



## **A novel DNA helicase with strand annealing activity from the crenarchaeon *Sulfolobus solfataricus***

Mariarita de Felice, Valentina Aria, Luca Esposito, Mariarosaria de Falco, Biagio Pucci, Mosé Rossi, Francesca M. Pisani

### **► To cite this version:**

Mariarita de Felice, Valentina Aria, Luca Esposito, Mariarosaria de Falco, Biagio Pucci, et al.. A novel DNA helicase with strand annealing activity from the crenarchaeon *Sulfolobus solfataricus*. *Biochemical Journal*, 2007, 408 (1), pp.87-95. <10.1042/BJ20070134>. <hal-00478745>

**HAL Id: hal-00478745**

**<https://hal.science/hal-00478745v1>**

Submitted on 30 Apr 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

## **A novel DNA helicase with strand annealing activity from the crenarchaeon *Sulfolobus solfataricus***

**Mariarita De Felice, Valentina Aria, Luca Esposito, Mariarosaria De Falco, Biagio Pucci, Mosè Rossi, and Francesca M. Pisani**

*Istituto di Biochimica delle Proteine, Consiglio Nazionale delle Ricerche. Via P. Castellino, 111. 80131-Napoli. Italy.*

**Short title:** *Archaeal DNA helicase with DNA-pairing activity*

**Keywords:** *Genome stability; DNA recombination; DNA helicase; DNA pairing activity; Archaea; Sulfolobus solfataricus.*

**Abbreviations:** *RQC, RecQ Conserved; HRDC, Helicase and RNase D Conserved; WH, winged helix; WRN, Werner syndrome helicase; BLM, Bloom syndrome helicase; TBE, Tris-Borate-EDTA buffer; ORF, Open Reading Frame; TSJ, Three-Strand Junction.*

**Correspondence:**

Dr Mariarita De Felice. Istituto di Biochimica delle Proteine, Consiglio Nazionale delle Ricerche. Via P. Castellino, 111. 80131-Napoli. Italy.

Tel: +39-0816132281

Fax: +39-0816132277

E-mail: [m.defelice@ibp.cnr.it](mailto:m.defelice@ibp.cnr.it)

## SYNOPSIS

To protect their genetic material cells adopt different mechanisms linked to DNA replication, recombination and repair. Several proteins function at the interface of these DNA transactions. In this study, we report on the identification of a novel archaeal DNA helicase. BlastP searches of *Sulfolobus solfataricus* genome database allowed us to identify an open reading frame (SSO0112, 875 amino acid residues) having sequence similarity with the human RecQ5 $\beta$ . The corresponding protein, named by us Hel112, was produced in *Escherichia coli* in soluble form, purified to homogeneity and characterised. Gel filtration chromatography and glycerol gradient sedimentation analyses revealed that Hel112 forms monomers and dimers in solution. Biochemical characterisation of the two oligomeric species revealed that only the monomeric form has an ATP-dependent 3'-5' DNA helicase activity, whereas, unexpectedly, both the monomeric and dimeric form possess DNA strand annealing capability. The Hel112 monomeric form is able to unwind forked and 3'-tailed DNA structures with high efficiency, whereas it is almost inactive on blunt-ended duplexes and bubble-containing molecules. This analysis reveals that *S. solfataricus* Hel112 shares some enzymatic features with the RecQ-like DNA helicases and suggests potential cellular functions of this protein.

## INTRODUCTION

The ability to faithfully duplicate genetic information is essential to all cells. To protect their genetic material cells adopt different mechanisms linked to DNA replication, recombination and repair. This is the reason why many human disorders are caused by defects in proteins involved in these processes. The importance of genome stability is evidenced by the difficulties that cells face when these mechanisms break down. There is a symbiotic relationship between DNA replication and recombination: the replication forks stalling due to damage on the genome is rescued through recombination events that ensure completion of the DNA replication process. Many aspects of DNA metabolic pathways, including both DNA replication and recombination, often involve steps that require unwinding of double-stranded DNA to form transient single-stranded intermediates, by the action of specific DNA helicases. They possess a distinct polarity, either 3' to 5' or 5' to 3' with respect to the strand on which the protein is presumed to translocate, deriving the energy from nucleotide triphosphate hydrolysis [1]. Particular attention has been given to the DNA helicases of the RecQ super-family [2]. They are named after the 3'-5' DNA helicase RecQ of *Escherichia coli* and are evolutionarily conserved from bacteria to humans [3]. RecQ-like DNA helicases share a centrally located helicase domain of about 400 amino acids that includes seven signature motifs characteristic of a wide variety of nucleic acids helicases. Included among these conserved sequences are the Walker A and B motifs, which are required for binding and hydrolysis of ATP. Outside of the helicase domain there is much less sequence similarity among the RecQ super-family members. Nevertheless, there are two additional identifiable domains that are found in some, but not all, RecQ-like proteins. A conserved region of about 80 residues, apparently unique to the RecQ super-family, is located C-terminally to the helicase domain and is designed as RecQ Conserved (RQC, [4]). This domain has been suggested to be involved in protein interaction [5] and DNA-binding [6, 7]. X-ray crystallographic analysis of *E. coli* RecQ protein has shown that the RQC region consists of an  $\alpha$ -helical  $\text{Zn}^{2+}$ -binding module followed by a winged-helix (WH) domain [8]. This latter acts as a double-stranded DNA-binding motif in several proteins including the catabolite gene activator protein [9]. A second sequence feature found in certain RecQ-like DNA helicases is the Helicase and RNase D Conserved (HRDC) motifs that is located C-terminally to the RQC domain. The HRDC motif is poorly conserved at the primary sequence level but has a well-defined structural fold comprising five  $\alpha$ -helices and a number of basic residues that constitute a putative DNA-binding surface [10]. Whereas unicellular organisms express a single RecQ-like protein, more complex organisms express two or more. There are five human RecQ DNA helicases: 1) WRN, defective in Werner syndrome; 2) BLM, defective in Bloom syndrome, 3) RTS, defective in Rothmund-Thomson syndrome; 4) RecQ1; and 5) RecQ5 [11]. Only the biochemical properties of BLM [12], WRN [13], RecQ1 [14], a *Drosophila* variant of RecQ5 [15],

*Saccharomyces cerevisiae* Sgs1 [16] and *E. coli* RecQ [17] have been carefully characterised. They all unwind DNA with a 3' to 5' polarity, but their substrate specificity is not identical. Several mechanisms have been proposed by which RecQ-like DNA helicases can help to prevent the breakdown of replication. These enzymes are thought to promote resolution of D-loops and double Holliday-junctions which originate when replication forks collapse due to the presence of a block in lagging strand [18]. In addition, RecQ-like DNA helicases have been implicated in reverse branch migration resolving the four-way junctions (or "chicken-foot" structures) that originate when a DNA lesion in the leading strand produces regression of the replication fork [19]. Unwinding of these secondary structures results in the leading strand bypassing the damage and avoids their processing by the action of Holliday junction resolvases and subsequent homologous recombination. On the other hand, BLM has been very recently proposed to promote the formation of the "chicken-foot" structures *in vivo* on the basis of its ability to catalyze regression of a model replication fork *in vitro* in an ATP-dependent manner [20].

Archaea, the third domain of life, share many similarities with Eukarya in their genetic information processing pathways even if they have a bacterial-like cellular structure [21]. Therefore, they represent an attractive model system to understand the more complex mechanisms of the eukaryotic cell. Whereas the DNA replication machinery of both Archaea and Eukarya has been extensively investigated, still little is known about the biological function of the various DNA helicases that are involved in repair or restart at damaged replication forks and no clear RecQ homolog has been identified so far in Archaea.

Here we report the identification and biochemical characterization of a novel archaeal DNA helicase. This protein, referred to as Hel112, has been identified as a putative homolog of human RecQ5 $\beta$  by sequence similarity searches of the *Sulfolobus solfataricus* genome database.

Recombinant Hel112 shows two predominant stable oligomeric states: monomer and dimer. Only the monomeric form of Hel112 has a 3'-5' DNA helicase activity, preferring a forked DNA as the substrate. Both the monomer and the dimer possess strand annealing activity. The biochemical properties of purified Hel112 are described and discussed.

## EXPERIMENTAL

**Materials** - All chemicals were of reagent grade. Restriction and modification enzymes were from New England Biolabs. Radioactive nucleotides were purchased from GE Healthcare Life Sciences. Oligonucleotides were synthesised by Proligo (Paris, France).

**Production of Hel112 in *E. coli*** - The Hel112 gene was amplified from *S. solfataricus* P2 genomic DNA by PCR using the Vent polymerase (New England Biolabs) with the oligonucleotide Hel112 NheI (5'- GGTTGCTAGC**TTG**AGCAATACTTACTTCTACTCTGATGAAG -3') as the 5' primer (the engineered NheI site is underlined and the Hel112 initiation codon is in bold) and the oligonucleotide Hel112 BamHI (5'- GGTTGGATCCTTAAACGGAAGTATACTCGAGCTCGATATTC-3') as the 3' primer (the engineered BamHI site is underlined). The PCR product was cloned into NheI/BamHI-linearised *E. coli* expression vector pET28a (Novagen) and sequenced. *E. coli* Rosetta cells (Novagen) transformed with the plasmid pET28a-Hel112 were grown at 37 °C in 5 litres of LB medium containing 30 µg/ml chloramphenicol and 100 µg/ml kanamycin. When the culture reached an  $A_{600\text{ nm}}$  of 0.8, protein expression was induced by adding isopropyl-1-thio- $\beta$ -D-galactopyranoside at 0.2 mM. The bacterial culture was incubated at 37 °C for additional 2 hours. Then cells were harvested by centrifugation and the pellet (10 g) was stored at -20 °C until use. The cell pellet was thawed and suspended in 50 ml of buffer A (25 mM Tris-HCl, pH 8.0, 2.5 mM MgCl<sub>2</sub>, 500 mM NaCl, 15% glycerol) supplemented with a protease inhibitors cocktail (Sigma). Cells were broken by two consecutive passages through a French pressure cell apparatus (Aminco Co., Silver Spring, MD) at 1500 p.s.i. The resulting cell extract was treated with DNase I (at 0.25 mg/ml) and Rnase A (at 0.1 mg/ml) for 30 min on ice with shaking. Then, protamine sulfate (at 2 mg/ml) was added and incubation was continued for additional 30 min on ice with shaking. The sample was centrifuged for 30 min at 30,000 rpm (Beckman rotor 70.0 Ti) at 10 °C. The supernatant was subjected to heat-treatment at 80 °C for 20 min, then kept in ice for 10 min. The thermo-precipitated proteins were removed by centrifugation for 30 min at 30,000 rpm (Beckman rotor 70.0 Ti) at 10 °C. The supernatant was passed through a 0.22-µm filter (Millipore), mixed with Ni<sup>2+</sup>-nitrilotriacetic acid Superflow-agarose resin (Qiagen) pre-equilibrated in buffer A and incubated for 1 hour in ice with gentle shaking. The resin was washed with buffer A and elution was carried out with 100 ml of an imidazole gradient (25–500 mM) in buffer A. 4-ml fractions were collected and analysed by SDS-PAGE to detect the Hel112 polypeptide. Fractions containing the recombinant protein were pooled and concentrated using an Amicon system up to 6 mg/ml. The protein sample was dialysed overnight against buffer A containing NaCl at 100 mM. The sample was loaded onto an Heparin Sepharose column (column volume: 8 ml) at 0.5 ml/min using an AKTA system (GE Healthcare Life Sciences). The column was washed using the same buffer and bound protein was eluted using

a linear gradient of NaCl in buffer (from 0.1 to 1 M NaCl). Fractions containing the recombinant Hel112 were pooled and concentrated. The sample was then loaded onto a Superdex 200 16/60 column (GE Healthcare Life Sciences) equilibrated in buffer A containing NaCl at 200 mM. The column was developed at a flow rate of 0.3 ml/min. 1-ml fractions were collected. A 20- $\mu$ l sample of each fraction was analysed by SDS-PAGE. Fractions containing the Hel112 protein were stored at 4 °C for subsequent biochemical analyses. The column was calibrated by running a set of gel filtration markers that included: thyroglobulin (669 kDa), ferritin (440 kDa), aldolase (158 kDa) and bovine serum albumin (BSA, 69 kDa).

**Glycerol gradient sedimentation** – Glycerol gradient ultracentrifugation was performed by applying the protein sample (60  $\mu$ g of fraction # 74 and 30  $\mu$ g of fraction # 65 in 150  $\mu$ l) to a 8 ml preformed 20-40 % glycerol gradient in buffer containing 25 mM Tris-HCl pH 8.0, 2.5 mM MgCl<sub>2</sub>, 200 mM NaCl. A parallel gradient was used to fractionate a mixture of standard proteins which included: catalase (232 kDa, 11.3 S), aldolase (158 kDa, 7.3 S), bovine serum albumin (67 kDa, 4.6 S) and ovalbumin (43 kDa, 3.5 S). After centrifugation at 38,000 rpm for 18 h in a Beckman SW41 rotor at 10 °C, fractions (200  $\mu$ l) were collected from the bottom of the tube and analysed by SDS-PAGE followed by Coomassie Blue Staining of the gel.

**DNA substrates** – The sequence of the synthetic oligonucleotides used to prepare the DNA substrates is reported in Table 1. Oligonucleotides were labelled using T4 polynucleotide kinase and [ $\gamma$ -<sup>32</sup>P]ATP after the labelling reaction they were purified using a Micro Bio-Spin P-30 Tris Chromatography Columns (Bio-Rad Laboratories), according to the manufacturer's instructions. To prepare duplexes DNA molecules, mixtures were incubated for 5 min at 95 °C which contained the labelled and the complementary unlabelled oligonucleotides at a 1:2 molar ratio. Then, the samples were slowly cooled at room temperature.

**DNA band-shift assays** – For each substrate, 10  $\mu$ l mixtures were prepared which contained: 200 fmol of [<sup>32</sup>P]-labelled DNA in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 2 mM EDTA, 0.5 mM MgCl<sub>2</sub>, 0.7 mM 2-mercaptoethanol, and the indicated amounts of Hel112. Following incubation for 10-15 min at room temperature, complexes were separated by electrophoresis through 5% polyacrylamide/bis gels (37.5:1) in 0.5 x TBE. Gels were dried down and analyzed using a Storm PhosphorImager (Molecular Dynamics; Image Quant software).

**DNA helicase assay** - Reaction mixtures (10  $\mu$ l) contained 50 fmol of [<sup>32</sup>P]-labelled substrate (about 1 x 10<sup>3</sup> cpm/fmol) and the indicated amounts of Hel112 (gel filtration peak fractions corresponding to the monomeric or the dimeric form of the protein) in helicase assay buffer (25 mM Tris-HCl, pH 8.0, 2.5 mM 2-mercaptoethanol, 50 mM sodium acetate, 5 mM MgCl<sub>2</sub>) containing 5

mM ATP. Samples were incubated for 30 min at 60 °C in a heated-top PCR machine to prevent evaporation and reactions were stopped by addition of 2.5 µl of 5 x Stop Solution (0.5% SDS, 40 mM EDTA, 0.5 mg/ml proteinase K, 20 % glycerol, 0.1% bromophenol blue). Reaction products were separated on a 8% polyacrylamide gel in TBE containing 0.1% SDS at a constant voltage of 150 V. After the electrophoresis the gel was soaked in 20% trichloroacetic acid and analyzed by a phosphorimager. Reaction products were quantified and any free oligonucleotide in the absence of enzyme was subtracted.

**Strand annealing assay** - Reaction mixtures (10 µl) contained partially or fully complementary synthetic oligonucleotides (each at a concentration of 2.5 nM), one of which was labelled at the 5'-end with T4 polynucleotide kinase and [ $\gamma$ -<sup>32</sup>P]ATP, and the indicated amounts of Hel112 in helicase assay buffer. Reaction mixtures were assembled on ice. Then, the protein was added to each mixture and the samples were incubated for the indicated times at 75 °C in a heated-top PCR machine to prevent evaporation and reactions were stopped by addition of 2.5 µl of 5 x Stop Solution. Samples were subjected to gel electrophoresis as described for the DNA helicase assay. Reaction products were quantified and any annealed oligonucleotide in the absence of enzyme was subtracted.

For the three-strand junction, samples (10 µl) contained a mixture of oligonucleotides (2.5 nM C80, 5 nM G80 and 5 nM G80bub; the C80 was labelled at the 5'-end with T4 polynucleotide kinase and [ $\gamma$ -<sup>32</sup>P]ATP), increasing amounts of Hel112 gel filtration peak fraction corresponding to the monomer (5, 15, 30, 50, 100 and 200 ng) in helicase buffer. Samples were incubated for 10 min at 60 °C. Then ATP was added at 5 mM into each sample, incubation was continued for additional 15 min at 60 °C and stopped with 2.5 µl of 5 x Stop Solution. Samples were subjected to gel electrophoresis as described for the DNA helicase assay.

**ATPase assay** – The ATPase reaction mixture (10 µl) contained 100 µM [ $\gamma$ -<sup>32</sup>P]ATP (0.5-1 µCi) and the indicated amounts of Hel112 (gel filtration peak fractions corresponding to the monomeric or dimeric form of the protein). Incubation was performed for 30 min at 60°C in a heated-top PCR machine to prevent evaporation and stopped in ice. A 1-µl sample of each mixture was spotted onto a polyethyleneimine-cellulose thin layer plate (Merck) and run in 0.5 M LiCl and 1 M formic acid. The [ $\gamma$ -<sup>32</sup>P]ATP hydrolysed to [<sup>32</sup>P]orthophosphate was quantified using a phosphorimager. The rate of ATP hydrolysis was determined in the linear range of reaction time and protein concentration dependence. The amount of spontaneously hydrolysed ATP was determined using blank reactions without enzyme and subtracted from the reaction rate values calculated as above.



## RESULTS

**Identification of a novel *Sulfolobus solfataricus* DNA helicase** - The aim of this study was to identify a DNA helicase belonging to the RecQ super-family in the hyper-thermophilic archaeon *Sulfolobus solfataricus* and investigate its biochemical properties. Since the archaeal molecular machineries responsible for DNA replication/recombination/repair are evolutionarily related to the eukaryal counterparts [21], we carried out BlastP [22] searches of the *S. solfataricus* genome database using each of the five human RecQ-like DNA helicases as query sequence in order to identify putative RecQ homologs in this archaeal species. Only when human RecQ5 $\beta$  was used as the query sequence in these searches, ORFs putatively coding for DNA helicases (SSO 0112, 0017, 0965, and 0394) were identified. Among these sequences, SSO0112 (875 amino acid residues; named by us Hel112) gave the highest similarity score with human RecQ5 $\beta$ . The most conserved region between the two proteins is the N-terminal portion which contains the sequence similarity motifs typically found in the Super-Family 2 DNA helicases ([23]; sequence similarity 41 %). Differently from other RecQ proteins, Hel112 does not seem to contain the RQC and the HRDC) sequence motifs that are located C-terminally to the conserved DNA helicase domain.

Hel112 was over-produced in *E. coli* as a protein containing an hexa-Histidine tag at the N-terminus using the vector pET28a. The protein was found to be expressed at high level in soluble form and was purified using a procedure that included a thermal treatment of the cell extracts, chromatographic steps on a Ni<sup>+</sup>-chelate resin, Heparin Sepharose, and gel filtration on a Superdex 200 column, as described under *Experimental*. The fractions of the gel filtration chromatography, carried out in a buffer containing 200 mM NaCl, were analysed by SDS-PAGE (Fig. 1A); this analysis revealed that Hel112 possesses two different oligomeric states with an apparent molecular mass of about 65 and 140 kDa. Since Hel112 theoretical molecular weight is calculated to be 100 kDa, we suppose that the two oligomeric states correspond to the monomeric and dimeric form of the protein in solution. In order to determine whether the two forms of Hel112 were stable, samples of the peak fractions (fraction # 65 corresponding to the dimer and fraction # 74 corresponding to the monomer) were subjected to analytical gel filtration chromatography on a Superdex 200 column (GE Healthcare Life Sciences; unpublished observations). They both eluted as single peaks corresponding to molecular globular species of 65 and 140 kDa suggesting that both the monomeric or the dimeric form are stable in solution. When the gel filtration analyses were carried out in different salt conditions (0.5 M and 1 M NaCl), the stability of the dimeric form was affected and it was converted into the monomeric one (data not shown). In addition, a sample of purified dimer was subjected to a prolonged incubation with  $\beta$ -mercaptoethanol (at 5 mM) and, then, analysed by gel filtration. Treatment with this sulphhydryl-reducing agent did not affect the oligomeric state of Hel112 revealing that disulphide bonds are not required for the formation of the protein dimeric assembly.

In order to analyse the biochemical properties of Hel112, DNA unwinding assays were carried out using different kinds of DNA duplex molecules. The activity assays were performed at 60 °C (and not at 75 °C) to reduce the spontaneous melting of the probe. We first analyzed the pool recovered from the Ni<sup>2+</sup>-chelate column but no DNA helicase activity was detected even using 1.5 µg of the protein sample on different substrates (data not shown). Then, we decided to test the DNA unwinding activity of the peak fractions eluted from the gel filtration column using two partially complementary oligonucleotides forming forked DNA molecules as the substrates (see Table 1). As shown in Fig. 2A, an evident displacement of the probe was observed when Hel112 in the monomeric state (gel filtration fraction # 74, see Fig.1) was assayed. In contrast, the dimeric species (corresponding to the gel filtration # 65, see Fig.1) did not display the ability to melt the fork DNA molecules (Fig. 2B-C). Surprisingly, the fraction # 65 seemed to catalyse annealing instead of unwinding of DNA molecules (Fig. 2B-C). Furthermore, DNA pairing activity was detected when fraction # 74 was assayed. To better investigate this issue strand annealing assays were carried out on both the gel filtration peak fractions using the fork DNA as the substrate. To avoid melting of double-stranded DNA due to the DNA unwinding activity of the protein, ATP was not present in the reaction mixtures. This analysis revealed that Hel112 in either monomeric or dimeric form possesses DNA pairing activity (see Fig 2 D-F).

To perform a more accurate analysis of Hel112 oligomeric state, glycerol gradient sedimentation analyses were carried out. Samples of the gel filtration peak fractions were subjected to ultracentrifugation through a preformed glycerol gradient. As shown in Fig. 1B, the protein of the peak fraction # 74 behaves as a monomer (expected molecular size of 100 kDa) and the protein of the peak fraction 65 was detected in a position of the gradient (between catalase and aldolase protein markers) compatible with a dimeric structure.

***The monomeric form of Hel112 has ATP-dependent 3'-5' DNA helicase activity*** – The gel filtration peak fraction corresponding to Hel112 monomeric form (fraction # 74 in Fig. 1) was used to more deeply characterise the DNA helicase activity. To determine the directionality of Hel112 helicase DNA unwinding assays were carried out using a substrate with a single-stranded region at either 3'- or 5'-end (3'-mis and 5'-mis, respectively). As shown in Fig. 3A, Hel112 is able to utilise only the substrate bearing a 3'-tail suggesting that the protein is able to translocate onto the DNA in the 3' to 5' direction. In addition, in the absence of ATP Hel112 was found to be devoid of any DNA unwinding capability indicating that the 3'-5 DNA helicase activity is ATP-dependent (Fig. 3B). To analyse the substrate preference of Hel112, various kinds of DNA substrates were used in the unwinding assays, such as fork, bubble and blunt-ended molecules. As shown in Fig. 4A, fork and 3'-tail DNA are the Hel112 favourite substrates, while no activity was detected on bubble or blunt-ended DNA.

**Both monomeric and dimeric forms of Hel112 possess ATPase activity** – The ATPase activity of the two Hel112 oligomeric forms was analysed, as shown in Fig. 5. The amount of [ $^{32}$ P] orthophosphate released was determined by means of a phosphorimager system and plotted versus the amounts of Hel112 used as either a monomer or a dimer (Fig. 5B). As indicated by the plots, the two oligomeric forms of Hel112 have a comparable level of ATPase activity. The ATP-hydrolysis function is not stimulated in the presence of either single- or double-stranded DNA (data not shown).

**Hel112 has strand annealing activity** – The Hel112 strand pairing activity was tested on different DNA substrates. An initial kinetic analysis was carried out using partially complementary oligonucleotides forming the fork DNA and a fixed amount of monomeric Hel112 (100 ng). As reported in Fig. 6A-B, Hel112 promoted annealing of about 50% of the single-stranded molecules by 2 min. On the other hand, in the absence of protein spontaneous melting of the paired oligonucleotides was observed, since samples were incubated at 75 °C due to the thermophilic nature of Hel112. To verify that the observed strand pairing was truly due to a specific catalytic function of Hel112, we carried out parallel control assays using *S. solfataricus* MCM DNA helicase, produced and characterised in our laboratory [24]. In the control reactions performed with this hexameric DNA helicase no difference was observed with the sample not containing protein (data not shown). In addition, DNA pairing assays were performed with increasing amounts of either monomeric or dimeric Hel112 at 75 °C for 2 min. Plots of annealed probe versus the amount of protein used are shown in Fig. 6D-E. It can be noticed that the strand annealing activity of Hel 112 monomeric and dimeric form is comparable on all the substrates used in these enzymatic assays (Fig. 6C).

Then, we decided to determine whether Hel112 was able to combine DNA pairing and unwinding to catalyse strand-exchange reactions, as recently reported for RecQ-like DNA helicases [25]. To this end, we utilised a fork substrate containing a 54-bp duplex region and 26-nt single-stranded 5'- and 3'-tails, which was made from a 5'-labelled 80-mer oligonucleotide (named C80) annealed to an unlabeled partially complementary 80-nt oligomer (named G80F). Assays were carried out with the Hel112 monomeric form using this fork substrate in combination with a third unlabeled 80-nt oligonucleotide (named G80B). The G80B oligonucleotide is complementary to both ends of C80, apart from a central non-complementary stretch (21-nt), and its 5'-end can pair with the 3'-tail of C80 to form a three-strand junction structure, as previously described [25]. Displacement of the G80F oligonucleotide from the three-strand junction structure together with annealing of G80B 3'-end to C80 5'-end produces a blunt-ended duplex with a central 21-nt bubble. In these assays, equimolar amounts of the above described fork substrate and the partially complementary G80B oligonucleotide were incubated with increasing amounts Hel112 monomer (from 5 to 200 ng) for 10 min at 60 °C in the absence of ATP. Then, ATP was added to the reaction mixtures and the

samples were incubated for additional 15 min. As shown in Fig 6F a slower migrating product was detected corresponding to a three-strand junction structure. On the other hand, the formation of the bubble-containing duplex was not observed as the result of enzyme-catalysed strand-exchange, even when the G80B was present in molar excess with respect to the forked DNA substrate and the reaction mixtures were incubated up to 30 min after the addition of ATP (data not shown). Therefore, Hel112 can promote strand-annealing also on DNA substrates with a more complex structure, but it seems unable to catalyse strand-exchange reactions, as recently reported for RecQ-like DNA helicases [25].

***Hel112 monomer and dimer show different DNA binding affinity*** – To further biochemically characterise Hel112, gel mobility shift assays were performed using either the monomeric or the dimeric form of the protein and various kinds of DNA ligands. As shown in Fig. 7, Fork and 3'-mis DNA are bound by Hel 112 in either monomeric or dimeric state with higher affinity. On the other hand, a clear difference in the DNA-binding capability is observed between the two forms of the protein on both single- and double-stranded DNA: the monomer binds these molecules with a very low affinity, whereas the DNA binding capability of the Hel 112 dimeric form is negligible on both these ligands.

Hel112 in the monomeric state is able to bind single-stranded DNA with an affinity lower than the one observed for the fork and 3'-mis DNA. In contrast, Hel112 in the dimeric form binds single-stranded DNA with very low affinity (only about 10% of the input DNA is shifted using 200 ng of protein, Fig. 7B-C). The difference in the DNA-binding capability between the two forms of the protein is more evident on double-stranded DNA: the monomer is able to bind these DNA molecules with low affinity (only 40 % of the total DNA is shifted by 400 ng of protein), whereas the dimer does not display any DNA-binding capability on this ligand.

## DISCUSSION

The Hel112 protein described in this study has been identified by a sequence similarity search of the *Sulfolobus solfataricus* genome database using human RecQ5 $\beta$  as the query sequence. Our biochemical analysis revealed that Hel112 possesses catalytic properties shared by RecQ super-family members. First of all, Hel112 can exist in two stable oligomeric states (monomer and dimer), but only the monomeric form displays a DNA unwinding activity which operates in the 3' to 5' direction. Similarly, RecQ proteins are monomeric, although *E. coli* RecQ, human BLM and WRN were also reported to form oligomeric assemblies in certain experimental conditions [17, 26-27]. Hel112 is able to unwind forked and 3'-tailed DNA structures with high efficiency, whereas it is almost inactive on blunt-ended duplexes and bubble-containing molecules. These properties are shared by human BLM and WRN and *Drosophila* and human RecQ5 $\beta$ . In contrast, *E. coli* RecQ is able to unwind blunt-ended DNA as well as bubble-containing structures [3]. Furthermore, *S. solfataricus* Hel112 is endowed with an efficient strand annealing activity that promotes the formation of various duplex DNA molecules starting from fully or partially complementary synthetic oligonucleotides. This peculiar feature has been recently described for RecQ-like proteins of metazoans, including WRN, BLM, RecQ1 and RecQ5 $\beta$  [28-30].

Nonetheless, Hel112 lacks both the QRC and HRDC sequence motifs that are typically found in RecQ super-family members [3]. The sequence similarity between human RecQ5 $\beta$  is limited to the N-terminal half of the polypeptide chain which contains the sequence similarity boxes of the Super-Family 2 DNA helicases [23]. All that considered, Hel112 can not be strictly classified within the RecQ Super-Family, although it displays biochemical features and catalytic properties that are hallmark of the RecQ-like DNA helicases. Despite the proposed critical cellular role and their ubiquitous distribution among Bacteria and Eukarya, clear RecQ homologs have not yet been identified in Archaea. Nevertheless, proteins responsible for Holliday junction migration were characterized from *Pyrococcus furiosus* (Hjm, [31]) and *M. thermoautotrophicus* (ORF Mth810, [32]). Both these enzymes belong to the Mus308 family of DNA helicases/translocases that support genome stability in metazoans, but are absent in bacteria and yeasts. This group includes human and *Drosophila* Hel308, human Pol  $\theta$  and *Drosophila* Mus308. These two latter contain a C-terminal polymerase module in addition to the N-terminal helicase domain. Hjm and Mth810 proteins, lacking the polymerase module, belong to the Hel308 subset. Similar sequences have been found in various sequenced archaeal genomes, including *S. solfataricus* (ORF # SSO2462) and *S. tokodaii* (ORF # STO0590). However, all these DNA helicases do not belong to the RecQ Super-Family.

Interestingly, Hel112 homologs are present in all the other archaeal species whose genome has been completely sequenced, suggesting a critical function for these proteins in DNA transactions of

both crenarchaeal and euryarchaeal species. A deeper analysis of the physical and/or functional interaction of Hel112 with other DNA replication/repair factors will help to unravel the biological role of this novel DNA helicase with strand pairing activity.

## **ACKNOWLEDGMENTS**

This work was partially supported by grants from Agenzia Spaziale Italiana (Project MoMa N° 1/014/06/0) and from ATIBB-BioTekNet Centro Regionale di Competenza in Biotecnologie Industriali.

## REFERENCES

1. Delagouette, E. and von Hippel, P. H. (2002) Helicase mechanisms and the coupling of helicases within macromolecular machines. Part I: Structures and properties of isolated helicases. *Q. Rev. Biophys.* **35**, 431-478.
2. Hickson, I. D. (2003) RecQ helicases: caretakers of the genome. *Nat. Rev. Cancer* **3**, 169-178.
3. Bennett, R. J. and Keck, J. L. (2004) Structure and function of RecQ DNA helicases. *Crit. Rev. Biochem. Mol. Biol.* **39**, 79-97.
4. Morozov, V., Mushegian, A. R., Koonin, E. V. and Bork, P. (1997) A putative nucleic acid-binding domain in Bloom's and Werner's syndrome helicases. *Trends Biochem. Sci.* **22**, 417-418.
5. 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. *EMBO J.* **20**, 5791-5801.
6. Von Kobbe, C., Thoma, N. H., Czyzewski, B. K., Pavletich, N. P., Bohr, V. A. (2003) Werner syndrome protein contains three structure-specific DNA binding domains. *J. Biol. Chem.* **278**, 52997-53006.
7. Liu, J. L., Rigolet, P., Dou, S. X., Wang, P. Y. and 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.
8. Bernstein, D. A., Zittel, M. C. and Keck, J. L. (2003) High-resolution structure of the *E. coli* RecQ helicase catalytic core. *EMBO J.* **22**, 4910-4921.
9. Gajiwala, K. S., and Burley, S. K. (2000) Winged helix proteins. *Curr. Opin. Struct. Biol.* **10**, 110-116.
10. Liu, Z., Macias, M. J., Bottomley, M. J., Stier, G., Linge, J. P., Nilges, M., Bork, P. and Sattler, M. (1999) The three-dimensional structure of the HRDC domain and implications for the Werner and Bloom syndrome proteins. *Structure* **7**, 1557-1566.
11. Harrigan, J. A. and Bohr, V. A. (2003) Human diseases deficient in RecQ helicases. *Biochimie* **85**, 1185-1193.
12. 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**, 2843-2849.
13. Brosh, R. M. Jr, Opresko, P. L. and Bohr, V. A. (2006) Enzymatic mechanism of the WRN helicase/nuclease. *Methods Enzymol.* **409**, 52-85.

14. Cui, S., Klima, R., Ochem, A., Arosio, D., Falaschi, A. and Vindigni, A. (2003) Characterization of the DNA-unwinding activity of human RecQ1, a helicase specifically stimulated by human replication protein A. *J. Biol. Chem.* **278**, 1424-1432.
15. Ozsoy, A. Z., Ragonese, H. R. and Matson, S. W. (2003) Analysis of helicase activity and substrate specificity of *Drosophila* RECQ5. *Nucleic Acids Res.* **31**, 1554-1564.
16. Bennett, R. J., Sharp, J. A. and Wang, J. C. (1998) Purification and characterization of the Sgs1 DNA helicase activity of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **273**, 9644-9650.
17. Harmon, F. G. and Kowalczykowski, S. C. (2001) Biochemical characterization of the DNA helicase activity of the *Escherichia coli* RecQ helicase. *J. Biol. Chem.* **276**, 232-243.
18. Wu, L. and Hickson, I. D. (2003) The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* **426**, 870-874.
19. Karow, J. K., Costantinou, A., Li, J. L., West, S. C. and Hickson, I. D. (2000) The Bloom's syndrome gene product promotes branch migration of Holliday junctions. *Proc. Natl. Acad. Sci. USA* **97**, 6504-6508.
20. Ralf, C., Hickson, I. D. and Wu, L. (2006) The Bloom's syndrome helicase can promote the regression of a model replication fork. *J. Biol. Chem.* **281**, 22839-22846.
21. Kelman, Z. and White, M. F. (2005) Archaeal DNA replication and repair. *Curr. Opin. Microbiol.* **8**, 669-676.
22. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402.
23. Gorbalenya, A. E. and Koonin, E. V. (1993) Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.* **3**, 419-429.
24. Carpentieri, F., De Felice, M., De Falco, M., Rossi, M. and Pisani, F. M. (2002) Physical and functional interaction between the mini-chromosome maintenance-like DNA helicase and the single-stranded DNA binding protein from the crenarchaeon *Sulfolobus solfataricus*. *J. Biol. Chem.* **277**, 12118-27.
25. 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**, 23397-23407.
26. Karow, J. K., Newman, R. H., Freemont, P. S. and Hickson, I. D. (1999) Oligomeric ring structure of the Bloom's syndrome helicase. *Curr. Biol.* **9**, 597-600.
27. Xue, Y., Ratcliff, G. C., Wang, H., Davis-Searles, P. R., Gray, M. D., Erle, D. A. and Redinbo, M. R. (2002) A minimal exonuclease domain of WRN forms hexamer on DNA and possesses both 3'-5' exonuclease and 5'-protruding strand endonuclease activities. *Biochemistry* **41**, 2901-2912.



28. Garcia, P. L., Liu, Y., Jiricny, J., West, S. C. and Janscak, P. (2004) Human RECQ5 $\beta$ , a protein with DNA helicase and strand-annealing activities in a single polypeptide. *EMBO J.* **23**, 2882-2891.
29. 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**, 28072-28084.
30. 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**, 3932-3941.
31. Fujikane, R., Komori, K., Shinagawa, H. and Ishino, Y. (2005) Identification of a novel helicase activity unwinding branched DNAs from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Biol. Chem.* **280**, 12351-12358.
32. Guy, C. P. and Bolt, E. D. (2005) Archaeal Hel308 helicase targets replication forks *in vivo* and *in vitro* and unwinds lagging strands. *Nucleic Acids Res.* **33**, 3678-3690.

## LEGEND TO FIGURES

**Fig. 1. Oligomeric state of Hel112.** (A) The Hel112 pool from the Heparin Sepharose chromatography was run on a Superdex 200 HR 16/60 gel filtration column, as described in *Experimental*. A 20- $\mu$ l sample of each fraction (1.0 ml) was loaded on a SDS polyacrylamide gel which was stained with Coomassie Blue. Arrows indicate the position of two protein markers used to calibrate the column (bovine serum albumin, BSA; and aldolase, Ald). (B) Samples of the gel filtration fraction # 65 and # 74 were fractionated on a glycerol gradient as described in *Experimental*. Aliquots of each fraction were analysed by SDS-PAGE followed by Coomassie Blue staining. Arrows indicate the position of protein markers fractionated on a parallel gradient. (C) A plot of the S values versus the position in the gradient for each marker protein is reported. S values for the protein in the peak fraction # 65 and # 74 were estimated to be 8.5 and 5.7, respectively.

**Fig. 2. Catalytic activities of Hel112 dimeric and monomeric form.** Enzymatic assays were carried out using the Hel112 dimeric and monomeric form. Increasing amounts of each protein (1.4, 7, 14, 70, 140 and 280 ng) were used either in DNA helicase (A-B) or in DNA pairing (D-E) assays. The labelled single-stranded DNA was quantified using a phosphorimager system and plotted versus the amount of protein used. The labelled single-stranded DNA present in the blank reaction was considered as 0 or 100 % in the plots shown in C and F, respectively. Plots obtained for the DNA helicase (C) and DNA pairing (F) assays are shown. Symbols used are as follows: *diamonds*, monomeric Hel112; *circles*, dimeric Hel112. ssDNA stands for single-stranded DNA. The enzymatic assays were performed at least in triplicate and the results were averaged. The error bars of the graphs are the standard error of the mean.

**Fig. 3. DNA helicase directionality and ATP-hydrolysis dependency.** (A) Increasing amounts of monomeric Hel112 (100, 200, 400 and 800 ng) were used in DNA unwinding assays with 3'-mis or 5'-mis DNA molecules as the substrates. (B) To check the DNA helicase dependency on ATP-hydrolysis, increasing amounts of monomeric Hel112 (100, 200, 400 and 800 ng) were assayed in the presence or absence of ATP on the same 3'-mis DNA substrate. Dots on the oligonucleotides represents the radio-labelled phosphate group at the 5'-end.

**Fig. 4. DNA helicase activity of Hel112 on various DNA substrates.** (A) Increasing amounts of Hel112 in the monomeric form (25, 50, 100, 200, 400 and 800 ng) were used to analyse the DNA unwinding activity on various DNA molecules, as described under *Experimental*. (B) The displaced oligonucleotide was quantified by a phosphorimager system and plotted against the amount of protein used. Symbols used are: *squares*, 3'-mis; *diamonds*, fork; *circles*, bubble; *triangles*, double-

stranded DNA. The enzymatic assays were performed at least in triplicate and the results were averaged. The error bars of the graphs are the standard error of the mean.

**Fig. 5. ATPase activity of Hel112.** ATPase assays were performed, as described in *Experimental*. (A) Increasing amounts (5, 10, 25, 50, 100 and 250 ng) of Hel112 monomer (*triangles*) or dimer (*squares*) were used. The orthophosphate released during the ATP hydrolysis was plotted versus the amount of protein used. The enzymatic assays were performed at least in triplicate and the results were averaged. The error bars of the graphs are the standard error of the mean.

**Fig. 6. DNA annealing activity of Hel112.** (A) Annealing of partially complementary oligonucleotides to form bubble DNA molecules was carried out in reaction mixtures containing (or not) monomeric Hel112 (100 ng). (B) The percentage of annealed probe was plotted versus the reaction time in the absence (*circles*) or presence of monomeric Hel 112 (*squares*). The annealed probe present in the blank reaction was considered as 0 %. The enzymatic assays were performed at least in triplicate and the results were averaged. The error bars of the graphs are the standard error of the mean. (C) Different DNA molecules were used to analyse the substrate specificity of monomeric or dimeric Hel112 in the annealing reaction. (D-E) The annealed probe was plotted versus the amount (0, 10, 25, 50, 100 ng) of Hel112 in the monomeric (D) or dimeric (E) form. The annealed probe present in the blank reaction was considered as 0 %. Symbols used are as follows: *circles*, double-stranded DNA; *squares*, fork; *triangles*; 3'-mis; *diamonds*, bubble DNA. The enzymatic assays were performed at least in triplicate and the results were averaged. The error bars of the graphs are the standard error of the mean. (F) Increasing amounts of monomeric Hel112 (5, 10, 30, 50, 100, 200 ng) were tested in strand exchange assays as described in the text. *TSJ* stands for three-strand junction.

**Fig. 7. DNA binding activity of Hel112.** The ability of Hel112 to bind different  $^{32}\text{P}$ -labeled synthetic oligonucleotides was analysed by gel shift assays as described under *Experimental*. (A) Increasing amounts of Hel112 (50, 100, 200 and 400 ng) were used in the presence of the indicated DNA molecules. For each DNA ligand used a plot is reported of the shifted DNA versus the amount of monomeric (B) or dimeric (C) protein used. Symbols used are as follows: *open squares*, bubble; *filled squares*, fork; *filled triangles*, double-stranded DNA; *open triangles*, single-stranded DNA; *circles*, 3'-mis. The enzymatic assays were performed at least in triplicate and the results were averaged. The error bars of the graphs are the standard error of the mean.

**Table 1.** Synthetic oligonucleotides used to prepare the DNA substrates used in this study.

Name used	Oligonucleotide sequence
ds-DNA	Strand 1: 5'-GCTCGGTACCCGGGGATCCTCTAGA-3' Strand 2: 5'-TCTAGAGGATCCCCGGGTACCGAGC-3'
3'-mis	Strand 1: 5'-CTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGC-3' Strand 2: 5'-TCCTCTAGAGTCGACCTGCAGGCATGCAAG-3'
5'-mis	Strand 1: as strand 1, 3'-mis Strand 2: 5'-TCACACAGGAAACAGCTATGACCATGATTA-3'
Bubble	Strand 1: 5'-TCTACCTGGACGACCGGGTTTTTTTTTTTTTTTTTTTGGGCCAGCAGGTCCATCA-3' Strand 2: 5'-TGATGGACCTGCTGGCCCTTTTTTTTTTTTTTTTTTTTCCCGGTCGTCCAGGTAGA-3'
Fork	Strand 1: 5'-GCTCGGTACCCGGGGATCCTCTAGATTTTTTTTTTTTTTTTTTT-3' Strand 2: 5'-TTTTTTTTTTTTTTTTTTTCTAGAGGATCCCCGGGTACCGAGC-3'
TSJ	C80: 5'-GCTGATCAACCCTACATGTGTAGGTAACCCTAACCCTAACCCTAAGGACAACCCTAGTGAAGCTTGTAACCCTAGGAGCT-3' G80F: 5'-CTTCACAGTCAGAGTCACAGTCTGCCGGGTTGTCTTAGGGTTAGGGTTAGGGTTACCTACACATGTAGGGTTGATCAGC-3' G80B: 5'-AGCTCCTAGGGTTACAAGCTTCACTAGGGTTGTCCAGTCACAGTCAGAGTCACAGTCCTACACATGTAGGGTTGATCAGC-3'

TSJ stands for three-strand junction.

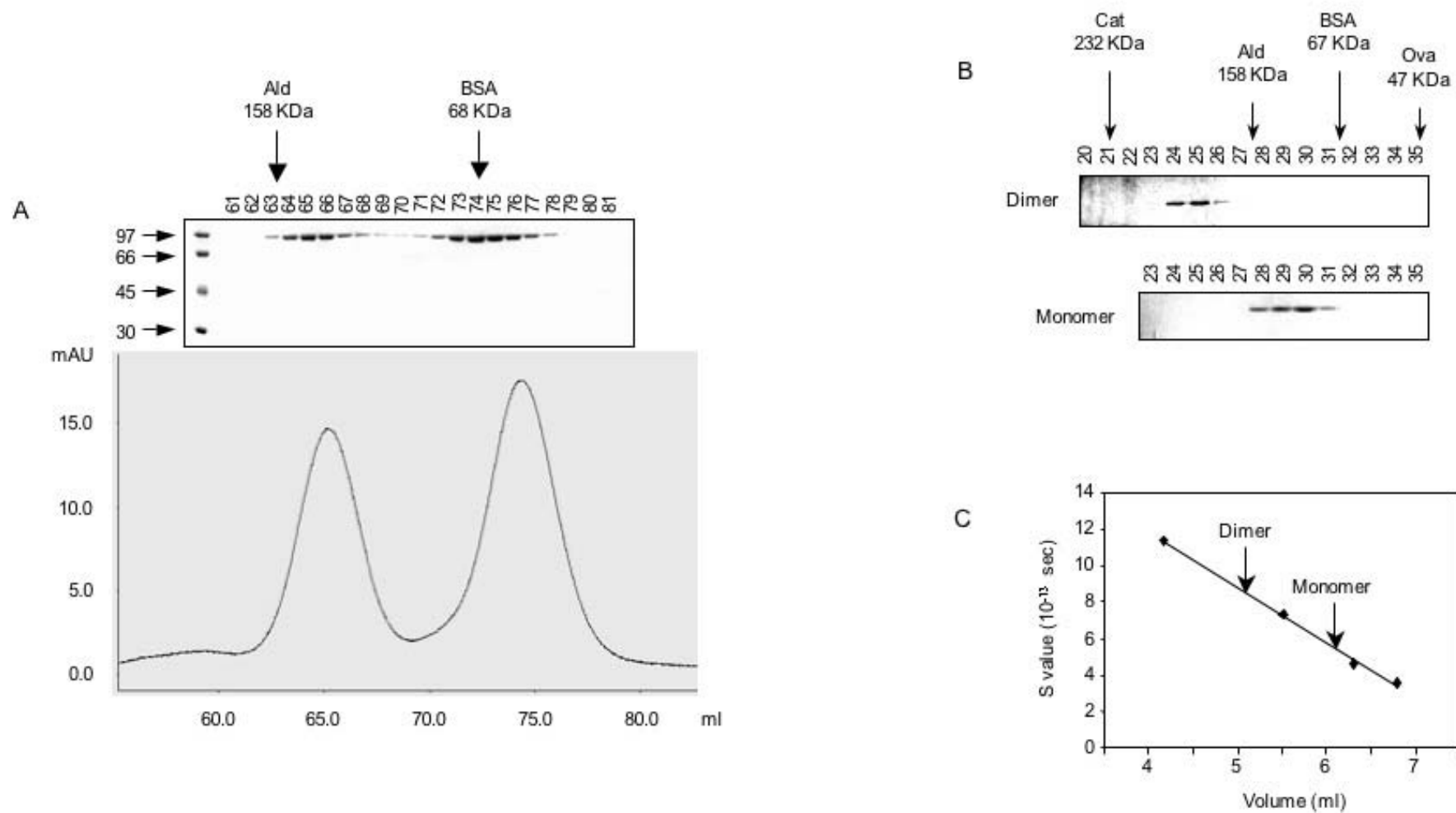


Fig.1

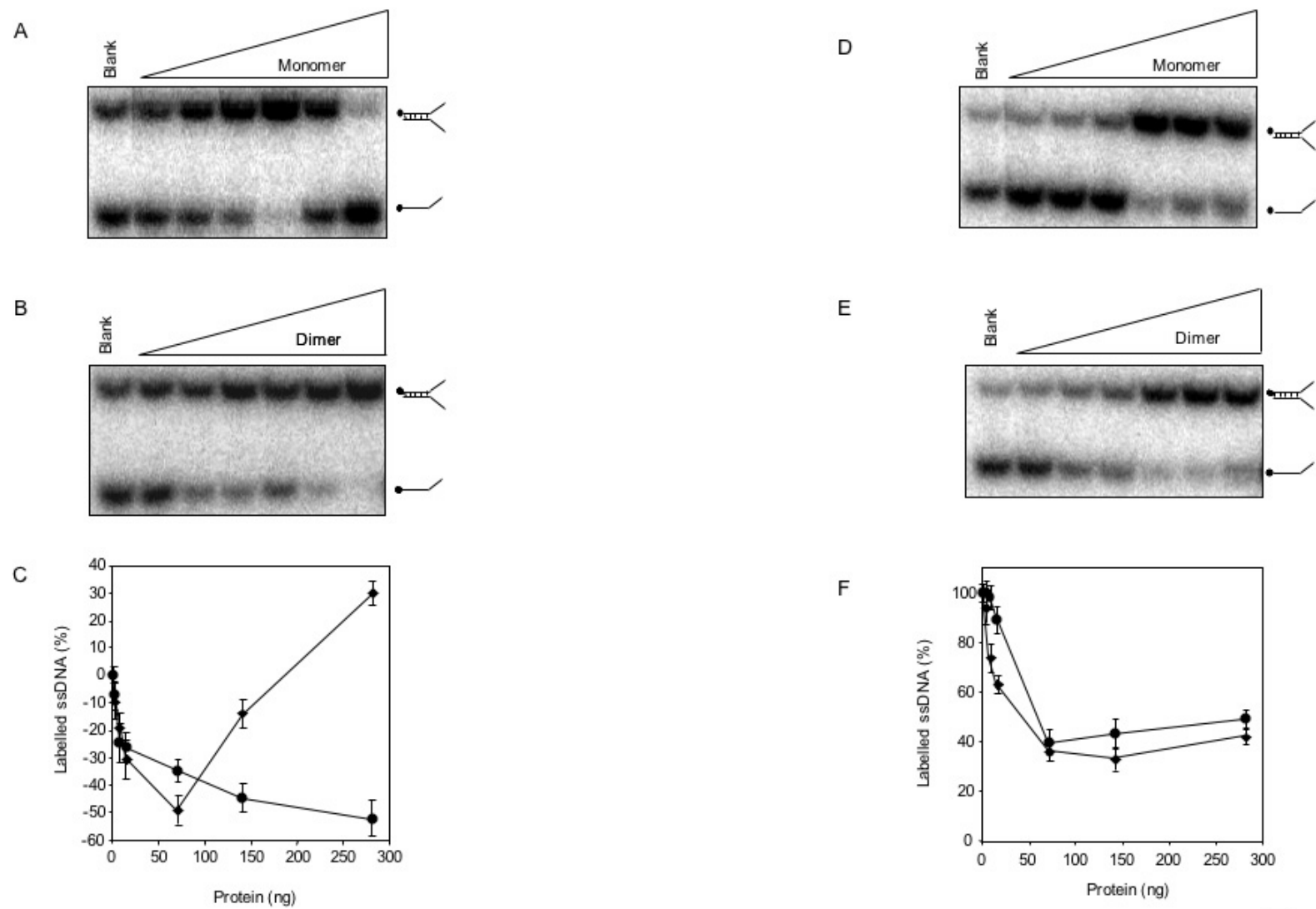


Fig. 2

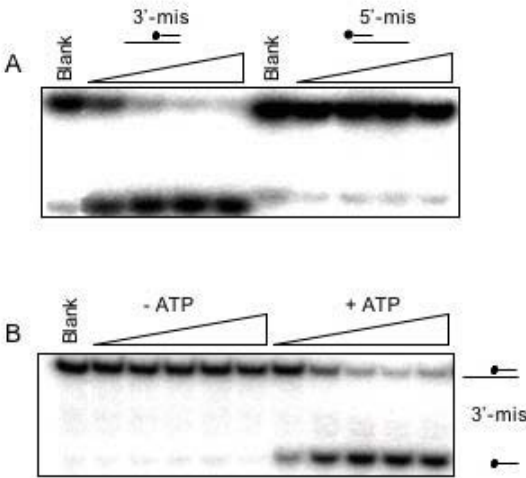


Fig. 3

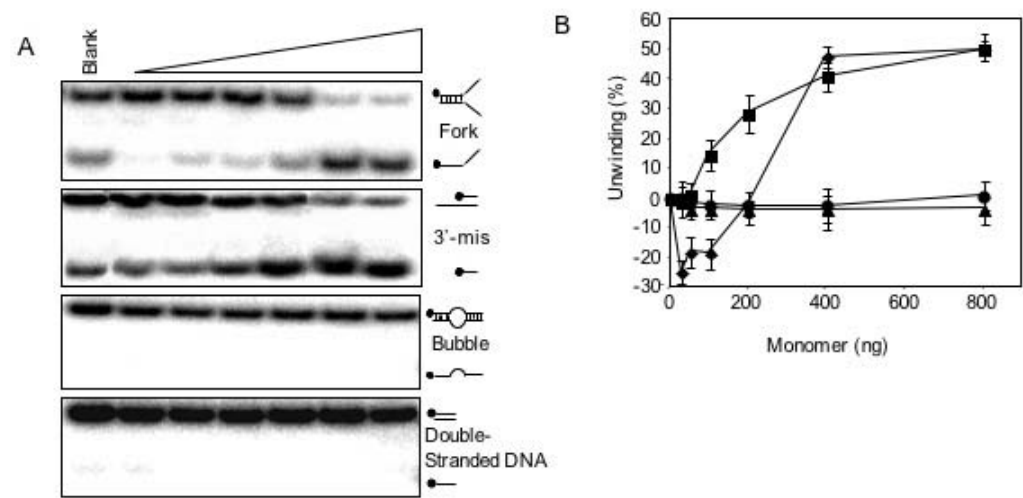


Fig. 4



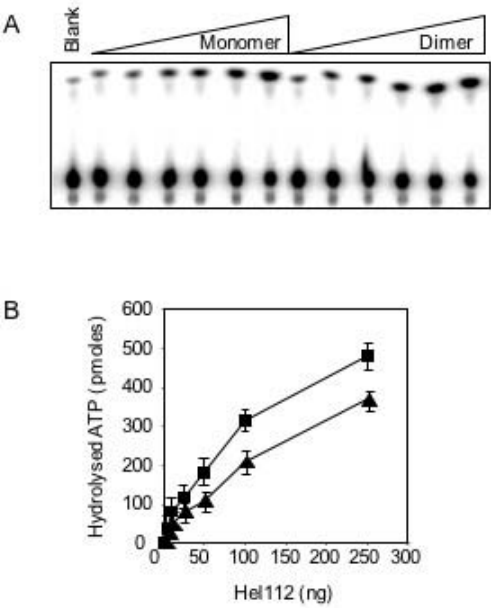


Fig. 5

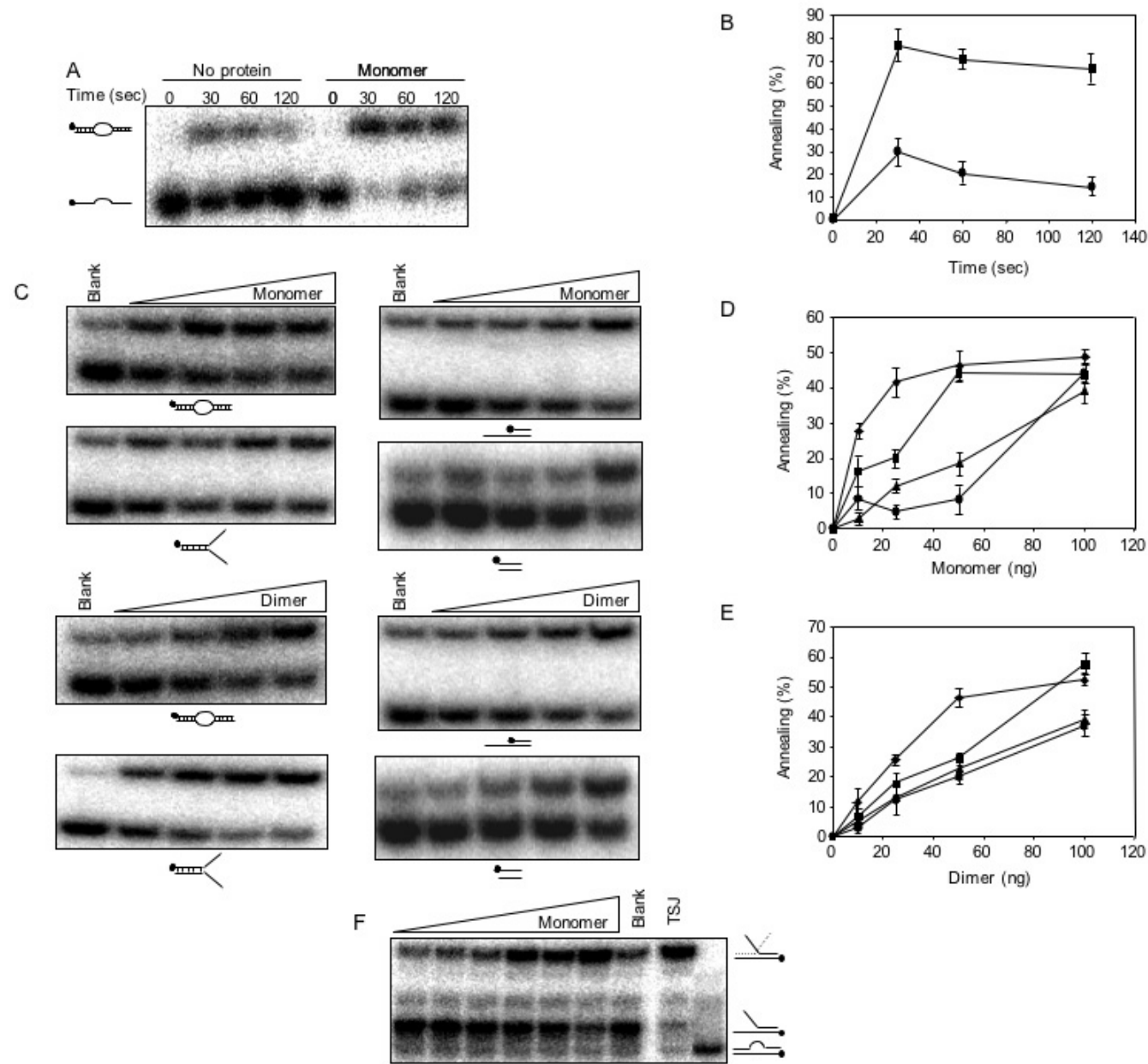


Fig. 6

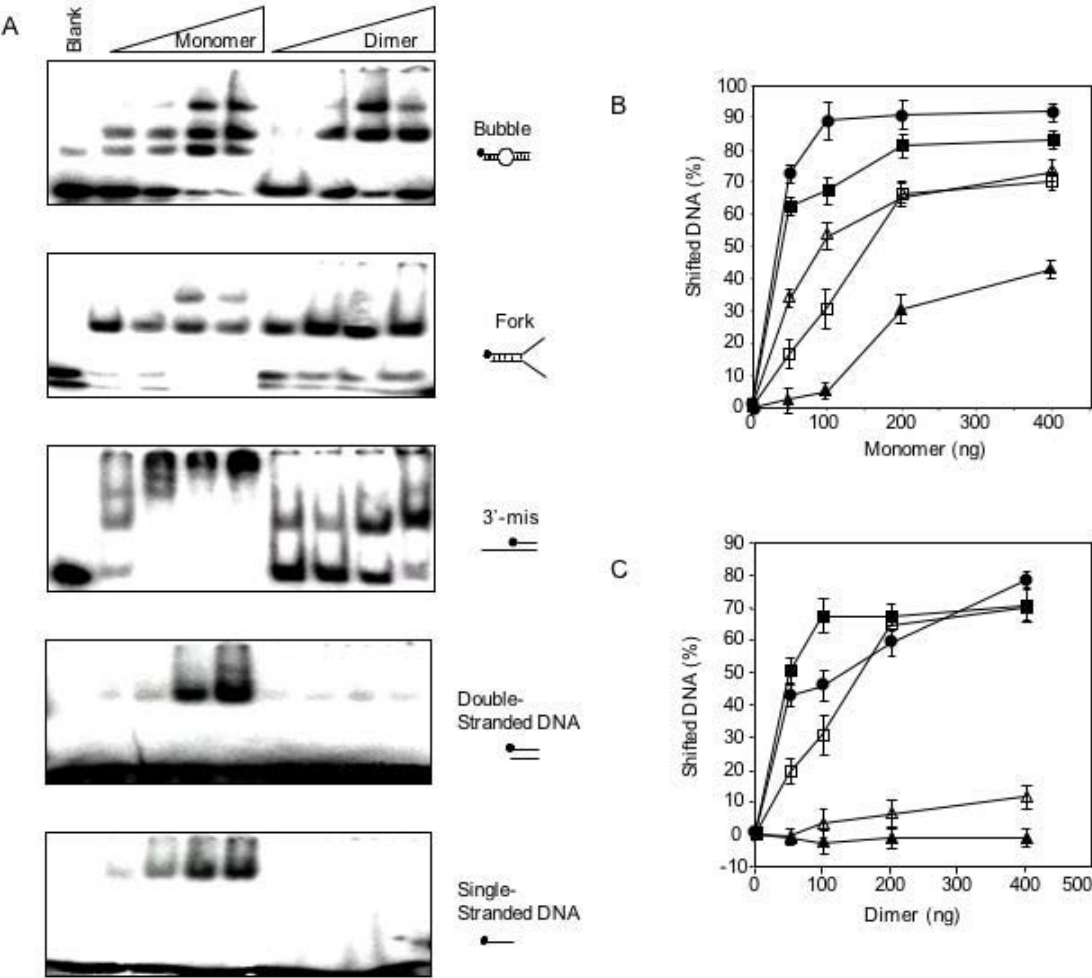


Fig. 7