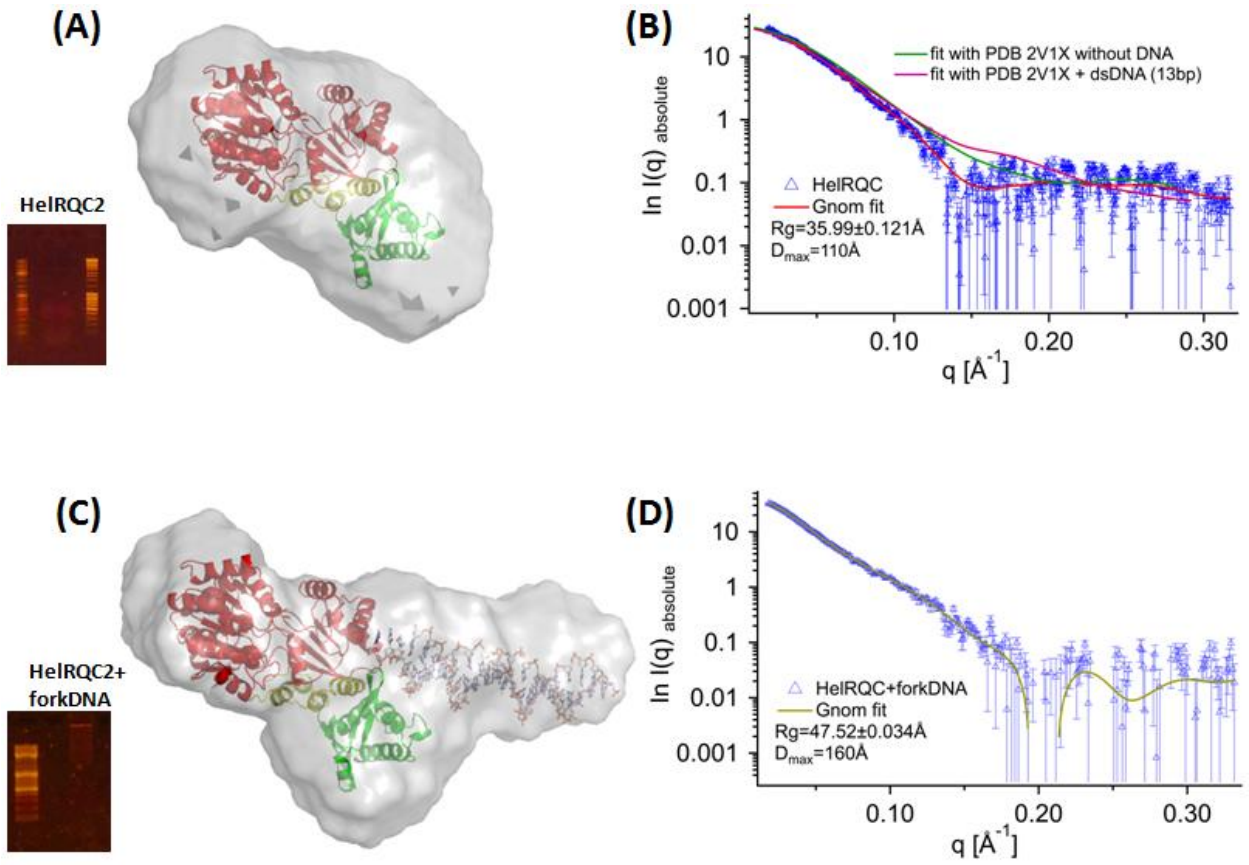

Figure 3.18. SAXS analysis of the catalytic core of human RecQ4 (HelRQC2 and HelRQC2+forkDNA). (A) Low resolution envelope with D_{max}=110Å of HelRQC2 (gray, vdw radius set to 5Å) averaged after 30 DAMMIF runs fitted with the RecQ1 crystal structure (PDB 2V1X) after removing DNA. The helicase domain of RecQ1 crystal structure is in red, the Zn-binding domain is in yellow and the winged helix domain is green. (lower left) Agarose gel of sample HelRQC2 with marker confirms the absence of DNA. (B) Experimental SAXS profile of HelRQC2 with standard deviation (blue triangle) and its best fit with GNOM (red line). Theoretical scattering curves simulated from the crystal structure of RecQ1 (PDB 2V1X) after removing DNA (green line) and from the original crystal structure with 13bp-dsDNA (PDB 2V1X) (magenta line). (C) Low resolution envelope with D_{max}=150Å of HelRQC+forkDNA (gray, vdw radius set to 5Å) averaged after 30 DAMMIF runs fitted with the RecQ1 crystal structure (PDB 2V1X) modeled with 21bp-dsDNA. The helicase domain of RecQ1 crystal structure is in red, the Zn-binding domain is in yellow and the winged helix domain is green. (lower left) Agarose DNA gel of sample HelRQC+forkDNA with marker confirms the presence of DNA. (D) Experimental SAXS profile of HelRQC+forkDNA with standard deviation (light violet triangle) and its best fit with GNOM (lemon line).

Based on, the fitting of the dsDNA model in the elongation region of both HelRQC+forkDNA and HelRQC1 models, the agarose gels and the earlier reports on the role RQC domain in duplex DNA binding and unwinding, we confirm the presence of DNA in these models and present a low resolution structure of RecQ4 bound to DNA. However, it is worth mentioning that the

SAXS data obtained for the last two samples (HelRQC2 and HelRQC2+forkDNA) were not in optimal conditions and because of this the resolution is much lower (30 – 35 Å). Due to heavy subscription of the beamline and time constraint no other data were obtained.

4. Conclusions and Future work

4.1. Conclusions

Despite the fact that N-terminal region of human RecQ4 has a role in initiation of DNA replication and is required for cell viability, and germline mutations in human *RECQ4* gene are associated to three autosomal recessive disorders and cancer, not much is known about the structural and biochemical aspects of the protein. Previous studies suggested that human RecQ4 has a well conserved helicase domain and an N-terminus homologous to the yeast replication factor Sld2 (Sangrithi et al., 2005). The only structure known for the protein is the NMR structure of the first 54 amino acids at the N-terminus of human RecQ4 which forms a helical bundle resembling a homeodomain (Ohlenschläger et al., 2012). A recent bioinformatic study done in our lab (Marino et al., 2013), predicted a second region of homology with Sld2 (in addition to the N-terminal 150 amino-acid residues), a cysteine-rich region classified as "retrovirus Zn finger like" or "Zn knuckle", located between the Sld2 homology region and the helicase domain.

Moreover, a puzzling feature of RecQ4 was the lack of a RQC domain, which is an essential pat of the catalytic core of all the RecQ helicases that have been characterize. As the initial reports (Macris et al., 2006) seemed to suggest that RecQ4 had no helicase activity, the absence of a RQC was blamed. However, more recently, several laboratories (Suzuki et al., 2009; Rossi et al., 2010) have shown a weak but reproducible helicase activity. Consistent with that, the presence of a RQC domain has been suggested by a bioinformatic analysis (Marino et al., 2013); such region contains in RecQ4 significant divergence, including two long insertions, which escaped the detection of automatic algorithms.

Based on these results several constructs were designed with different domain boundaries and an extensive cloning, expression and purification optimization was done on these constructs. Out of several constructs only three constructs (one comprising the N-terminal Zn knuckle, second construct encompassing the additional Sld2 homologous region with the Zn knuckle and the third consisting of the Helicase and RQC domain, Figure 3.1) were promising enough to extend to large scale expression and purification and furthermore to structural and biochemical characterization of the protein.

4.1.1. The N-terminal region

The work on the N-terminal region of human RecQ4 was focused on the structural and biochemical characterization of the predicted Zn-knuckle region and the role of the additional Sld2 homologous upstream region. The summary of this study is as follows:-

1. Although no folding was shown by the 25 amino acid long peptide comprising the non-canonical Zn-knuckle (Zn ligands: CNHC) corresponding to the N-terminal region of human RecQ4 in the presence of zinc or nucleic acid, its *Xenopus* counterpart (CCHC) folded on zinc addition.
2. A 3D NMR structure of *Xenopus* peptide was determined where three cysteines (Cys615, Cys 618 and Cys628) and one histidine (His623) coordinated one Zn^{2+} ion forming an α-helix between residues 624-628 and an antiparallel β-sheet, formed by two short β-strands (613-616 and 620-622).
3. The purified construct comprising of the GST tagged non-canonical Zn-knuckle (CNHC) region of human RecQ4 showed affinity towards both DNA and RNA with a preference towards forked substrates, while the *Xenopus* counterpart with the canonical motif (CCHC) binds only RNA substrates. The substitution of cysteine with asparagine might make the Zn-knuckle fold less tight and more adaptable towards various nucleic acid substrates.
4. A much higher affinity towards all the nucleic acid substrates was shown by the additional Sld2 homologous region upstream the Zn knuckle (from both species) with a preference towards forked and RNA substrates, compared to that of the Zn-knuckle region alone.
5. A decrease in the binding affinity was shown by the protein towards DNA but not RNA when the putative residues (Cys403, Asn406, His411 and Cys416) were mutated to alanine. To confirm the results, EDTA was used to chelate the metal ions in nucleic acid binding reactions which provided similar results. The mutant human protein in which Asn406 is replaced to cysteine to reconstruct the canonical "CCHC" motif behaves similar to other mutant proteins.

6. These results suggest that the additional Sld2 homologous region upstream of the Zn-knuckle plays an important role in nucleic acid binding and that the Zn-knuckle works synergistically with the upstream region.
7. The high affinity of these N-terminal regions towards RNA allow us to speculate the role of RecQ4 in various cellular processes involving RNA such as in initiation of DNA replication in higher eukaryotes, transcriptional/translational regulation and DNA repair, as discussed in section 3.2.5.

4.1.2. The catalytic core

As mentioned earlier, a recent bioinformatic study predicted the presence of RQC domain in RecQ4. To validate this prediction, an intensive cloning, expression and purification trials were done. Finally, the catalytic core of human RecQ4 comprising of helicase and the putative RQC domain (HelRQC, 445-1112) was purified, described in section 3.1.2. The structural and biochemical study of this fragment done in this work is summarized below:-

1. In the bioinformatic study various residues were predicted to be present within the RQC domain and play a significant role in the biochemical activity of the protein. As discussed in section 3.3.4, the residues including cysteines (Cys853, Cys855, Cys897, Cys925, Cys945 and Cys963) in the Zn binding domain, the aromatic residue (Phe1077) at the tip of the β-hairpin motif and the WalkerA lysine (Lys508) were mutated to alanine and the mutant proteins were checked for the biochemical assays parallel to the wild type protein.
2. The mutant proteins behaved similarly to the wild type protein in the purification and showed a similar folding with a melting temperature comparable with that of wild type. Thus, the mutations did not disrupt the folding of the protein.
3. The wild type protein showed the presence of two mole of Zn^{2+} for each mole of protein, indicating the presence of two Zn clusters within the catalytic core of human RecQ4. The mutated Phe1077 and Lys508 had a similar molar ratio of 1:2 of protein: Zn^{2+}, while the cysteine mutants show a loss of one Zn^{2+}. This result verifies the role of the predicted cysteines in Zn^{2+} binding and advocates the presence of the RQC Zn domain.
4. The recombinant wild type protein showed a higher binding affinity towards the forked DNA substrate compared to single stranded and blunt substrates. DNA binding activity

was lost by all the mutants except the WalkerA mutant (Lys508Ala), which is in-line with the earlier reports of the Zn-binding domain and the WH motif from other RecQ helicases to be involved in DNA binding.

5. The forked DNA substrate is the best substrate that stimulates the ATPase activity of the wild type protein. All the mutant proteins with mutated Zn ligands and aromatic β-hairpin residue had an equivalent activity, while the Walker A mutant (Lys508Ala) showed a loss of ATPase activity. This suggests that the ATPase activity is independent of the RQC domain in RecQ4.
6. The favored substrate to unwind for wild type protein was forked DNA substrate, while all the mutant proteins did not show any helicase activity. This confirms that RecQ4 displays an ATP-dependent DNA unwinding, and indicates the importance of Zn^{2+} binding for the helicase activity, as well as validates the role of phenylalanine at the tip of the β-hairpin in DNA unwinding.
7. Three SAXS models of the catalytic core of human RecQ4 were obtained which are in agreement with the 3D crystal structure of RecQ1 (2V1X/2WWY). The first sample was contaminated by nucleic acid, as shown by a long extension appeared in the 3D reconstruction, that was consistent with the position of DNA in the RecQ1-DNA complex (2WWY). When we further purify the protein from the contaminated DNA, we obtained a more compact structure; addition of exogenous fork substrates revealed the re-appearance of an extension compatible with the RecQ1-DNA model. However, these two latter data sets were collected in non-optimal conditions, due to the heavy subscription of the SAXS beamline, and are at very low resolution.
8. Besides carrying out a biochemical characterization of RecQ4, in this work we have definitely confirmed the prediction of a RQC domain in RecQ4 protein. These results enhance the knowledge about human RecQ4 and provide a base of a detailed study to dissect the mechanism of action of the protein in the cell.

4.2. Future work

These results give us a base to extend our research for a detailed structural and biochemical characterization of human RecQ4. We therefore want to continue our studies performing the following experiments:-

1. Carry out further mutagenesis study of N-terminal region, to test the prediction on the mode of nucleic acid binding from other Zn knuckle, and to assess the role of highly conserved residues within the Sld2 upstream region.
2. Despite extensive crystallization experiments we have been unable to obtain crystal. An obvious reason is the presence of two long and flexible insertions within the RQC Zn domain, which may interfere with lattice formation. We plan to eliminate by mutagenesis the long extra loops, leaving the Zn-ligands intact to produce a more compact protein, ameanable to crystallization. An alternative approach would be to co-crystallise with limited amounts of proteases *in situ* (Dong et al., 2007; Wernimont and Edwards, 2009).
3. Cloning, expressing and purifying other regions of RecQ4 such as the winged helix domain alone and the C-terminus, and trying a combination of various structure determination techniques such as crystallography, NMR and SAXS to characterize it better.
4. Further characterize biochemically the catalytic core using various substrates like R-loops, G-quadruplexes, D-loops, Holliday junctions, etc.
5. Carry out a detailed biochemical study of the annealing activity of RecQ4, to dissect the role of RecQ4 in annealing various substrates.
6. Identify the potential interaction partners of the protein. The laboratory can provide a large number of proteins involved in human DNA replication, including MCM2-7, MCM89, MCM10, Cdc45, GINS, AND-1, Cdt1, etc. As for substrates involved in DNA repair we will rely on collaborations with colleagues working in the field.

Appendix

A1. Reagents, buffers and solutions

Reagents used:

PfuTurbo DNA polymerase and PfuTurbo buffer were obtained from Stratagene. dNTPs, restriction endonucleases, restriction buffers, BSA (Bovine Serum Albumin) for restriction digestions and Calf Intestinal Phosphatase (CIP) were from New England Biolabs. The Rapid DNA Ligation Kit was purchased from Roche. Isopropyl β-D-1-thiogalactopyranoside or IPTG (dioxan free) and Dithiothreitol or DTT were from Melford Laboratories. Lysozyme (from chicken egg white), glutathione (reduced form), thrombin (lyophilised powder), PMSF (phenylmethanesulfonylfluoride), AEBSF (4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride), βME (beta mercaptoethanol), BSA (for Western Blots) and gel-filtration molecular weight markers were all purchased from Sigma.

Benzonase was from Merck. All other buffers, salts and reagents used for the preparation of solutions were purchased from Sigma.

Media and solutions:

LB medium (LB): LB powder (Sigma) was dissolved in de-ionized water (20 g per L of water) and autoclaved for 20 min at 121ºC. The medium was stored at room temperature and the necessary antibiotics were added to the specified final concentrations prior to use.

Terrific broth (TB): Terrific broth (Modified, Sigma) was dissolved in de-ionized water (47 g per L of water) and 8 ml glycerol per L of medium was also added prior to autoclaving (20 min at 121ºC). The medium was stored at room temperature and the necessary antibiotics were added to the specified final concentrations prior to use.

Autoinduction media (LBP-5052): it is a rich media for autoinduction. For 1L: 928 mL of LB, 1mL of 1M MgSO4, 20 mL of 50X 5052 and 50 mL od 20X NPS, antibiotic as needed (to avoid precipitation, 1M MgSO4 needs to be added before 20X NPS). Stock solutions are so composed: 20X NPS (NPS= 100mM PO4, 25mM SO4, 50mM NH4, 100mM Na, 50mM K). To make 100 ml: 90 ml water, 6.6 g of (NH4)2 SO4 0.5M, 13.6 g KH2PO4 1M, 14.2 g Na2HPO4 1M.

Solutions were added in sequence in a beaker and stirred until dissolved. The pH of a 20-fold dilution in water should be ~6.75.

50X 5052 (5052= 0.5% glycerol, 0.05% glucose, 0.2% a-lactose). To make 100 ml: 25 g glycerol (weigh in beaker), 73 ml water, 2.5 g glucose, 10 g α-lactose. Reagents were added in sequence in a beaker and stirred until dissolved. Process can be speeded up by heating in microwave. All stock solutions must be autoclaved for 20 minutes at 121°C.

LB Agar and agar plates: 5.25 g of LB agar powder (Sigma) were added to 150 ml de-ionized water and autoclaved for 20 min at 121ºC. The medium was allowed to cool to 50ºC before adding antibiotics to the specified final concentrations. 20 ml of medium were poured into 85 mm petri dishes. If necessary, the surface of the medium was flamed with a Bunsen burner to eliminate bubbles. The medium was left to harden for one hour and the agar plates were stored at 4ºC for ≤ 1 month.

Antibiotic stock solutions (1000x):

Kanamycin: 50 mg of kanamycin monosulphate (Melford Laboratories) were dissolved in 10 ml deionised water, filter-sterilised and stored at -20ºC. Kanamycin was added to growth medium to a final concentration of 50 μg/ml.

Chloramphenicol: 340 mg chloramphenicol (Sigma) were dissolved in 10 ml ethanol and the aliquots were wrapped in foil and stored at -20ºC. Chloramphenicol was added to growth medium to a final concentration of 34 μg/ml.

Buffers used for the preparation of competent *E. coli* cells:

TFB1: 100 mM RbCl, 50 mM MnCl2, 30 mM potassium acetate, 10 mM CaCl2, 15% glycerol, pH 5.8 (adjusted with acetic acid). The solution was filter-sterilized and stored at 4ºC.

TFB2: 10 mM MOPS, 10 mM RbCl, 75 mM CaCl2, 15% glycerol, pH 6.8 (adjusted with KOH). The solution was filter-sterilized and stored at 4ºC.

Buffers used for tag removal from proteins:

SUMO cleavage buffer: 25 mM Tris-HCl pH 8.0, 10 % glycerol

DNA agarose gels:

6x DNA loading Buffer: 50% (v/v) glycerol, 100 mM Na2EDTA·2H2O, 1% (w/v) SDS, 0.1% (w/v) bromophenol blue

10x TBE: 890 mM Tris Base, 890 mM Boric Acid, 20 mM EDTA pH 8

SDS-PAGE buffers:

2x SDS Sample Buffer: 0.125 M Tris-HCl, pH 6.8, 20% (v/v) glycerol, 0.04% (w/v) SDS, 10% (v/v) βME, 0.1% (w/v) bromophenol blue

6x SDS Sample Buffer: 0.35 M Tris-HCl pH 6.8, 10% SDS, 30% glycerol, 9.3% DTT

Tris Glycine SDS Separating gel solution: 0.375 M Tris-HCl, pH 8.8, 0.001% lauryl sulphate (SDS), , 10-15% Acrylamide/Bis, 0.0005% ammonium persulphate (APS) and 0.0001% N,N,N‘‘,N‘‘-tetramethylethylenediamine (TEMED).

Tris Glycine SDS Stacking gel solution: 0.125 M Tris-HCl, pH 6.8, 0.001% SDS, , 4% Acrylamide/Bis, 0.0005% APS and 0.0001% TEMED.

10x Tris-glycine Buffer: 0.25 M Tris Base, 1.92 M Glycine, 10% SDS. The buffer was diluted 10-fold to use for electrophoresis.

Primers: The cloning primers used in this study were supplied as desalted, lyophilized powders (Sigma-Genosys, UK). Primers were resuspended in deionised water to a final concentration of 100 μM. A further dilution was made to produce a working stock of 10 μM. Stock solutions were stored at -20ºC.

IPTG: IPTG powder (Carbosynth) was dissolved in deionised water to a final concentration of 1 M. The solution was filter sterilised, wrapped in foil and stored at -20ºC.

PMSF: PMSF (Sigma Aldrich) powder was dissolved in ethanol to a final concentration of 100 mM. The stock was stored at 4ºC.

AEBSF: PMSF (Sigma Aldrich) powder was dissolved in water to a final concentration of 200 mM. The stock was stored at -20ºC.

DTT: DTT powder (Sigma Aldrich) was dissolved in deionised water to a final concentration of 1 M. The solution was filter-sterilised, divided in 1.5 ml aliquots and stored at -20ºC.

A3. Experimental protocols

Preparation of competent *E. coli* cells

The following protocol was used for the preparation of the *E. coli* strains from commercial original stocks. All media used for the preparation of competent cells were not supplemented with antibiotics. However 34 μg/ml of chloramphenicol was added to some *E. coli* strains when needed according to manufacturer's instructions. The compositions of buffers TFB1 and TFB2 that were used in this protocol are provided in this Appendix. A trace of competent cells was removed from a stock vial with a sterile tip, streaked out on an agar plate and incubated at 37ºC overnight. Next, a single colony from the agar plate was inoculated in 10 ml LB medium and the culture was grown overnight at 37ºC with shaking at approximately 225 rpm. On the following day, 1 ml overnight culture was used to inoculate 100 ml prewarmed LB medium in a 250 ml flask. The cells were grown until an A600nm of 0.5 was reached. The culture was cooled on ice for 5 min and then transferred to a sterile 50 ml Falcon tube. The cells were pelleted by centrifugation at 4000 g for 5 min at 4ºC. The supernatant was carefully decanted and the cells were resuspended gently in 30 ml cold (4ºC) TFB1 buffer and incubated on ice for 90 min. The cells were collected by centrifugation at 4000 g for 5 min at 4ºC and the supernatant was carefully decanted. Using chilled pipettes, the cells were gently resuspended in 4 ml ice-cold TFB2 buffer. 200 μl aliquots of cells were prepared using prechilled sterile tubes. The aliquots were then frozen in a dry-ice/ethanol bath and stored at -80ºC.

A4. Vector Maps

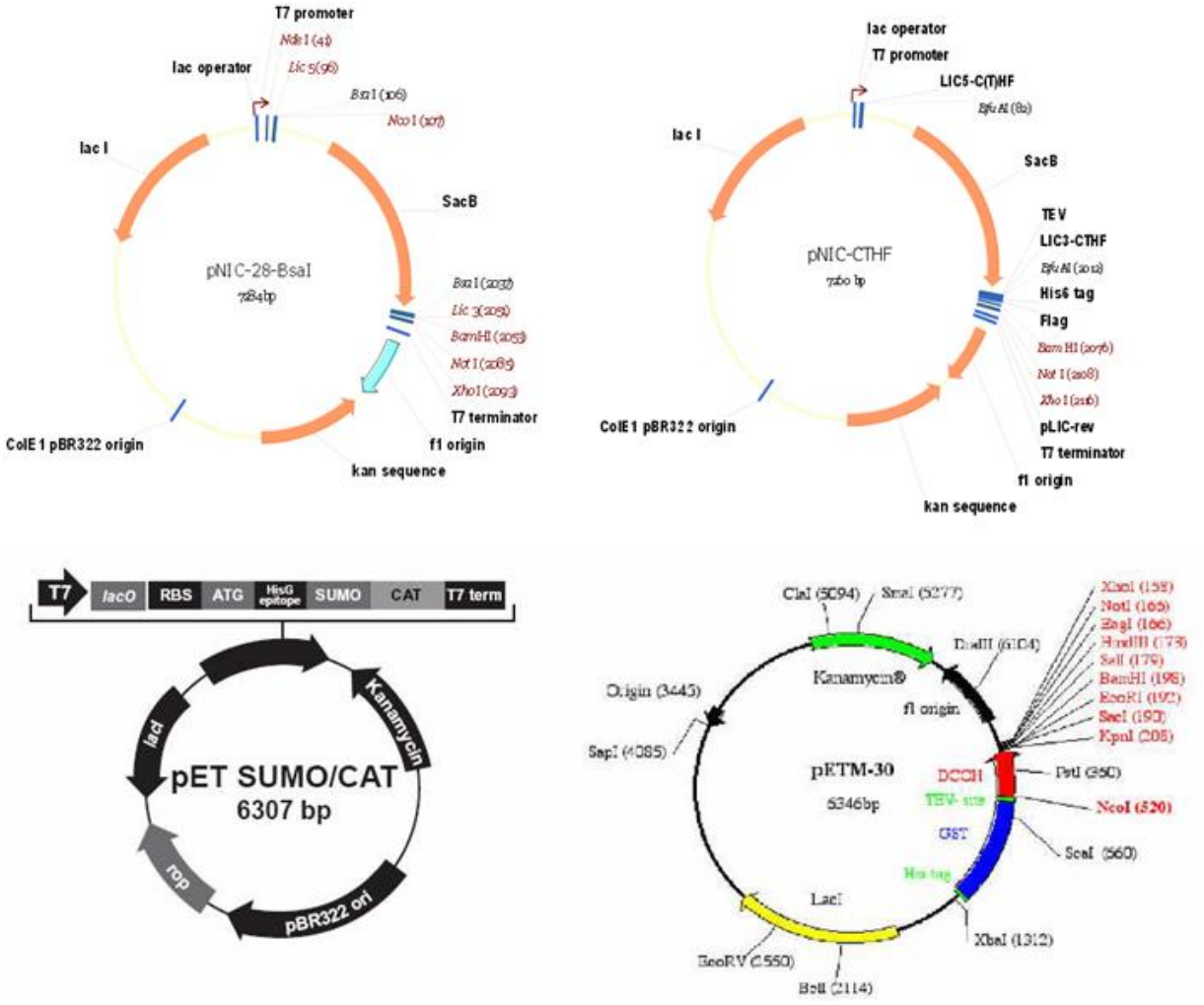

Figure A1 Map of the expression vectors used in this study. The following vectors were kindly provided by Dr Opher Gileadi from the Structural Genomics Consortium (SGC, Oxford): pNIC-Bsa4 and pNIC-CTHF. pET SUMO/CAT is from Invitrogen and pETM-30 was a kind gift from EMBL.

References

Abe, T., Yoshimura, A., Hosono, Y., Tada, S., Seki, M., Enomoto, T. (2011) The N-terminal region of RECQL4 lacking the helicase domain is both essential and sufficient for the viability of vertebrate cells. Role of the N-terminal region of RECQL4 in cells. *Biochim. Biophys. Acta*. 1813, 473-479.

Aggarwal M, Sommers JA, Morris C, Brosh RM Jr. (2010) Delineation of WRN helicase function with EXO1 in the replicational stress response. *DNA Repair* 9:765–76

Aggarwal M, Sommers JA, Shoemaker RH, Brosh RMJr. (2011) Inhibition of helicase activity by a small molecule impairs Werner syndrome helicase (WRN) function in the cellular response to DNA damage or replication stress. *Proc. Natl. Acad. Sci. USA* 108:1525–30

Aguilera, A. and Garcia-Muse, T. (2012) R loops: from transcription byoproducts to threats to genome stability. *Mol Cell* 46, 115-124

Ahn B, Harrigan JA, Indig FE, Wilson DM 3rd, Bohr VA. (2004) Regulation of WRN helicase activity in human base excision repair. *J Biol Chem*. Dec 17; 279(51):53465-74

Ahnert, P. and S. S. Patel (1997) Asymmetric interactions of hexameric bacteriophage T7 DNA helicase with the 5'- and 3'-tails of the forked DNA substrate. *J Biol Chem* 272(51): 32267-32273.

Ammazzalorso F, Pirzio LM, Bignami M, Franchitto A, Pichierri P. (2010) ATR and ATM differently regulate WRN to prevent DSBs at stalled replication forks and promote replication fork recovery. *EMBO J*. 29:3156–3169.

Amenitsch H., Bernstorff S., Kriechbaum M., Lombardo D., Mio H., Rappolt M., Laggner P. (1997) Performance and first results of the ELETTRA high-flux beamline for small-angle X-ray scattering. *J. Appl. Crystallogr.* 30 (5), 872-876

Andersen CB, Ballut L, Johansen JS, Chamieh H, Nielsen KH, Oliveira CL, Pedersen JS, Séraphin B, Le Hir H, Andersen GR. (2006) Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. *Science*. Sep 29;313(5795):1968-72

Ashworth A, Lord CJ, Reis-Filho JS. (2011) Genetic interactions in cancer progression and treatment. *Cell*.145(1):30–38.

Audebert M, Salles B, Calsou P. (2004) Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining. *J. Biol. Chem.* 279:55117–26

Aygun, O., J. Svejstrup, et al. (2008) A RECQ5-RNA polymerase II association identified by targeted proteomic analysis of human chromatin. *Proc Natl Acad Sci U S A* 105(25): 8580-8584.

Bachrati, C. Z., R. H. Borts, et al. (2006) Mobile D-loops are a preferred substrate for the Bloom's syndrome helicase. *Nucleic Acids Res* 34(8): 2269-2279.
Bachrati, C. Z. and I. D. Hickson (2003) RecQ helicases: suppressors of tumorigenesis and premature aging. *Biochem J* 374(Pt 3): 577-606.
Balajee, A. S., A. Machwe, et al. (1999) The Werner syndrome protein is involved in RNA polymerase II transcription. *Mol Biol Cell* 10(8): 2655-2668.
Barea F., Tessaro S., Bonatto D. (2008) In silico analyses of a new group of fungal and plant RecQ4-homologous proteins. *Comput Biol Chem.* Oct;32(5):349-58.

Barefield C, Karlseder J. (2012) The BLM helicase contributes to telomere maintenance through processing of late-replicating intermediate structures. *Nucleic Acids Res.* 40:7358–67

Baynton K, Otterlei M, Bjørås M, von Kobbe C, Bohr VA, Seeberg E. (2003) WRN interacts physically and functionally with the recombination mediator protein RAD52. *J Biol Chem*. Sep 19; 278(38):36476-86

Beamish H., Kedar P., Kaneko H., Chen, P., Fukao, T., Peng, C., Beresten, S., Gueven, N., Purdie, D., Lees-Miller, S., et al. (2002) Functional link between BLM defective in Bloom's syndrome and the ataxia-telangiectasia-mutated protein, ATM. *The Journal of boil. chem*. 277, 30515-30523.

Bennett, R. J. and J. L. Keck (2004) Structure and function of RecQ DNA helicases. *Crit Rev Biochem Mol Biol* 39(2): 79-97.

Bennett, R. J., J. L. Keck, et al. (1999) Binding specificity determines polarity of DNA unwinding by the Sgs1 protein of S. cerevisiae. *J Mol Biol* 289(2): 235-248.

Bennett, R. J., J. A. Sharp, et al. (1998) Purification and characterization of the Sgs1 DNA helicase activity of Saccharomyces cerevisiae. *J Biol Chem* 273(16): 9644-9650.

Bernstein DA, Keck JL. (2003) Domain mapping of Escherichia coli RecQ defines the roles of conserved N- and C-terminal regions in the RecQ family. *Nucleic Acids Res*. 31(11):2778-85

Bernstein DA, Keck JL. (2005) Conferring substrate specificity to DNA helicases: role of the RecQ HRDC domain. *Structure*. Aug;13(8):1173-82.

Bernstein DA, Zittel MC, Keck JL. (2003) High-resolution structure of the E.coli RecQ helicase catalytic core. *EMBO J.* Oct 1;22(19):4910-21.

Berti M, Chaudhuri AR, Thangavel S, Gomathinayagam S, Kenig S, et al. (2013) Human RECQ1 promotes restart of replication forks reversed by DNA topoisomerase I inhibition. *Nat. Struct. Mol. Biol.* 20:347–54

Blander, G., Zalle, N., Daniely, Y., Taplick, J., Gray, M.D., and Oren, M. (2002) DNA damageinduced translocation of the Werner helicase is regulated by acetylation. *The Journal of biological chemistry* 277, 50934-50940.

Bloom, D. (1954) Congenital telangiectatic erythema resembling lupus erythematosus in dwarfs; probably a syndrome entity. *AMA Am J Dis Child* 88(6): 754-758.

Blundred, R., K. Myers, et al. (2010) Human RECQL5 overcomes thymidine-induced replication stress. *DNA Repair (Amst)* 9(9): 964-975.

Bochman ML, Paeschke K, Chan A, Zakian VA. (2014) Hrq1, a homolog of the human RecQ4 helicase, acts catalytically and structurally to promote genome integrity. *Cell Rep*. Jan 30;6(2):346-56.

Böhm S, Bernstein KA. (2014) The role of post-translational modifications in fine-tuning BLM helicase function during DNA repair. *DNA Repair*. 22:z123–z132.

Bohr VA. (2008) Rising from the RecQ-age: the role of human RecQ helicases in genome maintenance. *Trends Biochem Sci*. Dec; 33(12):609-20.

Bothmer A, Robbiani DF, Feldhahn N, Gazumyan A, Nussenzweig A, Nussenzweig MC. (2010) 53BP1 regulates DNA resection and the choice between classical and alternative end joining during class switch recombination. *J. Exp. Med.* 207:855–65

Brosh RM Jr. (2013) DNA helicases involved in DNA repair and their roles in cancer. *Nat. Rev. Cancer* 13:542–58

Brosh RM Jr, Orren DK, Nehlin JO, Ravn PH, Kenny MK, Machwe A, Bohr VA. (1999) Functional and physical interaction between WRN helicase and human replication protein *A J Biol Chem*. Jun 25;274(26):18341-50.

Brosh, R.M., Jr., Li, J.L., Kenny, M.K., Karow, J.K., Cooper, M.P., Kureekattil, R.P., Hickson, I.D., and Bohr, V.A (2000) Replication protein A physically interacts with the Bloom's syndrome protein and stimulates its helicase activity. *The Journal of biological chemistry* 275, 23500-23508.

Brosh, R. M., Jr., A. Majumdar, et al. (2001) Unwinding of a DNA triple helix by the Werner and Bloom syndrome helicases. *J Biol Chem* 276(5): 3024-3030.

Brosh, R.M., Jr., von Kobbe, C., Sommers, J.A., Karmakar, P., Opresko, P.L., Piotrowski, J., Dianova, I., Dianov, G.L., and Bohr, V.A (2001) Werner syndrome protein interacts with human flap endonuclease 1 and stimulates its cleavage activity. *The EMBO journal* 20, 5791-5801.

Brough R, Frankum JR, Costa-Cabral S, Lord CJ, Ashworth A. (2011) Searching for synthetic lethality in cancer. *Curr Opin Genet Dev*.;21:34–41.

Bruck, I., Kanter, D.M., Kaplan, D.L. (2011) Enabling association of the GINS protein tetramer with the mini chromosome maintenance (Mcm) 2–7 protein complex by phosphorylated Sld2 protein and single-stranded origin DNA. *J Biol Chem.* 286: 36414–36426.

Brunger A.T. (2007) Version 1.2 of the Crystallography and NMR System, *Nature Protocols* **2**, 2728-2733.

Bugreev, D. V., R. M. Brosh, Jr., et al. (2008) RECQ1 possesses DNA branch migration activity. *J Biol Chem* 283(29): 20231-20242.

Bugreev DV, Yu X, Egelman EH, Mazin AV. (2007) Novel pro- and anti-recombination activities of the Bloom's syndrome helicase. *Genes Dev.* 21:3085–94

Bunting SF, Callen E, Wong N, Chen HT, Polato F, et al. (2010) 53BP1 inhibits homologous recombination in *Brca1*-deficient cells by blocking resection of DNA breaks. *Cell* 141:243–54

Buttner K, Nehring S and Hopfner KP. (2007) Structural basis for DNA duplex separation by a superfamily-2 helicase, *Nat Struct Mol Biol* 14; 647-52.

Byrd, A. K. and K. D. Raney (2005) Increasing the length of the single-stranded overhang enhances unwinding of duplex DNA by bacteriophage T4 Dda helicase. *Biochemistry* 44(39): 12990-12997.

Calado R, Young N. (2012) Telomeres in disease. *F1000 Med. Rep.* 4:8

Capp C, Wu J, Hsieh TS. (2009) Drosophila RecQ4 has a 3'-5' DNA helicase activity that is essential for viability. *J Biol Chem*. Nov 6; 284(45):30845-52

Cejka, P. and S. C. Kowalczykowski (2010) The full-length Saccharomyces cerevisiae Sgs1 protein is a vigorous DNA helicase that preferentially unwinds holliday junctions. *J Biol Chem* 285(11): 8290-8301.

Chakraborty, P. and Grosse, F. (2010) WRN helicase unwinds Okazaki fragment-like hybrids in a reaction stimulated by the human DHX9 helicase. *Nuc. Acids Res*. 38: 4722-4730.

Chan K.-M., Delfert D., Junger K. D. (1986) A direct colorimetric assay for Ca^{2+}-stimulated ATPase activity. *Anal. Biochem.* 157:375–380

Chang, C. C., I. C. Kuo, et al. (2004) Detection of quadruplex DNA structures in human telomeres by a fluorescent carbazole derivative. *Anal Chem* 76(15): 4490-4494.

Chaganti, R. S., S. Schonberg, et al. (1974) A manyfold increase in sister chromatid exchanges in Bloom's syndrome lymphocytes. *Proc Natl Acad Sci U S A* 71(11): 4508-4512.

Cheng WH, von Kobbe C, Opresko PL, Arthur LM, Komatsu K, Seidman MM, Carney JP, Bohr VA. (2004) Linkage between Werner syndrome protein and the Mre11 complex via Nbs1. *J Biol Chem*. May 14 ; 279(20):21169-76

Cheng WH, Sakamoto S, Fox JT, Komatsu K, Carney J, Bohr VA. (2005) Werner syndrome protein associates with γH2AX in a manner that depends upon Nbs1. *FEBS Lett.* 579:1350–56

Cheng WH, Kusumoto R, Opresko PL, Sui X, Huang S, Nicolette ML, Paull TT, Campisi J, Seidman M, Bohr VA (2006) Collaboration of Werner syndrome protein and BRCA1 in cellular responses to DNA interstrand cross-links. *Nucleic Acids Res*. May 19; 34(9):2751-60

Cheok C.F., Wu L., Garcia P.L., Janscak P. and Hickson I.D., (2005) The Bloom's syndrome helicase promotes the annealing of complementary single-stranded DNA, *Nucleic Acids Res* 33

Chester, N., Kuo, F., Kozak, C., O'Hara, C.D., and Leder, P. (1998) Stage-specific apoptosis, developmental delay, and embryonic lethality in mice homozygous for a targeted disruption in the murine Bloom's syndrome gene. *Genes & development* 12, 3382-3393.

Choi DH, Min MH, Kim MJ, Lee R, Kwon SH, Bae SH. (2014) Hrq1 facilitates nucleotide excision repair of DNA damage induced by 4-nitroquinoline-1-oxide and cisplatin in Saccharomyces cerevisiae. *J Microbiol*. Apr;52(4):292-8.

Ciccia A, Elledge SJ. (2010) The DNA damage response: making it safe to play with knives. *Mol. Cell* 40:179–204

Clark, K. L., E. D. Halay, et al. (1993) Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature* 364(6436): 412-420.

Collins R, Karlberg T, Lehtiö L, Schütz P, van den Berg S, Dahlgren LG, Hammarström M, Weigelt J, Schüler H. (2009) The DEXD/H-box RNA helicase DDX19 is regulated by an {alpha}-helical switch. *J Biol Chem.* Apr 17;284(16):10296-300

Connolly, B., C. A. Parsons, et al. (1991) Resolution of Holliday junctions in vitro requires the Escherichia coli ruvC gene product. *Proc Natl Acad Sci U S A* 88(14): 6063-6067.

Constantinou A, Tarsounas M, Karow J.K, Brosh R.M, Bohr V.A., Hickson I.D., West S.C. (2000) Werner's syndrome protein (WRN) migrates Holliday junctions and co-localizes with RPA upon replication arrest, *EMBO* Rep 1 80-4

Cooper MP, Machwe A, Orren DK, Brosh RM, Ramsden D, Bohr VA. (2000) Ku complex interacts with and stimulates the Werner protein. *Genes Dev.* 14:907–12

Cordin O, Banroques J, Tanner NK, Linder P. (2006) The DEAD-box protein family of RNA helicases. *Gene* 367:17-37

Costa A. and Onesti S. (2009) Structural biology of MCM helicases. *Crit. Rev. Biochem. Mol. Biol.* 44, 326-342.

Cox, M. M., M. F. Goodman, et al. (2000) The importance of repairing stalled replication forks. *Nature* 404(6773): 37-41.

Croteau, DL., Popuri, V., Opresko, PL., Bohr, VA., (2014) Human RecQ Helicases in DNA Repair, Recombination, and Replication. *Annu. Rev. Biochem.* 83: 9.1-9.34.

Croteau DL, Rossi ML, Canugovi C, Tian J, Sykora P, Ramamoorthy M, Wang Z, Singh DK, Akbari M, Kasiviswanathan R, Copeland WC, Bohr VA. (2012) RECQL4 localizes to mitochondria and preserves mitochondrial DNA integrity. *Aging Cell*. Feb 1.

Cui, S., Arosio, D., Doherty, K.M., Brosh, R.M., Jr., Falaschi, A., Vindigni, A. (2004) Analysis of the unwinding activity of the dimeric RECQ1 helicase in the presence of human replication protein A. *Nucleic acids research* 32, 2158-2170.

d'Adda di Fagagna, F. (2014) A direct role for small non-coding RNAs in DNA damage response. *Trends Cell Biol*. 24, 171-178.

D'Souza, V. and Summers, M.F. (2004) Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus. *Nature* 431: 586-590.

Das A, Boldogh I, Lee JW, Harrigan JA, Hegde ML, Piotrowski J, de Souza Pinto N, Ramos W, Greenberg MM, Hazra TK, Mitra S, Bohr VA. (2007) The human Werner syndrome protein stimulates repair of oxidative DNA base damage by the DNA glycosylase NEIL1. *J Biol Chem.* Sep 7;282(36):26591-602

Davies, S.L., North, P.S., Hickson, I.D. (2007) Role for BLM in replication-fork restart and suppression of origin firing after replicative stress. *Nat Struct Mol Biol* 14, 677-679.

De S, Kumari J, Mudgal R, Modi P, Gupta S, Futami K, Goto H, Lindor NM, Furuichi Y, Mohanty D, Sengupta S. (2012) RECQL4 is essential for the transport of p53 to mitochondria in normal human cells in the absence of exogenous stress. *J Cell Sci*. Feb 22.

Dietschy, T., Shevelev, I., Pena-Diaz, J., Huhn, D., Kuenzle, S., Mak, R., Miah, M.F., Hess, D., Fey, M., Hottiger, M.O., et al. (2009) p300-mediated acetylation of the Rothmund-Thomson-syndrome gene product RECQL4 regulates its subcellular localization. *Journal of cell science* . 122, 1258-1267.

Diffley JFX. (2011) Quality control in the initiation of eukaryotic DNA replication. *Philos Trans R Soc Lond B Biol Sci.* Dec 27;366(1584):3545-53. Review

Dinkelmann M, Spehalski E, Stoneham T, Buis J, Wu Y, et al. (2009) Multiple functions of MRN in end-joining pathways during isotype class switching. *Nat. Struct. Mol. Biol.* 16:808–13

Doherty KM, Sharma S, Uzdilla LA, Wilson TM, Cui S, et al. (2005) RECQ1 helicase interacts with human mismatch repair factors that regulate genetic recombination. *J. Biol. Chem.* 280:28085–94

Dong, A., et al. (2007) In situ proteolysis for protein crystallization and structure determination. *Nature Methods* 4:1019.

Dunderdale, H. J., F. E. Benson, et al. (1991) Formation and resolution of recombination intermediates by E. coli RecA and RuvC proteins. *Nature* 354(6354): 506-510.

Dutertre, S., Ababou, M., Onclercq, R., Delic, J., Chatton, B., Jaulin, C., Amor-Gueret, M. (2000) Cell cycle regulation of the endogenous wild type Bloom's syndrome DNA helicase. *Oncogene* 19, 2731-2738.

Eladad, S., Ye, T.Z., Hu, P., Leversha, M., Beresten, S., Matunis, M.J., Ellis, N.A. (2005) Intra-nuclear trafficking of the BLM helicase to DNA damage-induced foci is regulated by SUMO modification. *Human molecular genetics* 14, 1351-1365.

Ellis NA, Groden J, Ye TZ, Straughen J, Lennon DJ, Ciocci S, Proytcheva M., German J. (1995) The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell* 83: 655-666

Enemark EJ, Joshua-Tor L. (2006) Mechanism of DNA translocation in a replicative hexameric helicase. *Nature*. Jul 20;442(7100):270-5.

Epstein, C. J., G. M. Martin, et al. (1966) Werner's syndrome a review of its symptomatology, natural history, pathologic features, genetics and relationship to the natural aging process. *Medicine* (Baltimore) 45(3): 177-221.

Evers B, Helleday T, Jonkers J. (2010) Targeting homologous recombination repair defects in cancer. *Trends Pharmacol Sci*. 31:372–380.

Fan W, Luo J. (2008) RecQ4 facilitates UV light-induced DNA damage repair through interaction with nucleotide excision repair factor xeroderma pigmentosum group A (XPA). *J Biol Chem*. Oct 24;283(43):29037-44

Fattah F, Lee EH, Weisensel N, Wang Y, Lichter N, Hendrickson EA. (2010) Ku regulates the nonhomologous end joining pathway choice of DNA double-strand break repair in human somatic cells. *PLoS Genet.* 6:e1000855

Ferrarelli LK, Popuri V, Ghosh AK, Tadokoro T, Canugovi C, Hsu JK, Croteau DL, Bohr VA. (2013) The RECQL4 protein, deficient in Rothmund-Thomson syndrome is active on telomeric D-loops containing DNA metabolism blocking lesions. DNA Repair (Amst). Jul;12(7):518-28

Flaus A, Owen-Hughes T. (2004) Mechanisms for ATP-dependent chromatin remodeling: farewell to the tuna can octamer?. *Curr. Opin. Genet. Dev.* 14:165-73

Flaus A, Martin DM, Barton GJ, Owen-Hughes T. (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res.* 34:2887-905

Fogh, R. H., G. Ottleben, et al. (1994) Solution structure of the LexA repressor DNA binding domain determined by 1H NMR spectroscopy. *EMBO J* 13(17): 3936-3944.

Franchitto A, Pirzio LM, Prosperi E, Sapora O, Bignami M, Pichierri P. (2008) Replication fork stalling in WRN-deficient cells is overcome by prompt activation of a MUS81-dependent pathway. *J Cell Biol*. 183:241–252.

Franke, D. and Svergun, D. I. (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Cryst.*, 42, 342-346.

Fry, M. and L. A. Loeb (1994) The fragile X syndrome d(CGG)n nucleotide repeats form a stable tetrahelical structure. *Proc Natl Acad Sci U S A* 91(11): 4950-4954.

Futami K, Ishikawa Y, Goto M, Furuichi Y, Sugimoto M. (2008a) Role of Werner syndrome gene product helicase in carcinogenesis and in resistance to genotoxins by cancer cells. *Cancer Sci* 99: 843-848

Gai, D., R. Zhao, et al. (2004) Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. *Cell* 119(1): 47-60.

Gangloff, S., J. P. McDonald, et al. (1994) The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase. *Mol Cell Biol* 14(12): 8391-8398.

Garcia P.L., Liu Y., Jiricny J., West S.C., Janscak P. (2004) Human RECQ5beta, a protein with DNA helicase and strand-annealing activities in a single polypeptide, *Embo J* 23 2882-91

Gajiwala K.S., Chen H., Cornille F., Roques B.P., W. Reith, B. Mach, Burley S.K. (2000) Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding. *Nature* 403 916–921.

German, J. (1993) Bloom syndrome: a mendelian prototype of somatic mutational disease. *Medicine* (Baltimore) 72(6): 393-406.

German, J. (1995) Bloom's syndrome. *Dermatol Clin* 13(1): 7-18.

German J. (1997) Bloom's syndrome. XX. The first 100 cancers. *Cancer Genet Cytogenet*. Jan;93(1):100-6.

German, J., M. M. Sanz, et al. (2007) Syndrome-causing mutations of the BLM gene in persons in the Bloom's Syndrome Registry. *Hum Mutat* 28(8): 743-753.

Ghosh AK, Rossi ML, Singh DK, Dunn C, Ramamoorthy M, Croteau DL, Liu Y, Bohr VA. (2012) RECQL4, the protein mutated in Rothmund-Thomson syndrome, functions in telomere maintenance. *J Biol Chem*. Jan 2;287(1):196-209

Gileadi O, Burgess-Brown NA, Colebrook SM, Berridge G, Savitsky P, Smee CE, Loppnau P, Johansson C, Salah E, Pantic NH. (2008) High throughput production of recombinant human proteins for crystallography. *Methods Mol Biol*. 426:221-46.

Gogol EP, Seifried SE, von Hippel PH. (1991) Structure and assembly of the Escherichia coli transcription termination factor rho and its interaction with RNA. I. Cryoelectron microscopic studies. *J Mol Biol*. Oct 20; 221(4):1127-38

Gorbalenya, A.E., Koonin, E.V. (1993) Helicases: amino acid sequence comparisons and structure-function relationships. *Curr Opin Struct Biol*. 3, 419-429.

Gorbalenya AE, Koonin EV, Wolf YI. (1990) A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. *FEBS Lett*. Mar 12;262(1):145-8

Goss, K.H., Risinger, M.A., Kordich, J.J., Sanz, M.M., Straughen, J.E., Slovek, L.E., Capobianco, A.J., German, J., Boivin, G.P., Groden, J. (2002) Enhanced tumor formation in mice heterozygous for Blm mutation. *Science* 297, 2051-2053.

Goto, M. (1997) Hierarchical deterioration of body systems in Werner's syndrome: implications for normal ageing. *Mech Ageing Dev* 98(3): 239-254.

Gottipati P, Vischioni B, Schultz N, Solomons J, Bryant HE, et al. (2010) Poly(ADP-ribose) polymerase is hyperactivated in homologous recombination-defective cells. *Cancer Res.* 70:5389–98

Grandori C, Wu KJ, Fernandez P, Ngouenet C, Grim J, Clurman BE, Moser MJ, Oshima J, Russell DW, Swisshelm K, Frank S, Amati B, Dalla-Favera R, Monnat RJ Jr. (2003) Werner syndrome protein limits MYC-induced cellular senescence. *Genes Dev*. Jul 1;17(13):1569-74.

Gravel S, Chapman JR, Magill C, Jackson SP. (2008) DNA helicases Sgs1 and BLM promote DNA double-strand break resection. *Genes Dev.* 22:2767–72

Griffith, J. D., L. Comeau, et al. (1999) Mammalian telomeres end in a large duplex loop. *Cell* 97(4): 503-514.

Groocock LM, Prudden J, Perry JJ, Boddy MN. (2012) The RecQ4 orthologue Hrq1 is critical for DNA interstrand cross-link repair and genome stability in fission yeast. *Mol Cell Biol*. Jan;32(2):276-87.

Gu M, Rice CM. (2010) Three conformational snapshots of the hepatitis C virus NS3 helicase reveal a ratchet translocation mechanism *Proc Natl Acad Sci U S A*. Jan 12; 107(2):521-8.

Guo R.B., Rigolet P., Zargarian L., Fermandjian S., Xi X.G. (2005) Structural and functional characterizations reveal the importance of a zinc binding domain in Bloom's syndrome helicase. *Nucleic Acids Res*. 33 3109–3124.

Gyimesi, M., K. Sarlos, et al. (2010) Processive translocation mechanism of the human Bloom's syndrome helicase along single-stranded DNA. *Nucleic Acids Res* 38(13): 4404-4414.

Gyimesi M, Harami GM, Sarlós K, Hazai E, Bikádi Z, Kovács M. (2012) Complex activities of the human Bloom's syndrome helicase are encoded in a core region comprising the RecA and Zn-binding domains. *Nucleic Acids Res*. 40(9):3952-63.

Hanakahi, L. A., H. Sun, et al. (1999) High affinity interactions of nucleolin with G-Gpaired rDNA. *J Biol Chem* 274(22): 15908-15912.

Handa, N., K. Morimatsu, et al. (2009) Reconstitution of initial steps of dsDNA break repair by the RecF pathway of E. coli. *Genes Dev* 23(10): 1234-1245.

Harmon, F. G. and S. C. Kowalczykowski (1998) RecQ helicase, in concert with RecA and SSB proteins, initiates and disrupts DNA recombination. *Genes Dev* 12(8): 1134-1144.

Harrigan, J.A., Wilson, D.M., 3rd, Prasad, R., Opresko, P.L., Beck, G., May, A., Wilson, S.H., and Bohr, V.A (2006) The Werner syndrome protein operates in base excision repair and cooperates with DNA polymerase beta. *Nucleic acids research* 34, 745-7

Harrigan JA, Fan J, Momand J, Perrino FW, Bohr VA, Wilson DM 3rd (2007) WRN exonuclease activity is blocked by DNA termini harboring 3' obstructive groups. *Mech Ageing Dev*. Mar; 128(3):259-66

Helleday T, Petermann E, Lundin C, Hodgson B, Sharma RA. (2008) DNA repair pathways as targets for cancer therapy. *Nat Rev Cancer*. 8:193–204.

Hickman AB, Dyda F. (2005) Binding and unwinding: SF3 viral helicases. *Curr Opin Struct Biol*. Feb; 15(1):77-85.

Hidenori T., Osamu F., Atsushi F., Yasuyoshi M., Ryo M., Nobuyoshi A., Junichi T., Satoshi T., Yuji S., Naohiro N. (2010) Real-time monitoring of RNA helicase activity using fluorescence resonance energy transfer in vitro. *Biochem. and Biophys. Res. Comm.* 393, 131–136

Hoki, Y., Araki, R., Fujimori, A., Ohhata, T., Koseki, H., Fukumura, R., Nakamura, M., Takahashi, H., Noda, Y., Kito, S., et al. (2003) Growth retardation and skin abnormalities of the Recql4-deficient mouse. *Human molecular genetics* 12, 2293-2299.

Holliday R.(1964) A mechanism for gene conversion in fungi. *Genet. Res*. 5:282-304

Hu JS, Feng H, Zeng W, Lin GX, Xi XG. (2005) Solution structure of a multifunctional DNA- and protein-binding motif of human Werner syndrome protein. *Proc Natl Acad Sci U S A*. Dec 20;102(51):18379-84.

Hu, Y., X. Lu, et al. (2005) Recql5 and Blm RecQ DNA helicases have nonredundant roles in suppressing crossovers. *Mol Cell Biol.* 25(9): 3431-3442.

Hu, Y., Lu, X., Zhou, G., Barnes, E.L., and Luo, G. (2009) Recql5 plays an important role in DNA replication and cell survival after camptothecin treatment. *Molecular biology of the cell*. 20, 114-123.

Hu Y, Raynard S, Sehorn MG, Lu X, Bussen W, Zheng L, Stark JM, Barnes EL, Chi P, Janscak P, Jasin M, Vogel H, Sung P, Luo G (2007) RECQL5/Recql5 helicase regulates homologous recombination and suppresses tumor formation via disruption of Rad51 presynaptic filaments. *Genes Dev*. Dec 1;21(23):3073-84

Huang Y., Liu Z.-R. (2002) The ATPase, RNA Unwinding, and RNA Binding Activities of Recombinant p68 RNA Helicase. *J. Biol. Chem.*, 277, 12810-12815.

Huang S., Li B., Gray M.D., Oshima J., Mian I.S. and Campisi J. (1998) The premature ageing syndrome protein, WRN, is a 3'-->5' exonuclease. *Nat Genet.* 20: 114-6

Huber, M. D., M. L. Duquette, et al. (2006) A conserved G4 DNA binding domain in RecQ family helicases. *J Mol Biol*. 358(4): 1071-1080.

Ichikawa, K., Noda, T., and Furuichi, Y. (2002) Preparation of the gene targeted knockout mice for human premature aging diseases, Werner syndrome, and Rothmund-Thomson syndrome caused by the mutation of DNA helicases. *Nippon yakurigaku zasshi*. 119, 219-226.

Imamura O, Fujita K, Shimamoto A, Tanabe H, Takeda S, Furuichi Y, Matsumoto T. (2001) Bloom helicase is involved in DNA surveillance in early S phase in vertebrate cells. *Oncogene.* 20: 1143-1151

Islam, M. N., D. Fox, 3rd, et al. (2010) RecQL5 promotes genome stabilization through two parallel mechanisms--interacting with RNA polymerase II and acting as a helicase. *Mol Cell Biol.* 30(10): 2460-2472.

Johnson, F. B., Marciniak R. A., et al. (2001) The Saccharomyces cerevisiae WRN homolog Sgs1p participates in telomere maintenance in cells lacking telomerase. *EMBO J.* 20(4): 905-913.

Kamath-Loeb, A. S., L. A. Loeb, et al. (2001) Interactions between the Werner syndrome helicase and DNA polymerase delta specifically facilitate copying of tetraplex and hairpin structures of the d(CGG)n trinucleotide repeat sequence. *J Biol Chem*. 276(19): 16439-16446.

Kanagaraj, R., D. Huehn, et al. (2010) RECQ5 helicase associates with the C-terminal repeat domain of RNA polymerase II during productive elongation phase of transcription. *Nucleic Acids Res*. 38(22): 8131-8140.

Kanagaraj, R., N. Saydam, et al. (2006) Human RECQ5beta helicase promotes strand exchange on synthetic DNA structures resembling a stalled replication fork. *Nucleic Acids Res*. 34(18): 5217-5231.

Kantar, D.M., Kaplan, D.M. (2011) Sld2 binds to origin single-stranded DNA and stimulates DNA annealing. *Nucleic Acids Res*. 39(7): 2580-92

Kaplan, D. L., M. J. Davey, et al. (2003) Mcm4,6,7 uses a "pump in ring" mechanism to unwind DNA by steric exclusion and actively translocate along a duplex. *J Biol Chem*. 278(49): 49171-49182.

Karmakar, P., Snowden, C.M., Ramsden, D.A., Bohr, V.A. (2002) Ku heterodimer binds to both ends of the Werner protein and functional interaction occurs at the Werner Nterminus. *Nucleic Acids Res.* 30: 3583-3591.

Karow J.K., Constantinou A., Li J.L., West S.C., Hickson I.D. (2000) The Bloom's syndrome gene product promotes branch migration of holliday junctions, *Proc Natl Acad Sci U S A.* 97: 6504-8

Killoran MP, Keck JL. (2008) Structure and function of the regulatory C-terminal HRDC domain from Deinococcus radiodurans RecQ. *Nucleic Acids Res*. May;36(9):3139-49

Kim YM, Choi BS. (2010) Structure and function of the regulatory HRDC domain from human Bloom syndrome protein. *Nucleic Acids Res*. Nov;38(21):7764-77

Kim SY, Hakoshima T, Kitano K. (2013) Structure of the RecQ C-terminal domain of human Bloom syndrome protein. *Sci Rep*. Nov 21;3:3294.

Kitano K, Kim SY, Hakoshima T. (2010) Structural basis for DNA strand separation by the unconventional winged-helix domain of RecQ helicase WRN. *Structure*. Feb 10;18(2):177-87.

Kitano K, Yoshihara N, Hakoshima T. (2007) Crystal structure of the HRDC domain of human Werner syndrome protein, WRN. *J Biol Chem*. Jan 26;282(4):2717-28.

Kitao, S., A. Shimamoto, et al. (1999) Mutations in RECQL4 cause a subset of cases of Rothmund-Thomson syndrome. *Nat Genet.* 22(1): 82-84.

Kohzaki M, Chiourea M, Versini G, Adachi N, Takeda S, et al. (2012) The helicase domain and Cterminus of human RecQL4 facilitate replication elongation on DNA templates damaged by ionizing radiation. *Carcinogenesis*. 33:1203–10

Korolev S, Hsieh J, Gauss GH, Lohman TM,Waksman G. (1997) Major domain swiveling revealed by the crystal structures of complexes of E. coli Rep helicase bound to singlestranded DNA and ADP. *Cell*. 90:635–47

Kowalczykowski, S. C. (2000) Initiation of genetic recombination and recombination dependent replication. *Trends Biochem.* Sci 25(4): 156-165.

Kozin M. and Svergun D. I. (2001) Automated matching of high- and low-resolution structural models *J Appl Cryst*. 34: 33-41.

Krissinel, E. and Henrick, K. (2004) Secondary-structure matching (PDBeFold), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst*. D60: 2256-2268

Krude, T. (2010) Non- coding RNAs: new players in the field of eukaryotic DNA replication *Subcell Biochem.* 50: 105-118

Kumata, Y., Tada, S., Yamanada, Y., Tsuyama, T., Kobayashi, T., Dong, Y.P., Ikegami, K., Murofushi, H., Seki, M., and Enomoto, T. (2007) Possible involvement of RecQL4 in the repair of double-strand DNA breaks in Xenopus egg extracts. *Biochimica et biophysica acta.* 1773: 556-564.

Kunkel TA, Burgers PM. (2008) Dividing the workload at a eukaryotic replication fork. *Trends Cell Biol.* Nov;18(11): 521-7.

Kusumoto R, Dawut L, Marchetti C, Wan Lee J, Vindigni A, et al. (2008) Werner protein cooperates with the XRCC4-DNA ligase IV complex in end-processing. *Biochemistry*. 47: 7548–56

Kusumoto-Matsuo R, Opresko PL, Ramsden D, Tahara H, Bohr VA. (2010) Cooperation of DNA-PKcs and WRN helicase in the maintenance of telomeric D-loops. *Aging.* 2: 274–84

Kwon SH, Choi DH, Lee R, Bae SH. (2012) Saccharomyces cerevisiae Hrq1 requires a long 3'-tailed DNA substrate for helicase activity. *Biochem Biophys Res Commun*. Oct 26;427(3): 623-8.

Larsen NB, Hickson ID. (2013) RecQ Helicases: Conserved Guardians of Genomic Integrity. *Adv Exp Med Biol*. 767: 161-84.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR*. 8: 477-486.

Lauper JM, Krause A, Vaughan TL, Monnat RJ Jr. (2013) Spectrum and risk of neoplasia in Werner syndrome: a systematic review. *PLoS One*. 8(4): e59709.

Lebel, M., and Leder, P. (1998) A deletion within the murine Werner syndrome helicase induces sensitivity to inhibitors of topoisomerase and loss of cellular proliferative capacity. *Proc Nat Acad of Sci U S A.* 95: 13097-13102.

Lebel, M., Cardiff, R.D., Leder, P. (2001) Tumorigenic effect of nonfunctional p53 or p21 in mice mutant in the Werner syndrome helicase. *Cancer Res.* 61: 1816-1819.

Lee, J. W., J. Harrigan, et al. (2005) Pathways and functions of the Werner syndrome protein. *Mech Ageing Dev.* 126(1): 79-86.

Lee, J. W., R. Kusumoto, et al. (2005) Modulation of Werner syndrome protein function by a single mutation in the conserved RecQ domain. *J Biol Chem.* 280(47): 39627-39636.

Lee JY, Yang W (2006) UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. *Cell*. Dec 29;127(7): 1349-60

Lesburg, C.A., Huang, C., Christianson, D.W. and Fierke, C.A. (1997) Histidine --> carboxamide ligand substitutions in the zinc binding site of carbonic anhydrase II alter metal coordination geometry but retain catalytic activity. *Biochemistry*. 36: 15780-15791.

Levin, M. K., Y. H. Wang, et al. (2004) The functional interaction of the hepatitis C virus helicase molecules is responsible for unwinding processivity. *J Biol Chem.* 279(25): 26005-26012.

Li B, Comai L. (2000) Functional interaction between Ku and the Werner syndrome protein in DNA end processing. *J. Biol. Chem.* 275:28349–52

Li D, Frazier M, Evans DB, Hess KR, Crane CH, Jiao L, Abbruzzese JL. (2006) Single nucleotide polymorphisms of RecQ1, RAD54L, and ATM genes are associated with reduced survival of pancreatic cancer. *J Clin Oncol.* 24: 1720-1728

Li, D., R. Zhao, et al. (2003) Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen. *Nature* 423(6939): 512-518.

Lillard-Wetherell K, Machwe A, Langland GT, Combs KA, Behbehani GK, et al. (2004) Association and regulation of the BLM helicase by the telomere proteins TRF1 and TRF2. *Hum. Mol. Genet.* 13:1919–32

Linder P, Lasko P. (2006) Bent out of shape: RNA unwinding by the DEAD-box Helicase Vasa. *Cell.* 125: 219-21

Lindor NM, Furuichi Y, Kitao S, Shimamoto A, Arndt C, Jalal S. (2000) Rothmund Thomson syndrome due to RECQ4 helicase mutations: report and clinical and molecular comparisons with Bloom syndrome and Werner syndrome. *Am J Med Genet.* 90: 223-228

Littlefield, O. and H. C. Nelson (1999) A new use for the 'wing' of the 'winged' helix turn-helix motif in the HSF-DNA cocrystal. *Nat Struct Biol.* 6(5): 464-470.

Liu FJ, Barchowsky A, Opresko PL. (2010) The Werner syndrome protein suppresses telomeric instability caused by chromium (VI) induced DNA replication stress. *PLoS ONE* 5: e11152

Liu J.L., Rigolet P., Dou S.X., Wang P.Y., Xi X.G. (2004) The zinc finger motif of Escherichia coli RecQ is implicated in both DNA binding and protein folding, *J. Biol. Chem*. 279: 42794–42802.

Liu Z, Macias MJ, Bottomley MJ, Stier G, Linge JP, Nilges M, Bork P, Sattler M. (1999) The three-dimensional structure of the HRDC domain and implications for the Werner and Bloom syndrome proteins. *Structure*. Dec 15;7(12): 1557-66.

Lohman, T. M. (1992) Escherichia coli DNA helicases: mechanisms of DNA unwinding. *Mol Microbiol.* 6(1): 5-14.

Lohman TM, Bjornson KP. (1996) Mechanisms of helicase-catalyzed DNA unwinding. *Annu. Rev. Biochem.* 65:169–214

Lohman, T. M., E. J. Tomko, et al. (2008) Non-hexameric DNA helicases and translocases: mechanisms and regulation. *Nat Rev Mol Cell Biol*. 9(5): 391-401.

Lu H, Fang EF, Sykora P, Kulikowicz T, Zhang Y, Becker KG, Croteau DL, Bohr VA. (2014) Senescence induced by RECQL4 dysfunction contributes to Rothmund-Thomson syndrome features in mice. *Cell Death*. 5: e1226.

Lu L, Harutyunyan K, Jin W, Wu J, Yang T, Chen Y, Joeng KS, Bae Y, Tao J, Dawson BC, Jiang MM, Lee B, Wang LL. (2015) RECQL4 Regulates p53 Function in vivo During Skeletogenesis. *J Bone Miner Res.* Jan 1.

Lucic B, Zhang Y, King O, Mendoza-Maldonado R, Berti M, Niesen FH, Burgess-Brown NA, Pike AC, Cooper CD, Gileadi O, Vindigni A. (2011) A prominent β-hairpin structure in the winged-helix domain of RECQ1 is required for DNA unwinding and oligomer formation. *Nucleic Acids Res*. Mar; 39(5): 1703-17

Luo, G., Santoro, I.M., McDaniel, L.D., Nishijima, I., Mills, M., Youssoufian, H., Vogel, H., Schultz, R.A., and Bradley, A. (2000) Cancer predisposition caused by elevated mitotic recombination in Bloom mice. *Nature genetics.* 26: 424-429.

Luo D, Xu T, Watson RP, Scherer-Becker D, Sampath A, Jahnke W, Yeong SS, Wang CH, Lim SP, Strongin A, Vasudevan SG, Lescar J. (2008) Insights into RNA unwinding and ATP hydrolysis by the flavivirus NS3 protein. *EMBO J*. Dec 3;27(23): 3209-19.

Machwe A., Xiao L., Groden J., Matson S.W. and Orren D.K. (2005) RecQ family members combine strand pairing and unwinding activities to catalyze strand exchange. *J Biol Chem.* 280: 23397-407

Macris M.A., Krejci L., Bussen W., Shimamoto A and Sung P. (2006) Biochemical characterization of the RECQ4 protein, mutated in Rothmund-Thomson syndrome. *DNA Repair (Amst).* Feb 3;5(2): 172-80.

Maiorano D., Lutzmann M., Méchali M. (2006) MCM proteins and DNA replication. *Curr Opin Cell Biol.* Apr;18(2): 130-6.

Mann, M.B., Hodges, C.A., Barnes, E., Vogel, H., Hassold, T.J., and Luo, G. (2005) Defective sister-chromatid cohesion, aneuploidy and cancer predisposition in a mouse model of type II Rothmund-Thomson syndrome. *Human molecular genetics.* 14, 813-825.

Marino F, Vindigni A, Onesti S. (2013) Bioinformatic analysis of RecQ4 helicases reveals the presence of a RQC domain and a Zn knuckle. *Biophys Chem*. Jul-Aug;177-178: 34-9.

Masai H, Matsumoto S, You Z, Yoshizawa-Sugata N, Oda M (2010) Eukaryotic chromosome DNA replication: where, when, and how? *Annu Rev Biochem*. 79:89-130.

Matson, S. W., D. W. Bean, et al. (1994) DNA helicases: enzymes with essential roles in all aspects of DNA metabolism. *Bioessays* 16(1): 13-22.

Matsuno, K., Kumano, M., Kubota, Y., Hashimoto, Y., and Takisawa, H. (2006) The Nterminal noncatalytic region of Xenopus RecQ4 is required for chromatin binding of DNA polymerase alpha in the initiation of DNA replication. *Molecular and cellular biology.* 26: 4843-4852.

Mer, G., A. Bochkarev, et al. (2000) Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA. *Cell*. 103(3): 449-456.

Mohaghegh P., Karow J.K., Brosh R.M. Jr., Bohr V.A. and Hickson I.D. (2001) The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases, *Nucleic Acids Res* 29: 2843-9

Morozov, V., A. R. Mushegian, et al. (1997) A putative nucleic acid-binding domain in Bloom's and Werner's syndrome helicases. *Trends Biochem Sci*. 22(11): 417-418.

Moser, M. J., A. S. Kamath-Loeb, et al. (2000) WRN helicase expression in Werner syndrome cell lines. *Nucleic Acids Res.* 28(2): 648-654.

Moyer SE, Lewis PW, Botchan MR. (2006) Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc Natl Acad Sci U S A*. Jul 5;103(27): 10236-41

Muftuoglu M, Kusumoto R, Speina E, Beck G, Cheng WH, Bohr VA. (2008a) Acetylation regulates WRN catalytic activities and affects base excision DNA repair. *PLoS One*. Apr 9;3(4): e1918.

Murfuni I, De Santis A, Federico M, Bignami M, Pichierri P, Franchitto A. (2012) Perturbed replication induced genome wide or at common fragile sites is differently managed in the absence of WRN. *Carcinogenesis*. 33: 1655–1663.

Nakayama H, Nakayama K, Nakayama R, Irino N, Nakayama Y. and Hanawalt PC. (1984) Isolation and genetic characterization of a thymineless death-resistant mutant of *Escherichia coli* K12: identification of a new mutation (recQ1) that blocks the RecF recombination pathway. *Mol Gen Genet.* 195: 474-80

Nam, Y., Chen, C., Gregory, R.I., Chou, J.J., Sliz, P. (2011) Molecular basis for interaction of let-7 microRNAs with Lin28. *Cell*. 147: 1080-1091.
Nguyen GH, Dexheimer TS, Rosenthal AS, Chu WK, Singh DK, et al. (2013) A small molecule inhibitor of the BLM helicase modulates chromosome stability in human cells. *Chem. Biol.* 20: 55–62

Nimonkar AV, Genschel J, Kinoshita E, Polaczek P, Campbell JL, et al. (2011) BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev.* 25: 350–62

Nimonkar AV, Ozsoy AZ, Genschel J, Modrich P, Kowalczykowski SC. (2008) Human exonuclease 1 and BLM helicase interact to resect DNA and initiate DNA repair. *Proc. Natl. Acad. Sci. USA.* 105: 16906–11

O'Sullivan, R. J. and J. Karlseder (2010) Telomeres: protecting chromosomes against genome instability. *Nat Rev Mol Cell Biol.* 11(3): 171-181.

Ohlenschläger O, Kuhnert A, Schneider A, Haumann S, Bellstedt P, Keller H, Saluz HP, Hortschansky P, Hänel F, Grosse F, Görlach M, Pospiech H. (2012) The N-terminus of the human RecQL4 helicase is a homeodomain-like DNA interaction motif. *Nucleic Acids Res*. Jun 22.

Onoda F, Seki M, Miyajima A, Enomoto T. (2000) Elevation of sister chromatid exchange in Saccharomyces cerevisiae sgs1 disruptants and the relevance of the disruptants as a system to evaluate mutations in Bloom's syndrome gene. *Mutat Res*. 459(3): 203-9.

Opresko PL, Mason PA, Podell ER, Lei M, Hickson ID, Cech TR, Bohr VA. (2005) POT1 stimulates RecQ helicases WRN and BLM to unwind telomeric DNA substrates. *J Biol Chem*. Sep 16;280(37): 32069-80

Opresko, P. L., M. Otterlei, et al. (2004) The Werner syndrome helicase and exonuclease cooperate to resolve telomeric D loops in a manner regulated by TRF1 and TRF2. *Mol Cell*. 14(6): 763-774.

Opresko, P. L., C. von Kobbe, et al. (2002) Telomere-binding protein TRF2 binds to and stimulates the Werner and Bloom syndrome helicases. *J Biol Chem.* 277(43): 41110-41119.

Orren, D. K., S. Theodore, et al. (2002) The Werner syndrome helicase/exonuclease (WRN) disrupts and degrades D-loops in vitro. *Biochemistry*. 41(46): 13483-13488.

Otterlei M, Bruheim P, Ahn B, Bussen W, Karmakar P, Baynton K, Bohr VA. (2006) Werner syndrome protein participates in a complex with RAD51, RAD54, RAD54B and ATR in response to ICL-induced replication arrest. *J Cell Sci*. Dec 15;119(Pt 24): 5137-46.

Ouyang KJ, Woo LL, Zhu J, Huo D, Matunis MJ, Ellis NA. (2009) SUMO modification regulates BLM and RAD51 interaction at damaged replication forks. *PLoS Biol*. 7(12): e1000252.

Ouyang KJ, Yagle MK, Matunis MJ, Ellis NA. (2013) BLM SUMOylation regulates ssDNA accumulation at stalled replication forks. *Front Genet*. 4: 167.

Park, S.J., Lee, Y.J., Beck, B.D., Lee, S.H. (2006) A positive involvement of RecQL4 in UVinduced S-phase arrest. *DNA and cell biology*. 25: 696-703.

Parvathaneni S, Stortchevoi A, Sommers JA, Brosh RM Jr, Sharma S. (2013) Human RECQ1 interacts with Ku70/80 and modulates DNA end-joining of double-strand breaks. *PLoS ONE*. 8: e62481

Petkovic, M., Dietschy, T., Freire, R., Jiao, R., Stagljar, I. (2005) The human Rothmund-Thomson syndrome gene product, RECQL4, localizes to distinct nuclear foci that coincide with proteins involved in the maintenance of genome stability. *Journal of cell science.* 118: 4261-4269

Petoukhov M.V., Konarev P.V., Kikhney A.G., Svergun D.I. (2007) ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis. *J Appl Cryst.* 40: s223-s228.

Petoukhov M.V. et al. (2012) New developments in the ATSAS program package for small-angle scattering data analysis J. Appl. Cryst. 45: 342-350

Phatnani, H. P. and A. L. Greenleaf (2006) Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev.* 20(21): 2922-2936.

Pichierri P, Franchitto A, Mosesso P, Palitti F. (2001) Werner's syndrome protein is required for correct recovery after replication arrest and DNA damage induced in S-phase of cell cycle. *Mol Biol Cell.* 12: 2412-2421

Pichierri P, Nicolai S, Cignolo L, Bignami M, Franchitto A. (2012) The RAD9-RAD1-HUS1 (9.1.1) complex interacts with WRN and is crucial to regulate its response to replication fork stalling. *Oncogene*. 31: 2809–2823.

Pike AC, Shrestha B, Popuri V, Burgess-Brown N, Muzzolini L, Costantini S, Vindigni A, Gileadi O. (2009) Structure of the human RECQ1 helicase reveals a putative strand-separation pin. *Proc Natl Acad Sci U S A.* 106(4): 1039-44.

Plank, J. L., J. Wu, et al. (2006) Topoisomerase IIIalpha and Bloom's helicase can resolve a mobile double Holliday junction substrate through convergent branch migration. *Proc Natl Acad Sci U S A.* 103(30): 11118-11123.

Popp O, Veith S, Fahrer J, Bohr VA, Burkle A, Mangerich A. (2013) Site-specific noncovalent interaction of the biopolymer poly(ADP-ribose) with the Werner syndrome protein regulates protein functions. *ACS Chem. Biol.* 8:179–88

Popuri V., Bachrati C.Z., Muzzolini L., Mosedale G., Costantini S., Giacomini E., Hickson I.D., Vindigni A. (2008) The Human RecQ helicases, BLM and RECQ1 display distinct DNA substrate specificities, *J Biol Chem.* 283: 17766-76

Popuri V, Huang J, Ramamoorthy M, Tadokoro T, Croteau DL, Bohr VA. (2013) RECQL5 plays co-operative and complementary roles with WRN syndrome helicase. *Nucleic Acids Res.* 41: 881–99

Ralf, C., Hickson, I.D., Wu, L. (2006) The Bloom's syndrome helicase can promote the regression of a model replication fork. *J biol chem.* 281: 22839-22846.

Rao, V.A., Conti, C., Guirouilh-Barbat, J., Nakamura, A., Miao, Z.H., Davies, S.L., Sacca, B., Hickson, I.D., Bensimon, A., and Pommier, Y. (2007) Endogenous gamma-H2AXATM-Chk2 checkpoint activation in Bloom's syndrome helicase deficient cells is related to DNA replication arrested forks. *Mol Cancer Res.* 5: 713-724.

Ren, H., S. X. Dou, et al. (2008) The zinc-binding motif of human RECQ5beta suppresses the intrinsic strand-annealing activity of its DExH helicase domain and is essential for the helicase activity of the enzyme. *Biochem J.* 412(3): 425-433.

Rieping, W., Habeck, M., Bardiaux, B., Bernard, A., Malliavin, T.E. and Nilges, M. (2007) ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics.* 23: 381-382.

Rooney S, Chaudhuri J, Alt FW. (2004) The role of the non-homologous end-joining pathway in lymphocyte development. *Immunol. Rev.* 200: 115–31

Rossi ML, Ghosh AK, Kulikowicz T, Croteau DL, Bohr VA (2010) Conserved helicase domain of human RecQ4 is required for strand annealing-independent DNA unwinding. *DNA Repair (Amst).* Jul 1;9(7): 796-804

Rothstein, R., B. Michel, et al. (2000) Replication fork pausing and recombination or "gimme a break". *Genes Dev*. 14(1): 1-10.

Saikrishnan K, Powell B, Cook NJ, Webb MR, Wigley DB. (2009) Mechanistic basis of 5'-3' translocation in SF1B helicases, *Cell.* 137: 849-59.

Sallmyr A, Tomkinson AE, Rassool FV. (2008) Up-regulation of WRN and DNA ligase IIIαin chronic myeloid leukemia: consequences for the repair of DNA double-strand breaks. *Blood* 112: 1413–23

Sangrithi MN, Bernal JA, Madine M, Philpott A, Lee J, Dunphy WG, Venkitaraman AR. (2005) Initiation of DNA replication requires the RECQL4 protein mutated in Rothmund-Thomson syndrome. *Cell.* 121: 887-898

Sato A, Mishima M, Nagai A, Kim SY, Ito Y, Hakoshima T, Jee JG, Kitano K. (2010) Solution structure of the HRDC domain of human Bloom syndrome protein BLM. *J Biochem*. Oct;148(4): 517-25.

Scheffzek K, Ahmadian MR, Kabsch W, Wiesmuller L, Lautwein A, et al. (1997) The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science.* 277: 333–38

Schuler, W., Dong, C.Z., Wecker, K. and Roques, B.P. (1999) NMR Structure of the complex between the zinc finger protein NCp10 of Moloney Murine Leukemia virus and the single-stranded pentanucleotide d(ACGCC): comparison with HIV-NCp7 complexes. *Biochemistry*. 38: 12984-12994.

Schurman SH, Hedayati M, Wang Z, Singh DK, Speina E, Zhang Y, Becker K, Macris M, Sung P, Wilson DM 3rd, Croteau DL, Bohr VA. (2009) Direct and indirect roles of RECQL4 in modulating base excision repair capacity. *Hum Mol Genet*. Sep 15;18(18): 3470-83

Schwendener S, Raynard S, Paliwal S, Cheng A, Kanagaraj R, et al. (2010) Physical interaction of RECQ5 helicase with RAD51 facilitates its anti-recombinase activity. *J. Biol. Chem.* 285: 15739–45

Sengoku T, Nureki O, Nakamura A, Kobayashi S, Yokoyama S. (2006) Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. *Cell*. Apr 21;125(2): 287-300

Shamanna RA, Singh DK, Lu H, Mirey G, Keijzers G, Salles B, Croteau DL, Bohr VA. (2014) RECQ helicase RECQL4 participates in non-homologous end joining and interacts with the Ku complex. Carcinogenesis. Nov;35(11): 2415-24.

Sharma, S., Otterlei, M., Sommers, J.A., Driscoll, H.C., Dianov, G.L., Kao, H.I., Bambara, R.A., and Brosh, R.M., Jr. (2004a) WRN helicase and FEN-1 form a complex upon replication arrest and together process branchmigrating DNA structures associated with the replication fork. *Molecular biology of the cell.* 15: 734-750.

Sharma S, Phatak P, Stortchevoi A, Jasin M, Larocque JR. (2012) RECQ1 plays a distinct role in cellular response to oxidative DNA damage. *DNA Repair.* 11: 537–49

Sharma, S., Sommers, J.A., Wu, L., Bohr, V.A., Hickson, I.D., Brosh, R.M., Jr. (2004b) Stimulation of flap endonuclease-1 by the Bloom's syndrome protein. *J biol chem.* 279: 9847-9856.

Sharma S., Sommers J.A., Choudhary S., Faulkner J.K., Cui S., Andreoli L., Muzzolini L., Vindigni A. and Brosh R.M., Jr., (2005) Biochemical analysis of the DNA unwinding and strand annealing activities catalyzed by human RECQ1. *J Biol Chem* 280: 28072-84.

Sharma, S., and Brosh, R.M., Jr. (2007) Human RECQ1 is a DNA damage responsive protein required for genotoxic stress resistance and suppression of sister chromatid exchanges. *PLoS One.* 2: e1297.

Sharma, S., Stumpo, D.J., Balajee, A.S., Bock, C.B., Lansdorp, P.M., Brosh, R.M., Jr., and Blackshear, P.J. (2007) RECQL, a member of the RecQ family of DNA helicases, suppresses chromosomal instability. *Mol Cell Biol.* 27: 1784-1794.

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

Shen J.C., Gray M.D., Oshima J., Kamath-Loeb A.S., Fry M. and Loeb L.A. (1998) Werner syndrome protein. I. DNA helicase and dna exonuclease reside on the same polypeptide, *J Biol Chem* 273: 34139-44

Shereda, R. D., Reiter N. J., et al. (2009) Identification of the SSB binding site on E. coli RecQ reveals a conserved surface for binding SSB's C terminus. *J Mol Biol.* 386(3): 612-625.

Sidorova, J.M., Li, N., Folch, A., and Monnat, R.J., Jr. (2008) The RecQ helicase WRN is required for normal replication fork progression after DNA damage or replication fork arrest. *Cell cycle* Georgetown, Tex 7: 796-807.

Siddiqui-Jain, A., C. L. Grand, et al. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc Natl Acad Sci U S A*. 99(18): 11593-11598.

Siitonen HA, Kopra O, Kaariainen H, Haravuori H, Winter RM, Saamanen AM Peltonen L, Kestila M. (2003) Molecular defect of RAPADILINO syndrome expands the phenotype spectrum of RECQL diseases. *Hum Mol Genet.* 12: 2837-2844

Siitonen HA, Sotkasiira J, Biervliet M, Benmansour A, Capri Y, Cormier-Daire V, Crandall B, Hannula-Jouppi K, Hennekam R, Herzog D, Keymolen K, Lipsanen-Nyman M, Miny P, Plon SE, Riedl S, Sarkar A, Vargas FR, Verloes A, Wang LL, Kääriäinen H, Kestilä M. (2009) The mutation spectrum in RECQL4 diseases. *Eur J Hum Genet*. Feb;17(2): 151-8

Simonsson, T., P. Pecinka, et al. (1998) DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res*. 26(5): 1167-1172.

Simpson, R.J., Cram, E.D., Czolij, R., Matthews, J.M., Crossley, M. and Mackay, J.P. (2003) CCHX zinc finger derivatives retain the ability to bind Zn(II) and mediate protein-DNA interactions. *J Biol Chem*. 278: 28011-28018.

Singh, T. R., A. M. Ali, et al. (2008) BLAP18/RMI2, a novel OB-fold-containing protein, is an essential component of the Bloom helicase-double Holliday junction dissolvosome. *Genes Dev*. 22(20): 2856-2868.

Singh DK, Karmakar P, Aamann M, Schurman SH, May A, Croteau DL, Burks L, Plon SE, Bohr VA. (2010) The involvement of human RECQL4 in DNA double-strand break repair. *Aging Cell*. Jun;9(3): 358-71

Singh DK, Popuri V, Kulikowicz T, Shevelev I, Ghosh AK, et al. (2012) The human RecQ helicases BLM and RECQL4 cooperate to preserve genome stability. *Nucleic Acids Res.* 40: 6632–48

Singleton MR, Dillingham MS, Wigley DB. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem*. 76: 23-50.

Singleton, M. R., M. R. Sawaya, et al. (2000). Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell*. 101(6): 589-600.

Singleton, M. R. and D. B. Wigley (2002) Modularity and specialization in superfamily 1 and 2 helicases. *J Bacteriol.* 184(7): 1819-1826.

Skordalakes E, Berger JM. (2006) Structural insights into RNA dependent ring closure and ATPase acrivation by the Rho termination factor. *Cell*. Nov 3;127(3): 553-64.

Skordalakes E, Berger JM. (2003) Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell*. Jul 11;114(1): 135-46.

Soultanas, P. and Wigley D.B. (2001) Unwinding the 'Gordian knot' of helicase action. *Trends Biochem Sci*. 26(1): 47-54.

Sousa FG, Matuo R, Soares DG, Escargueil AE,Henriques JA, et al. (2012) PARPs and the DNA damage response. *Carcinogenesis*. 33: 1433–40

South, T.L. and Summers, M.F. (1993) Zinc- and sequence-dependent binding to nucleic acids by the N-terminal zinc finger of the HIV-1 nucleocapsid protein: NMR structure of the complex with the Psi-site analog, dACGCC. *Protein Science*. 2: 3-19.

Speina, E., L. Dawut, et al. (2010) Human RECQL5beta stimulates flap endonuclease 1. *Nucleic Acids Res*. 38(9): 2904-2916.

Storici, F., Bebenek, K., Kunkel, T.A., Gordenin, D.A. and Resnick M.A. (2007) RNA-templated DNA repair. *Nature*. 447: 338-341

Sun, H., R. J. Bennett, et al. (1999) The Saccharomyces cerevisiae Sgs1 helicase efficiently unwinds G-G paired DNAs. *Nucleic Acids Res*. 27(9): 1978-1984.

Sun H., Karow J.K., Hickson I.D., Maizels N. (1998) The Bloom's syndrome helicase unwinds G4 DNA. *J Biol Chem.* 273: 27587-92

Suzuki, N., M. Shiratori, et al. (1999) Werner syndrome helicase contains a 5'-->3' exonuclease activity that digests DNA and RNA strands in DNA/DNA and RNA/DNA duplexes dependent on unwinding. *Nucleic Acids Res.* 27(11): 2361-2368.

Svergun D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Cryst.* 25: 495-503.

Svergun D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J.* 2879-2886.

Svergun D.I., Barberato C., Koch M.H.J. (1995) CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Cryst.* 28: 768-773.

Svergun D.I. & Koch M.H.J. (2003) Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys*. 66: 1735–1782

Tackett, A. J., Y. Chen, et al. (2005) Multiple full-length NS3 molecules are required for optimal unwinding of oligonucleotide DNA in vitro. *J Biol Chem*. 280(11): 10797-10806.

Tadokoro T, Ramamoorthy M, Popuri V, May A, Tian J, et al. (2012) Human RECQL5 participates in the removal of endogenous DNA damage. *Mol. Biol. Cell.* 23: 4273–85

Tanaka T, Umemori T, Endo S, Muramatsu S, Kanemaki M, Kamimura Y, Obuse C, Araki H. (2011) Sld7, an Sld3-associated protein required for efficient chromosomal DNA replication in budding yeast. *EMBO J*. May 18; 30(10): 2019-30.

Temime-Smaali, N., L. Guittat, et al. (2008) Topoisomerase IIIalpha is required for normal proliferation and telomere stability in alternative lengthening of telomeres. *EMBO J.* 27(10): 1513-1524.

Thangavel S, Mendoza-Maldonado R, Tissino E, Sidorova JM, Yin J, Wang W, Monnat RJ Jr, Falaschi A, Vindigni A. (2010) Human RECQ1 and RECQ4 helicases play distinct roles in DNA replication initiation. *Mol Cell Biol*. Mar; 30(6): 1382-96.

Tikoo S, Madhavan V, Hussain M, et al. (2013) Ubiquitin-dependent recruitment of the Bloom Syndrome helicase upon replication stress is required to suppress homologous recombination. *EMBO J.* 32: 1778–1792.

Thomas C, Tulin AV. (2013) Poly-ADP-ribose polymerase: machinery for nuclear processes. *Mol. Asp. Med.* 34: 1124–37

Thomson, M. S. (1936) Poikiloderma Congenitale: Two Cases for Diagnosis. *Proc R Soc Med.* 29(5): 453-455.

Thomsen ND, Berger JM. (2009) Running in reverse: the structural basis for translocation polarity in hexameric helicases. *Cell.* Oct 30;139(3): 523-34.

Toth EA, Li Y, Sawaya MR, Cheng Y, Ellenberger T. (2003) The crystal structure of the bifunctional primase-helicase of bacteriophage T7. *Mol Cell*. Nov;12(5): 1113-23.

Trego KS, Chernikova SB, Davalos AR, Perry JJ, Finger LD, et al. (2011) The DNA repair endonuclease XPG interacts directly and functionally with the WRN helicase defective in Werner syndrome. *Cell Cycle.* 10: 1998–2007

Umezu, K., K. Nakayama, et al. (1990) Escherichia coli RecQ protein is a DNA helicase. *Proc Natl Acad Sci U S A.* 87(14): 5363-5367.

Unger T, Jacobovitch Y, Dantes A, Bernheim R, Peleg Y. (2010) Applications of the Restriction Free (RF) cloning procedure for molecular manipulations and protein expression. *J Struct Biol*. Oct;172(1): 34-44

van den Ent F, Löwe J. (2006) RF cloning: a restriction-free method for inserting target genes into plasmids. *J Biochem Biophys Methods*. Apr 30;67(1): 67-74

Van Maldergem L, Siitonen HA, Jalkh N, Chouery E, De Roy M, Delague V Muenke M, Jabs EW, Cai J, Wang LL Plon SE, Fourneau C, Kestilä M, Gillerot Y, Mégarbané A, Verloes A. (2006) Revisiting the craniosynostosis radial ray hypoplasia association: Baller-Gerold syndrome caused by mutations in the RECQL4 gene. *J Med Genet.* 43: 148-152

Vaughn, J. P., S. D. Creacy, et al. (2005) The DEXH protein product of the DHX36 gene is the major source of tetramolecular quadruplex G4-DNA resolving activity in HeLa cell lysates. *J Biol Chem*. 280(46): 38117-38120.

Velankar SS, Soultanas P, Dillingham MS, Subramanya HS, Wigley DB. (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism *Cell*. Apr 2;97(1): 75-84

Vennos EM, Collins M, James WD. (1992) Rothmund-Thomson syndrome: review of the world literature. *J Am Acad Dermatol.* 27: 750-762

Vindigni A, Hickson ID. (2009) RecQ helicases: multiple structures for multiple functions? *HFSP J*. Jun;3(3): 153-64

Vindigni A, Marino F, Gileadi O. (2010) Probing the structural basis of RecQ helicase function. *Biophys Chem.* 149(3): 67-77

Volkov V.V. and Svergun D. I. (2003) Uniqueness of ab-initio shape determination in small-angle scattering. *J. Appl. Cryst.* 36: 860-864.

von Hippel, P. H. (2004) Helicases become mechanistically simpler and functionally more complex. *Nat Struct Mol Biol*. 11(6): 494-496.

von Kobbe C, Harrigan JA, May A, Opresko PL, Dawut L, Cheng WH, Bohr VA. (2003) Central role for the Werner syndrome protein/poly(ADP-ribose) polymerase 1 complex in the poly(ADP-ribosyl)ation pathway after DNA damage. *Mol Cell Biol*. Dec;23(23): 8601-13.

von Kobbe, C., J. A. Harrigan, et al. (2004) Poly(ADP-ribose) polymerase 1 regulates both the exonuclease and helicase activities of the Werner syndrome protein. *Nucleic Acids Res.* 32(13): 4003-4014.

von Kobbe C, Karmakar P, Dawut L, Opresko P, Zeng X, et al. (2002) Colocalization, physical, and functional interaction between Werner and Bloom syndrome proteins. *J. Biol. Chem.* 277: 22035–44

von Kobbe, C., N. H. Thoma, et al. (2003) Werner syndrome protein contains three structure-specific DNA binding domains. *J Biol Chem*. 278(52): 52997-53006.

Vranken W.F., Boucher W., Stevens T.J., Fogh R.H., Pajon A., Llinas, M., Ulrich, E.L., Markley J.L., Ionides J., Laue E.D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins.* 59: 687-696.

Walker JE, Saraste M, Runswick MJ, Gay NJ. (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1: 945–51

Wang H, Rosidi B, Perrault R, Wang M, Zhang L, et al. (2005) DNA ligase III as a candidate component of backup pathways of nonhomologous end joining. *Cancer Res.* 65: 4020–30

Wang LL, Levy ML, Lewis RA, Chintagumpala MM, Lev D, Rogers M, Plon SE. (2001) Clinical manifestations in a cohort of 41 Rothmund-Thomson syndrome patients. *Am J Med Genet.* 102: 11-17

Wang L, Kaku H, Huang P, Xu K, Yang K, Zhang J, Li M, Xie L, Wang X, Sakai A, Watanabe M, Nasu Y, Shimizu K, Kumon H, Na Y. (2011) Single nucleotide polymorphism WRN Leu1074Phe is associated with prostate cancer susceptibility in Chinese subjects. *Acta Med Okayama*. Oct;65(5): 315-23.

Wang M, Wu W, Wu W, Rosidi B, Zhang L, et al. (2006) PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Res.* 34: 6170–82

Wang Z, Xu Y, Tang J, Ma H, Qin J, Lu C, Wang X, Hu Z, Wang X, Shen H. (2009b) A polymorphism in Werner syndrome gene is associated with breast cancer susceptibility in Chinese women. *Breast Cancer Res Treat.* Nov;118(1): 169-75

Watt, P. M., E. J. Louis, et al. (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(2): 253-260.

Wernimont, A. and Edwards, A. (2009) In Situ Proteolysis to Generate Crystals for Structure Determination: An Update. *PLoS ONE* 4:e5094.

Werner, O., (1985) On cataract in conjunction with scleroderma. Otto Werner, doctoral dissertation, 1904, Royal Ophthalmology Clinic, Royal Christian Albrecht University of Kiel. *Adv Exp Med Biol.* 190: 1-14.

Werner SR, Prahalad AK, Yang J, Hock JM. (2006) RECQL4-deficient cells are hypersensitive to oxidative stress/damage: Insights for osteosarcoma prevalence and heterogeneity in Rothmund-Thomson syndrome. *Biochem Biophys Res Commun.* 345: 403-409

West SC. (1996) The RuvABC proteins and Holliday junction processing in Escherichia coli. *JBacteriol*. Mar;178(5): 1237-41

Wilson DM III, Bohr VA. (2007) The mechanics of base excision repair, and its relationship to aging and disease. *DNA Repair.* 6: 544–59

Woo, L.L., Futami, K., Shimamoto, A., Furuichi, Y., Frank, K.M. (2006) The Rothmund-Thomson gene product RECQL4 localizes to the nucleolus in response to oxidative stress. *Experimental cell research.* 312: 3443-3457.

Wu, L., C. Z. Bachrati, et al. (2006) BLAP75/RMI1 promotes the BLM-dependent dissolution of homologous recombination intermediates. *Proc Natl Acad Sci U S A.* 103(11): 4068-4073.

Wu L, Chan KL, Ralf C, Bernstein DA, Garcia PL, Bohr VA, Vindigni A, Janscak P, Keck JL, Hickson ID. (2005) The HRDC domain of BLM is required for the dissolution of double Holliday junctions. *EMBO J*. Jul 20;24(14): 2679-87

Wu, L., Davies, S.L., Levitt, N.C., Hickson, I.D. (2001) Potential role for the BLM helicase in recombinational repair via a conserved interaction with RAD51. *J Biol Chem.* 276: 19375-19381.

Wu L, Hickson ID. (2003) The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature*. Dec 18;426(6968): 870-4

Wu, X. and N. Maizels (2001) Substrate-specific inhibition of RecQ helicase. *Nucleic Acids Res.* 29(8): 1765-1771.

Xie A, Kwok A, Scully R. (2009) Role of mammalian Mre11 in classical and alternative nonhomologous end joining. *Nat. Struct. Mol. Biol.* 16: 814–18

Xu, D., R. Guo, et al. (2008) RMI, a new OB-fold complex essential for Bloom syndrome protein to maintain genome stability. *Genes Dev.* 22(20): 2843-2855.

Xu, X. and Y. Liu (2009) Dual DNA unwinding activities of the Rothmund-Thomson syndrome protein, RECQ4. *EMBO J*. 28(5): 568-577.

Xu, X., Rochette, P.J., Feyissa, E.A., Su, T.V., Liu, Y. (2009) MCM10 mediates RECQ4 association with MCM2-7 helicase complex during DNA replication. *EMBO J.* 28: 3005-3014.

Yannone SM, Roy S, Chan DW, Murphy MB, Huang S, et al. (2001) Werner syndrome protein is regulated and phosphorylated by DNA-dependent protein kinase. *J. Biol. Chem.* 276: 38242–48

Yu, C. E., J. Oshima, et al. (1996) A YAC, P1, and cosmid contig and 17 new polymorphic markers for the Werner syndrome region at 8p12-p21. *Genomics*. 35(3): 431-440.

Zhang, X. D., S. X. Dou, et al. (2006) Escherichia coli RecQ is a rapid, efficient, and monomeric helicase. *J Biol Chem*. 281(18): 12655-12663.

Zhang Y, Jasin M. (2011) An essential role for CtIP in chromosomal translocation formation through an alternative end-joining pathway. *Nat. Struct. Mol. Biol.* 18: 80–84

Zhou, J., Bean, R.L., Vogt, V.M., Summers, M. (2007) Solution structure of the Rous sarcoma virus nucleocapsid protein: muPsi RNA packaging signal complex. *J Mol Biol*. 365: 453-467