Published online 9 October 2017 *Nucleic Acids Research, 2017, Vol. 45, No. 21* **12551–12564**
doi: 10.1093/nar/gkx887

The crystal structure of *Pyrococcus furiosus* RecJ implicates it as an ancestor of eukaryotic Cdc45

Min-Jun Li[1,†], Gang-Shun Yi[2,†], Feng Yu[1], Huan Zhou[1], Jia-Nan Chen[2], Chun-Yan Xu[1], Feng-Ping Wang[2,3], Xiang Xiao[2,3], Jian-Hua He[1,*] and Xi-Peng Liu[2,3,*]

[1]Shanghai Institute of Applied Physics, Chinese Academy of Sciences, No. 239 Zhangheng Road, Shanghai 201204, China, [2]State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, No. 800 Dongchuan Road, Shanghai 200240, China and [3]State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, No. 800 Dongchuan Road, Shanghai 200240, China

Received April 10, 2017; Revised September 20, 2017; Editorial Decision September 21, 2017; Accepted October 01, 2017

ABSTRACT

RecJ nucleases specifically degrade single-stranded (ss) DNA in the 5′ to 3′ direction. Archaeal RecJ is different from bacterial RecJ in sequence, domain organization, and substrate specificity. The RecJ from archaea *Pyrococcus furiosus* (PfuRecJ) also hydrolyzes RNA strands in the 3′ to 5′ direction. Like eukaryotic Cdc45 protein, archaeal RecJ forms a complex with MCM helicase and GINS. Here, we report the crystal structures of PfuRecJ and the complex of PfuRecJ and two CMPs. PfuRecJ bind one or two divalent metal ions in its crystal structure. A channel consisting of several positively charged residues is identified in the complex structure, and might be responsible for binding substrate ssDNA and/or releasing single nucleotide products. The deletion of the complex interaction domain (CID) increases the values of k_{cat}/K_m of 5′ exonuclease activity on ssDNA and 3′ exonuclease activity on ssRNA by 5- and 4-fold, respectively, indicating that the CID functions as a regulator of enzymatic activity. The DHH domain of PfuRecJ interacts with the C-terminal beta-sheet domain of the GINS51 subunit in the tetrameric GINS complex. The relationship of archaeal and bacterial RecJs, as well as eukaryotic Cdc45, is discussed based on biochemical and structural results.

INTRODUCTION

Nucleases, including endonucleases and exonucleases, hydrolyze phosphodiester bonds and play an important role in various metabolic processes of nucleic acids, such as DNA replication and repair, degradative recycling of DNA and RNA, and maturation of RNA and Okazaki fragments (1). Bacterial RecJ nuclease, as a 5′-3′ single-stranded (ss) DNA-specific exonuclease (6), mainly participates in DNA homologous recombination and mismatch repair (8,9). Structurally, most bacterial RecJ proteins identified to date, such as *Escherichia coli* RecJ, feature an N-terminal catalytic core, consisting of two domains DHH and DHHA1, and an oligonucleotide/oligosaccharide-binding (OB) fold domain, located at the C-terminus. Domains DHH and DHHA1 are interconnected by a long helix and form the catalytic core. The k_{cat} value of the catalytic core is approximately that of full-length RecJ, whereas the K_m value of the catalytic core is ∼500 times higher than that of full-length RecJ (11). These results proved that the OB fold domain mainly plays a role in improving the ssDNA-binding capability. Interestingly, a minority of bacterial RecJs identified to date have an additional C-terminal domain. For example, *Thermus thermophilus* RecJ (TthRecJ) has four domains labeled from the N- to C-terminus as domains DHH, DHHA1, OB fold and IV (11,12). The C-terminal domain IV of *Deinococcus radiodurans* RecJ (DrRecJ) can increase the 5′-3′ nuclease activity by promoting ssDNA substrate binding and interact with the HerA helicase, which promotes the nuclease activity of RecJ (13). Recently, the complex structure of DrRecJ and ssDNA revealed the 5′-3′ polarity of ssDNA substrate by DrRecJ (14). The terminal 5′-phosphate-binding pocket, which is consisted of conserved residues R109, S371 and R373, is a key factor for determining its 5′-3′ polarity on ssDNA. The OB-fold domain is also critical to the efficient hydrolysis of ssDNA, during which its residues of R475, Y496 and W517 participate in the binding of ssDNA by DrRecJ (12,14).

Little is known about archaeal RecJ nuclease, especially its functions *in vivo*. Archaeal RecJ nuclease only has two domains, which correspond to the bacterial catalytic core,

*To whom correspondence should be addressed. Tel: +86 21 3420 7205; Email: xpliu@sjtu.edu.cn
Correspondence may also be addressed to Jian-Hua He. Email: hejh@sinap.ac.cn
†These authors contributed equally to this work as first authors.

and include residues 40–425 of the *T. thermophilus* RecJ (3,11,12,15). The archaeal RecJ nuclease, such as PfuRecJ and TkoGAN, is longer than the bacterial RecJ catalytic core by ∼100 residues. In contrast to the bacterial RecJ nuclease, TkoGAN and PfuRecJ exhibit 5′-3′ exonuclease activity on ssDNA, and the latter shows 3′-5′ exonuclease activity on ssRNA (2,3,15). In addition to nuclease activity, TkoGAN also interacts with some subunits of the DNA replisome, such as the GINS complex (a key component of the archaeal DNA replication fork) via its GINS51 subunit, and the large subunit of DNA polymerase D complex (3,16). In Crenarchaeota *Sulfolobus solfataricus* encodes a protein that is homologous to the DNA binding domain of bacterial RecJ (referred as RecJdbh) and interacts with GINS complex (17).

Despite the high sequence conservation and broad distribution of RecJ nucleases in prokaryotes, no orthologue of RecJ exists in eukaryotes. Bioinformatic analysis has shown that Cdc45, an essential replication initiation protein whose site mutations result in partial defects in DNA replication (18), has significant sequence similarity to the conserved N-terminal DHH domain of RecJ-family proteins (5,19). Cdc45 lacks most of the conserved motifs and residues that are essential for bacterial and archaeal enzyme activities (19). Human Cdc45 (hCdc45) loses nuclease activity but retains the capability of binding ssDNA and ssRNA and functions as a molecular wedge for DNA unwinding (19,20). Similar to archaeal RecJ, eukaryotic Cdc45 interacts with MCM2–7 and GINS to form a complex, Cdc45–MCM–GINS (CMG), that is believed to act as the DNA helicase at the replication fork (21–25). In addition to forming the CMG complex, Cdc45 also interacts with other replication factors, such as the DNA polymerases (26). Recently, the crystal structure of hCdc45 showed that an additional peptide consisting of 100 aa residues forms a separate CMG interaction domain (CID) that interacts with the MCM2 and MCM5 subunits of MCM helicase (27). An interaction mechanism of hCdc45 with other subunits in the CMG complex were proposed based on the crystal structures of hCdc45 and cryo-EM structure of CMG (27–28).

In addition to the crystal structures of bacterial RecJs and human Cdc45, during the preparation of this paper, a crystal structure of an archaeal RecJ, known as *Thermococcus kodakarensis* GAN (TkoGAN), was reported (29). TkoGAN exhibits a topological structure similar to hCdc45, and possesses an additional 100 residues domain, termed the CID domain, that is not found in bacterial RecJ proteins. The crystal structure of TkoGAN reveals some clues as to its interaction with GINS and the MCM helicase (29). The TkoGAN forms a complex with GINS via an interaction between its DHH domain and the C-terminal B domain of GINS51 subunit. The two domains bind each other via a mixture of hydrophobic (conserved residues I140, M164, I166, V184, I186 from GINS51, and residues L313, V315, A316, L319 from TkoGAN) and hydrophilic (hydrogen bonds between D163 from GINS51 and Q56, S58 from TkoGAN) interactions. The interaction leads to a shift of DHHA1 of TkoGAN toward the N-terminal DHH by ∼10 Å (∼30° rotation), and the approach of two domains might be the reason of promoting the nuclease activity of TkoGAN by GINS.

Although the first structure of archaeal RecJ has been solved (29), the structure of archaeal RecJ with a substrate or product is still required to be solved for interpreting the catalytic mechanism, especial the binding mode of substrate or product. Furthermore the function of CID is not clear and still need to be fully elucidated. We report the crystal structure of *Pfu*RecJ, which has an overall fold similar to TkoGAN, the catalytic core of bacterial RecJ, and hCdc45. Like the structure of TkoGAN, PfuRecJ also has the corresponding CID. Our work focused on the hydrolysis mechanism of archaeal RecJ as well as the specific function of CID during the hydrolysis of oligo(deoxy)nucleotides. Our results show that the deletion of CID increases the k_{cat}/K_m of both 5′ exonuclease and 3′ exonuclease activities on ssDNA and ssRNA by five and four folds, respectively. Divalent metal ions Mg^{2+}, Mn^{2+} and Zn^{2+} can been bound in PfuRecJ crystals. However, Mn^{2+} is a preferred cofactor for nuclease activity and Zn^{2+} is an inhibitor to activity. The DNA-binding mode (a positive charged cleft) was inferred by the complex structure of PfuRecJ and CMP and confirmed by the site-directed mutations of conserved residues. Based on the biochemical and structural results, we discuss the possible function of archaeal RecJ during nucleic acid metabolism in the evolutionary context of prokaryotic RecJ and eukaryotic Cdc45.

MATERIALS AND METHODS

Materials

RNase A inhibitor was purchased from Takara (Dalian, China). KOD-plus DNA polymerase was purchased from Toyobo (Shanghai, China). Nickel-nitrilotriacetic acid resin was purchased from Bio-Rad (Shanghai, China). Oligodeoxyribonucleotides and oligoribonucleotides (Supplementary Table S1) were synthesized by Invitrogen (Shanghai, China) and Takara (Dalian, China), respectively. Expression vectors pDEST17 and pCDFDuet-1 were used throughout this study. *Escherichia coli* strain DH5α was used for cloning and Rosetta 2(DE3)pLysS strain was used to express *P. furiosus* protein. Mononucleotide CMP was purchased from Sigma. All other chemicals and reagents were of analytical grade.

Protein purification and crystallization

PfuRecJ was overexpressed and purified primarily by Ni-NTA column, as described previously (15). The affinity-purified PfuRecJ was fully dialyzed against buffer A (20 mM Tris–HCl pH 7.5, 50 mM NaCl, 5 mM β-mercaptoethanol, and 5% glycerol), and loaded onto a cation exchange HiTrap SP-sepharose HP column (GE Healthcare) preequilibrated with buffer A. PfuRecJ was eluted with a linear gradient from 50 to 1000 mM NaCl. Fractions containing PfuRecJ were pooled and concentrated using a 10-kDa Amicon Ultra-15 centrifugal filter (Millipore), and purified further using a 120 ml Hiload Superdex 200 column (GE Healthcare) with a buffer of 20 mM HEPES (pH 7.0), 100 mM NaCl, 1 mM DTT, 0.1 mM EDTA, and 2% glycerol. The fractions containing PfuRecJ were pooled and concentrated to 20 mg/ml for crystallization. Selenomethionine-labeled (SeMet) PfuRecJ was ex-

pressed using the methionine auxotrophic *E. coli* strain B834 (DE3) in a defined medium, and purified similarly to native protein.

The crystals of PfuRecJ were grown at 18°C using the hanging drop vapor-diffusion method, by mixing equal volumes of protein and reservoir solution. The reservoir solution contained 4.3% (w/v) PEG 2000 MME, 50 mM Bicine pH 8.8, 28.6% (w/v) PEG 600. After 1 week, crystals were harvested and then mounted and flash-frozen in liquid nitrogen for diffraction test and data collection. Divalent metal ions Mn^{2+} and Zn^{2+} were introduced by soaking the crystals in a solution that contains the desired metal ion. Complex crystal of PfuRecJ with ribonucleotide were obtained by soaking apo-protein crystals into solution containing CMP at 4°C.

Structure determination and refinement

The crystals were stabilized and cryoprotected by socking into a reservoir solution containing 20% glycerol and then flash cooled in liquid nitrogen. All X-ray diffraction data sets were collected at 100 K at BL17U1 of the Shanghai Synchrotron Radiation Facility. Indexing, integration, scaling and merging of the diffraction data were performed by using the HKL2000 program suites (30,31).

The structure of SeMet-labeled apo-PfuRecJ-D83A was determined using the single-wavelength anomalous dispersion (SAD) method. Other related structures of PfuRecJ, including the structure of wt PfuRecJ and these with the bound Mg^{2+}, Mn^{2+}, Zn^{2+} and CMP, were determined by molecular replacement. The initial structure was solved by the autoSHARP pipeline (32). Then, maximum likelihood-based refinement of the atomic positions and temperature factors were performed with Phenix (33). The atomic model was fit with the program Coot (34). The stereochemical quality of the final model was assessed with MolProbity (35). The data collection statistics and the refinement statistics of the PfuRecJ structures are shown in Table 1. Figures were prepared with PyMOL [Schrodinger LLC (2012) The PyMOL Molecular Graphics System, version 1.5.0.3.].

Preparation of proteins used in nuclease assay

All expression plasmids for PfuRecJ, PfuRecJ CID-deleted mutant (PfuRecJΔCID), DHH domain of PfuRecJ (PfuRecJ_N), DHH and CID fused domain of PfuRecJ (PfuRecJ_N+CID), CID domain of PfuRecJ (PfuCID), PfuGINS, PfuGINS51 and B domain of PfuGINS51 (PfuGINS51_B) were constructed according to previous methods (15). The expression of HisTag-free GINS were obtained by inserting the *gins*/*gins*51 genes just downstream the first start cordon of expression frame of vector. The expression vectors for site-directed PfuRecJ mutants were constructed based on the pDEST17-PfuRecJ vector, as described previously (15), using their respective primers (Supplementary Table S1). All proteins, including RecJ and its mutants and GINS and their truncated versions, were over-expressed and purified through immobilized-Ni^{2+} affinity chromatography as described previously (15). The HisTag-free GINS or its GINS51 subunit were used to identify the interaction between GINS and PfuRecJ by co-purification with a HisTag PfuRecJ or PfuRecJΔCID. For the RecJ-GINS complexes, the affinity-purified proteins were further purified by a 120 ml Hiload Superdex 200 column before the nuclease assay.

Nuclease assay of RecJ and its mutants

Pyrococcus furiosus RecJ and its mutants were characterized in a buffer of 20 mM Tris–HCl (pH 7.5), 30 mM NaCl, 10 mM KCl, 1 mM dithiothreitol (DTT), 1.0 mM $MnCl_2$ and 100 ng/μl BSA. The oligoribonucleotides and oligodeoxyribonucleotides used in the nuclease activity assays are listed in Supplementary Table S1. After incubation for the specified time at 50°C (Tm of the RNA/DNA hybrid is 53°C), an equal volume of a stopping buffer (90% formamide, 100 mM EDTA and 0.2% SDS) was added to the reaction. Subsequently, the reactions were subjected to 18% 8 M urea-denatured PAGE. Nucleic acid binding experiments of PfuRecJ, and PfuRecJΔCID were performed at 37°C for 10 min with the same buffer as the enzyme activity assay, but Ca^{2+} was substituted for Mn^{2+} to stop the cleavage. After loading a 1/10 volume of 50% sucrose to the reactions, the reaction mixtures were subjected to native 12% PAGE in 0.5× TBE buffer. After electrophoresis, bands in the gels were quantitated using a Typhoon 9500 fluorescent scanner. The percentage of degraded ssDNA was plotted against each incubation period to obtain the initial rates at each substrate concentration. Then, the kinetic parameters (K_m and k_{cat}) of wt and mutant RecJs and their complexes with GINS were calculated using double reciprocal plotting.

Determining the interaction between *P. furiosus* RecJ and GINS

The qualitative physical interaction between PfuRecJ or its derivatives and PfuGINS or its derivatives were analyzed through an Octet QK instrument (ForteBio, Inc.), which provides a continuous real-time display of biomolecular interactions. Biotinylated PfuGINS, PfuGINS51, and PfuGINS51_B proteins were loaded onto streptavidin biosensors. Octet analysis was performed using binding buffer (25 mM Tris–HCl pH 8.5, 50 mM NaCl, and 5 mM $MgCl_2$). Each solution of PfuRecJ, PfuRecJΔCID, PfuRecJ_N+CID, PfuRecJ_N was used to interact with the immobilized PfuGINS or its derivatives on sensors, and the binding buffer was used as blank.

Co-purification of all combinations of PfuRecJ or PfuRecJΔCID and GINS or its two subunits, GINS51 and GINS23, was used to identify the interactions between proteins. Because the soluble expression of PfuRecJ and PfuRecJΔCID was better than that of the GINS-related proteins, to achieve a better soluble expression the His tags were fused to the N-termini of PfuRecJ and PfuRecJΔCID, not to the GINS proteins.

RESULTS

PfuRecJ has an overall structure similar to TkoGAN and human Cdc45

The crystal structure of wt PfuRecJ was solved by single-wavelength anomalous diffraction with selenium (SeMet-

Table 1. Data collection and refinement statistics

Parameter	Wild-type	D83A		
		Mn	Zn	Zn + CMP
Data collection				
Wavelength(Å)	0.9792	0.9792	0.9792	0.9792
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell dimensions				
a, *b*, *c* (Å)	58.4, 66.6, 59.9	58.6, 67.1, 60.3	58.6, 67.3, 60.5	58.4, 66.2, 59.9
β (°)[b]	112.1	112.4	112.5	112.2
Resolution (Å)	50–2.80 (2.85–2.80)	50–2.10 (2.14–2.10)	50–2.03 (2.07–2.03)	50–1.75 (1.78–1.75)
No. reflections	10538	37590	45535	72646
R_{merge} (%)[a]	7.7 (24.3)	8.5 (70.4)	7.2 (17.8)	6.9 (43.7)
Mean $I/\sigma(I)$[a]	26.4 (10.2)	23.5 (10.2)	24.8 (13.5)	27.1 (4.6)
Completeness (%)[a]	99.9 (100)	100 (100)	94.1 (85.4)	97.9 (90.3)
Redundancy[a]	7.5 (7.6)	7.5 (3.1)	7.3 (6.7)	7.5 (7.2)
Refinement				
R_{work}/R_{free} (%)[b]	20.65/22.33	17.37/21.00	17.70/21.17	16.30/18.88
No. atoms				
Protein	3713	3702	3708	3722
Water	–	178	341	455
Ligand	1	2	5	66
R.M.S. Deviation				
Bond lengths (Å)	0.002	0.002	0.002	0.003
Bond angles (°)	0.420	0.473	0.432	0.624
Ramachandran plot (%)				
Favored	97.01	98.29	99.36	98.51
Allowed	2.56	1.28	0.64	1.49
Outliers	0.43	0.43	0.00	0.00

[a]The Values in parentheses are for the outermost shell.
[b] $R_{work} = \sum_{hkl}||F_{obs}| - |F_{calc}||/\sum_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are observed and calculated structure factors, respectively.
[b] R_{free}, calculated the same as R_{work}, but from a test set containing 5% of data excluded from the refinement calculation.
[b] $R_{merge} = \sum_{hkl}\sum_i|I_i(hkl) - <I(hkl)>|/\sum_{hkl}\sum_i|I_i(hkl)$, where $<I(hkl)>$ is the mean intensity of a set of equivalent reflections.

SAD). The PfuRecJ adopts a topology (Figure 1A), which is similar to those of human Cdc45 (Figure 1B), TkoGAN (Figure 1C), and bacterial RecJ (Figure 1D). Similar to TkoGAN and hCdc45, PfuRecJ possesses the CMG-Interaction Domain (CID, residues 189–281), which does not exist in bacterial RecJ (11,12). The archaeal CID consists of five α-helixes and two anti-parallel β-sheets, which are positioned at the two ends. Although the CID is located between DHH and DHHA1 domains, it is in fact inserted into the DHH domain (residues 1–320) as a separate domain (residues 189–281). Interestingly, the CID takes a similar steric orientation to the OB-fold domain of TthRecJ nuclease (11). However, compared with archaeal CID, the bacterial OB-fold occurs after DHHA1 domain and is quite far from the catalytic core, which consists of DHH and DHHA1 domain (Figure 1D).

Due to the large conformational deviation of the linker region (residues 282–351), which is characterized by a long α-helix, the relative positions of the conserved domains (DHH, CID and DHHA1) are very different among the reported RecJ proteins (Figure 1). Although archaeal PfuRecJ and TkoGAN share a high sequence identity (72.8% by Clustal O), and their DHH-CID domains can be well superimposed with RMSD of 0.7 Å over 219 C_α atoms, the long 'linker' α-helix of PfuRecJ makes a ∼65° bending towards DHH domain (Figure 1E). This remarkable conformation shift makes the substrate-binding domain DHHA1 close to the catalytic domain DHH to form a closed conformation for substrate hydrolysis.

However, despite that PfuRecJ has little sequence identity with human Cdc45 (15.3% by Clustal O), it shares a remarkable 3D conformational similarity with its eukaryotic homology (Figure 1F). Their DHH-CID domains can be well superimposed with RMSD of 3 Å over 156 C_α atoms. Furthermore, the linker α-helices of Cdc45 is almost parallel to that of PfuRecJ, which lead to the overall structure of both protein in the similar closed conformation. The motif VI GGGHxxAAG corresponding to the last β-sheet in the structures of PfuRecJ and TkoGAN has been mutated completely to a loop in hCdc45 (Supplementary Figure S1). This DHHA1 loop of hCdc45 is proposed to mediate the intramolecular interaction between the N-terminal DHH and C-terminal DHHA1 domains by inserting the conserved large hydrophobic residue F542 of DHHA1 into a hydrophobic pocket on the DHH domain surface (27). Instead in PfuRecJ, the residue H440 from DHHA1 domain interacts with the residue S130 from DHH domain via a hydrogen-bond within 2.3 Å. This interaction and the substrate binding probably contribute to the closed conformation of DHH and DHHA1 domains in PfuRecJ. Compared with structures of three proteins, it is found that the PfuRecJ is more similar to its eukaryotic homolog Cdc45 in three dimensional structure (Figure 1 and Supplementary Figure S2).

After superimposing the structures of bacterial and archaeal N-terminal DHH domain, we found that there is a 14 Å gap between domain DHH and DHHA1 in archaeal RecJ that results from the deviation of the long alpha helix, creating a narrower ssDNA-binding groove than that of

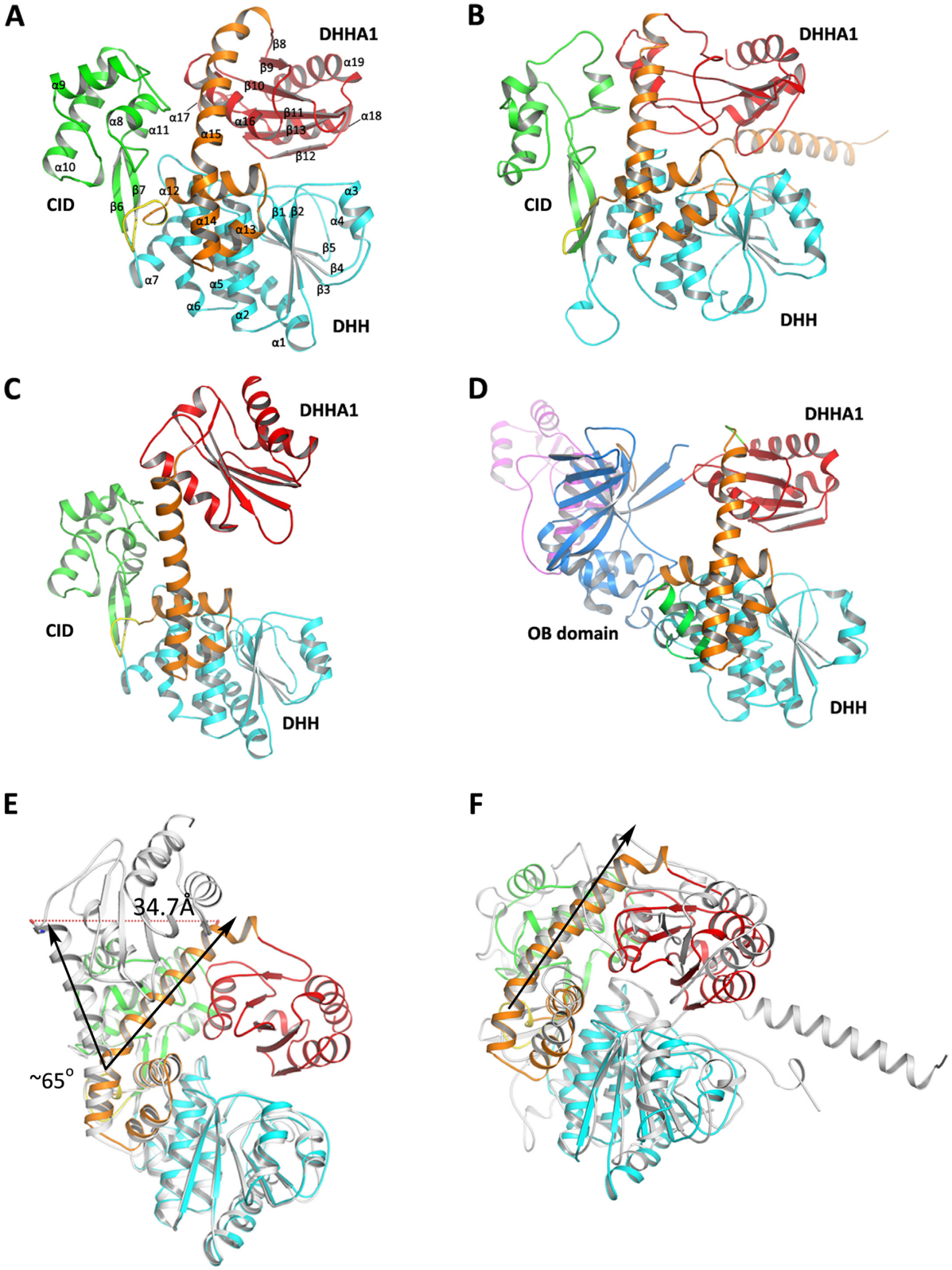

Figure 1. The crystal structures of archaeal *Pyrococcus furiosus* (PfuRecJ) and its homologs. The crystal structure of archaeal PfuRecJ (**A**), human Cdc45 (**B**, PDB ID: 5DGO), archaeal TkoGAN (**C**, PDB ID: 5GHT) and bacterial TthRecJ (**D**, PDB ID: 2ZXP) are shown as cartoon model. The DHH domain of these proteins is colored as cyan, CID domain as green, linker region as orange, and DHHA1 domain as red. The secondary structures of PfuRecJ are labeled according to DSSP analysis. (**E**) The 3D superimposing of PfuRecJ (colored) and TkoGAN (grey). Compared with its relatives from TkoGAN, the long linker α-helix of PfuRecJ adopts a large conformational change, which leads to the substrate-binding domain DHHA1 maintained in a 'closed' conformation related to the catalytic domain DHH. (**F**) The 3D superimposing of PfuRecJ (colored) and Human Cdc45 (grey) (5DGO).

bacterial RecJ (Supplementary Figure S3A). There are two hydrogen bonds responsible for the deviation of the alpha helix. One is formed between the side chain N atom of Q338 and the hydroxyl of Y334 (Y335 in TkoGAN); the other is formed between the side chain O atom of Q338 and amino N atom of K335 (Supplementary Figure S3B). The disruption of the hydrogen bonds between Q338 and Y334/K335 increases the nuclease activities on both ssDNA and ssRNA by about 50% (Supplementary Figure S3C). A possible reason for the increased activity of Q338A is that the disruption of hydrogen bond results in a wide gap, which is better for the entry of substrates.

Divalent metal ion binding sites

Although archaeal and bacterial RecJs show some sequence conservation with hCdc45 (Figure 2A and Supplementary Figure S1), the residues that coordinate catalytic metal ions in prokaryotic RecJs are not completely conserved in Cdc45. Residues of prokaryotic RecJs (D136 and H160 in TthRecJ, D83 and H106 in PfuRecJ) are changed to N76 and T100, respectively, in hCdc45 (14,27). The structures of the PfuRecJ D83A mutant complexed with metal ions were the same as those of wt PfuRecJ, as shown (Figure 2B–D). In the substrate-free state, only one Mn^{2+} or Mg^{2+} was found in the active site of PfuRecJ (Figure 2B and C); however, two Zn^{2+} were identified in the metal ion binding site (Figure 2D). The key residues involved in metal ion binding were identified. Mn^{2+}, Mg^{2+}, and one Zn^{2+} were bound by the same residues (D36, H106, and D165) but the second Zn^{2+} was bound by H32 and D34 (Figure 2D). From the structure of metal-bound PfuRecJ, D83 was shown to possibly be involved in binding the first ions.

When these conserved residues, which are required to bind divalent metal ions, were mutated to alanine, all the mutants displayed significantly reduced activity on ssDNA and ssRNA substrates (Figure 2E). These results indicated both the residue of D36A, which is required to bind all three kinds of metal ions, and the residues of H32 and D34, that are responsible for binding the second Zn^{2+}, were essential for the nuclease activity (Figure 2E).

Archaeal RecJs have conserved product-binding motifs

To characterize the hydrolysis mechanism of oligonucleotides, we tried to resolve the structure of co-crystals of PfuRecJ and oligonucleotides. Despite repeated efforts, a co-crystal of PfuRecJ and ssDNA or ssRNA was not obtained. Soaking 3 nt ssDNA into apo-PfuRecJ crystals also failed. Finally, soaking ribonucleotide CMP into the apo-PfuRecJ crystal was successful. Two single nucleotides are bound by several positive conserved residues that are located in the substrate-binding cleft (Figure 3A). The first CMP is bound tightly via hydrogen bonds with three residues (K406, S408 and R410) conserved in archaeal RecJs, while the second CMP is bound loosely, forming hydrogen bonds with only one residue His440 (Figure 3B). Because only single nucleotides, not oligonucleotides, are bound in the PfuRecJ structure, it is difficult to confirm that these residues are responsible for binding substrate oligonucleotides or product single nucleotide. The site-directed mutations show that the effects of nucleotide-binding residues on the hydrolysis of ssDNA and ssRNA are different (Figure 3C). The mutants of residues (K406, S408 and R410) that bind the first CMP have higher activities on ssRNA and ssDNA (Figure 3C, lanes 5 and 6). The H440A mutant lost almost all activity on ssDNA and ssRNA, indicating that H440 is critical for hydrolyzing both substrates. These results suggest that the first ribonucleotide CMP1 is a product and the second ribonucleotide CMP2 is possibly the first nucleotide of a long oligonucleotide substrate that is in state of being hydrolyzed.

The conserved positively charged residues responsible for binding ssDNA

As the initial and an essential step, oligonucleotide binding is the basis for hydrolyzing the phosphodiester bond. To identify the key residues that bind ssDNA, the structure of the PfuRecJ-CMP complex was compared with the structure of the DrRecJ-ssDNA complex (Figure 4A). Based on the functional similarity of residues conserved between the two RecJs, some conserved positively charged residues of PfuRecJ, including N302, T304, R414 and R306, strongly interacted with substrate via hydrogen bonds, salt bridges etc. These residues are also conserved in bacterial RecJs (11,12,14). Therefore, we mutated them into alanines to check their effects on nuclease activity. All the tested mutants, N302A+T304A, R414A and N302A+R306A, demonstrate clearly decreased activity on ssDNA (Figure 4B, left panel). In comparison with the ssDNA substrate, the effect of mutation on hydrolyzing the ssRNA substrate was clearly different. Although both mutants R414A and N302A + R306A completely abolished the activity, the N302A + T304 mutant hydrolyzed the ssRNA with a higher activity than the wt RecJ (Figure 4B, right panel, lanes 5 and 6). The different effects of the same residues on binding ssDNA and ssRNA substrate suggest that the PfuRecJ requires a different binding model for two substrates. The R307A mutant also abolished the activity of TkoGAN, which is the same as the mutant of N302A and R306A. The corresponding residues R280 and R373 of bacterial DrRecJ also involve in the binding of ssDNA and their mutation lead to the inactivation of nuclease (14).

The DHH domain, not CID, interacts with GINS

Similar to Cdc45, but unlike bacterial RecJ, archaeal RecJs intrinsically interact with the GINS complex (3,16). The crystal structure of the complex of TkoGAN and B domain of TkoGINS51 showed that the interaction unquestionably exists between the DHH domain of TkoGAN and the B domain of TkoGINS51 (29). Since PfuRecJ and TkoGAN, as well as PfuGINS and TkoGINS, have a sequence similarity higher than 70%, especially the same conserved residues for interaction, the PfuRecJ and PfiGINS should have the same interaction mode. We also confirmed the interaction between DHH domain of PfuRecJ and B domain of PfuGINS51 through an Octet QK instrument (Supplementary Figure S4), which is the same as that of TkoGAN and TkoGINS (29). We further characterized the function of CID during the interaction with GINS. The deletion of CID has no effect on forming the PfuRecJ–GINS complex. Both the

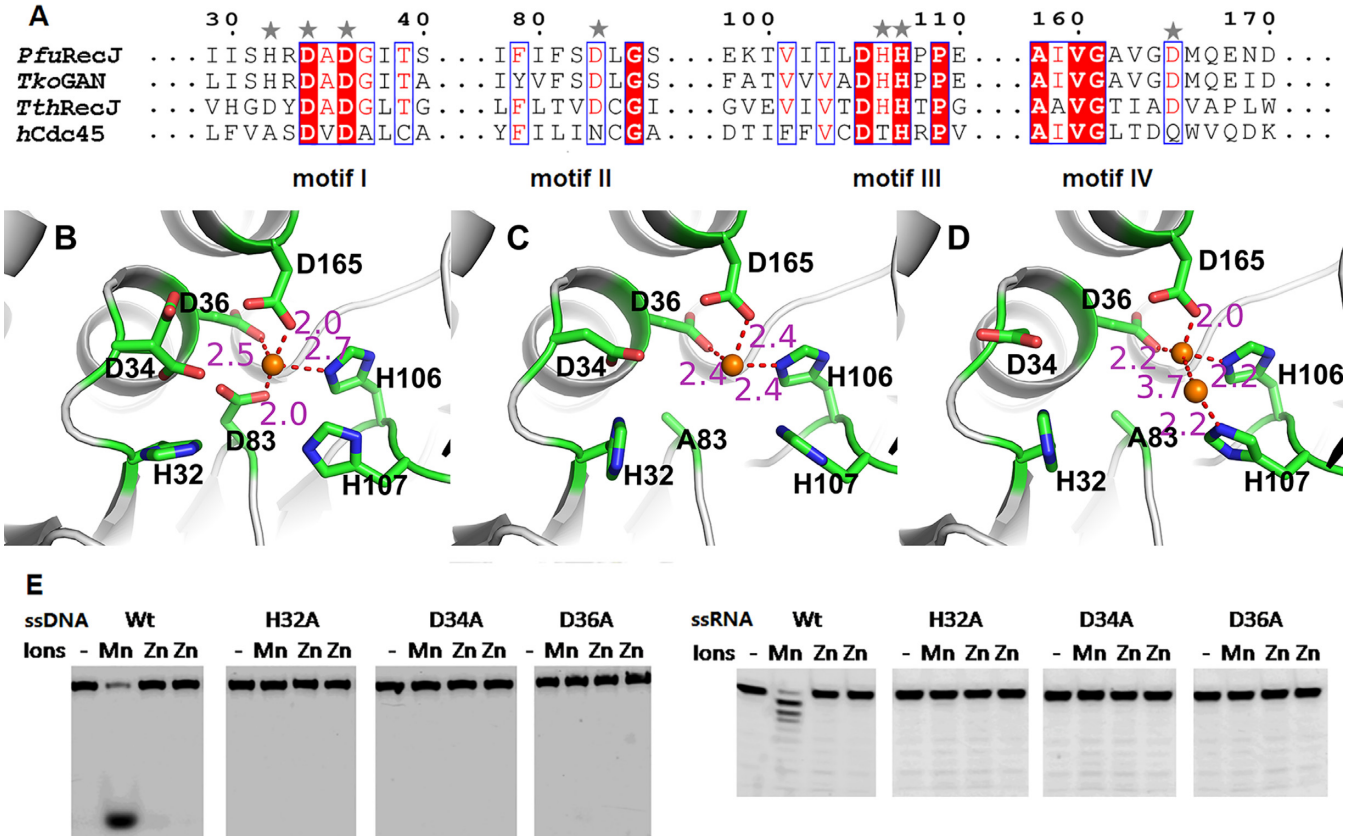

Figure 2. Comparison of binding model of different divalent ions. (**A**) Multialignment of prokaryotic RecJ and hCdc45. The conserved motifs involved in binding the divalent metal ions are compared. (**B**) Structure of wt *Pfu*RecJ with a magnesium ion bound. (**C**) Structure of *Pfu*RecJ D83A with a manganese ion bound. (**D**) Structure of *Pfu*RecJ D83A with two zinc ions bound. The residues forming the ion-binding sites are shown in stick form. (**E**) The exonuclease activity of wt and mutant PfuRecJs on ssDNA and ssRNA. The activities were determined with 42 nt ssDNA and 16 nt ssRNA as substrates in a buffer consisting of 20 mM Tris–HCl (pH 7.5), 30 mM NaCl, 10 mM KCl, 1 mM DTT, 100 ng/μl BSA and 4 U Rnsin. The mutations and added divalent ions are listed at the top of each lane. Substrates (50 nM) were incubated with 50 nM wt PfuRecJ or 100 nM mutant PfuRecJs at 50°C for 30 min. Mn^{2+} (1 mM) or Zn^{2+} (0.1, 1 mM) were included in reaction buffer.

HisTag full-length and CID-deleted PfuRecJs could pull-down the GINS complex and its GINS51 subunit (Figure 5A), indicating that the CID is not involved in the association of GINS and PfuRecJ.

Since the archaeal GINS complex can stimulate the 5′ exonuclease activity of euryarchaeal RecJ (3,15), it must be confirmed whether the CID is involved in stimulating the nuclease activity of PfuRecJ by GINS. Although the complexes of PfuRecJ-GINS and PfuRecJ-GINS51 show higher 5′ exonuclease activity on ssDNA (Figure 5B, left panel), the PfuRecJΔCID-GINS and PfuRecJΔCID-GINS51 do not clearly show increased activity (Figure 5B, right panel), suggesting that CID is an essential element for stimulating nuclease activity by GINS in spite of not being required for subunit interaction. In comparison with ssDNA substrate, the hydrolysis of ssRNA is more weakly promoted by GINS (Figure 5C).

CID functions as a negative regulator of nuclease activity

Although the additional CID is located near the possible entry of ssDNA and ssRNA substrates, its exact function is still unknown. To determine its function, especially during the substrate binding and/or catalytic steps, we deleted it and characterized the effect on the nuclease activities on ssDNA and ssRNA. Since the CID takes a similar spatial orientation as the OB-fold in the bacterial RecJ structure (11), we speculated that it had a function similar to bacterial OB-fold, which stimulates the nuclease activity and is involved in binding ssDNA (11). Our results showed that, very interestingly, the removal of CID increases the 5′ exonuclease activity on ssDNA (Figure 6A, left panel). The kinetic parameters of PfuRecJ and PfuRecJΔCID showed that the value of k_{cat}/K_m of PfuRecJΔCID increases about 5 times (Table 2, the representative gel images for calculated the kinetic parameters of PfuRecJ and PfuRecJΔCID were shown in Supplementary Figure S5). In comparison with ssDNA, the promotion on ssRNA is lower (Figure 6A, right panel), with a 3.9 times increase of the value of k_{cat}/K_m (Table 2). The deletion of CID domain also largely decreased the promotion of nuclease activity by GINS or GINS51 (Figure 6B, right two panels), suggesting that the promotion is dependent on the CID domain of PfuRecJ. In comparison with ssDNA, the promotion on ssRNA digestion by GINS/GINS51 is so weaker that no clear promotion can be detected (Figure 6B, four panels at bottom). Meanwhile, the presence of high concentrations of a single peptide CID has no clear inhibitory effect on the nuclease

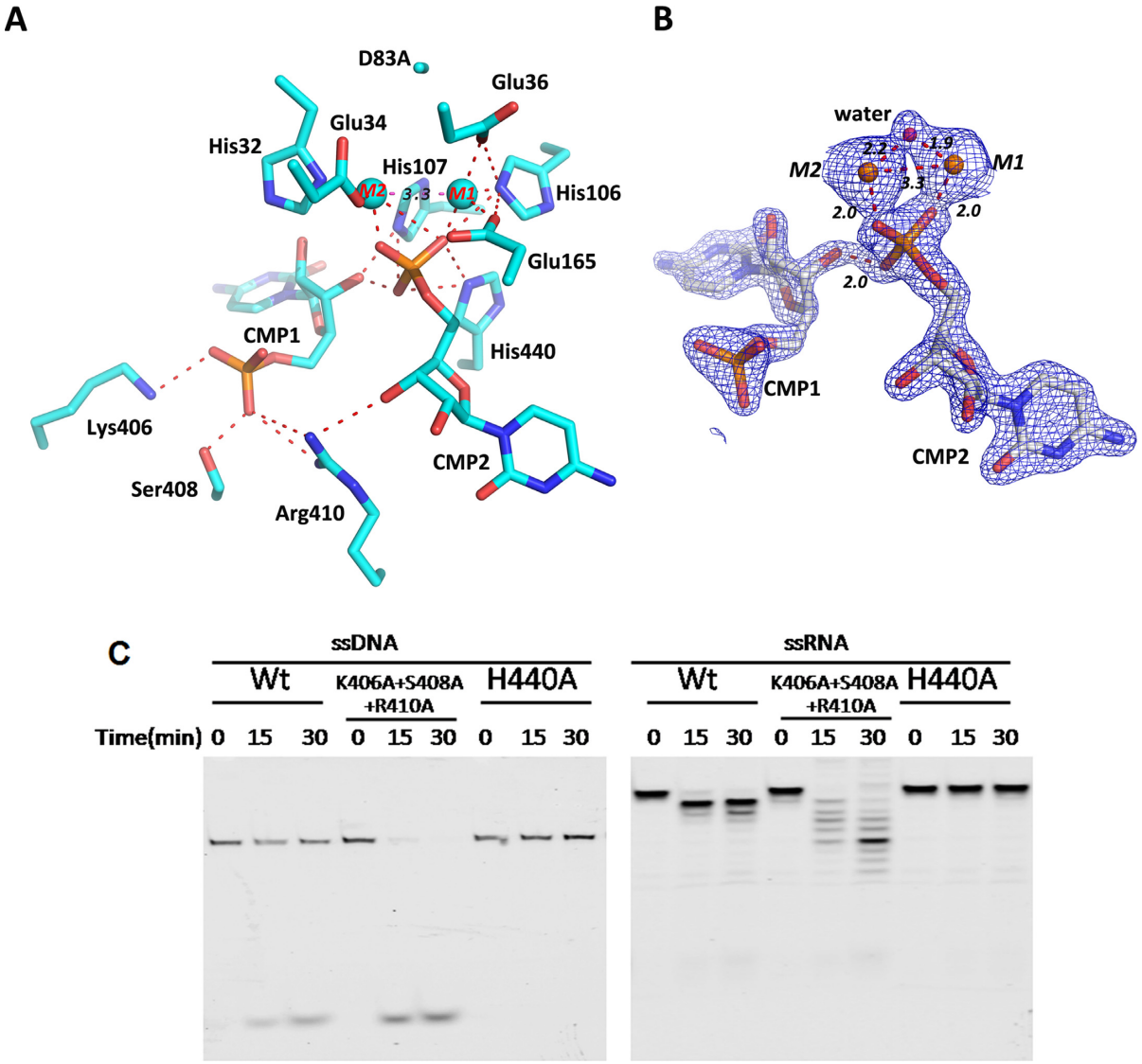

Figure 3. Structures of the complex of PfuRecJ and ribonucleotides. (**A**) The binding pocket and the position of bound ribonucleotides in PfuRecJ. CMP is shown in stick form. The residues in the structure of PfuRecJD83A–Zn^{2+}–CMP complex are shown as cyan sticks. The hydrogen bonds between CMP and residues are shown as red dash lines. The coordination bonds between Zn^{2+} and residues are also shown as red dashed lines. The spheres M1 and M2 denote the two zinc ions. (**B**) The electron-density map of two ribonucleotides bound in *Pfu*RecJ. The $2F_o - F_c$ map is contoured at 1.5σ and shown as blue mesh. The color scheme of the hydrogen bonds are the same as listed in (A). (**C**) The exonuclease activity of wt and mutant PfuRecJs on ssDNA and ssRNA. The activities were determined with 42 nt ssDNA and 16 nt ssRNA as substrates in a buffer consisting of 20 mM Tris–HCl (pH 7.5), 30 mM NaCl, 10 mM KCl, 1.0 mM Mn^{2+}, 1 mM DTT, 100 ng/μl BSA, and 4 U Rnsin. Site-directed PfuRecJ mutants include CMP1-binding mutant K406A + S408A + R410A and CMP2-binding mutant H440A. Substrates (50 nM) were incubated with wt or mutated PfuRecJ (50 nM) at 50°C for 0, 15 and 30 min.

activities of PfuRecJ and PfuRecJΔCID (Figure 6C). CID also did not bind substrate ssDNA (data not shown). The EMSA showed that the PfuRecJΔCID bound the ssDNA with a higher affinity (Figure 6D). The removal of CID also weakly increases the activity on other substrates, including the RNA/DNA hybrid (Supplementary Figure S6). Based on our results, we propose that the CID functions as a special regulatory domain for nuclease activity. The k_{cat} of PfuRecJΔCID is almost the same as that of PfuRecJ, but the K_m of truncated RecJ is five times smaller, indicating that the removal of CID facilitates the entry and binding of substrates.

Consistent with the promotion of nuclease by GINS, the complexes of PfuRecJ-GINS/GINS51 have higher values of k_{cat}/K_m than PfuRecJ alone, but a little lower than that of PfuRecJΔCID (Table 2). However, PfuRecJΔCID and its complexes with GINS/GINS51 have the similar values of k_{cat}/K_m, indicating that the promotion is dependent on the CID domain. Although the GINS51_B domain forms complex with PfuRecJ (Supplementary Figure S4 and ref. 29), it does not promote the nuclease activity of PfuRecJ. In other words, the complexes of PfuRecJ-GINS51_B has the same value of k_{cat}/K_m as that of PfuRecJ (Table 2).

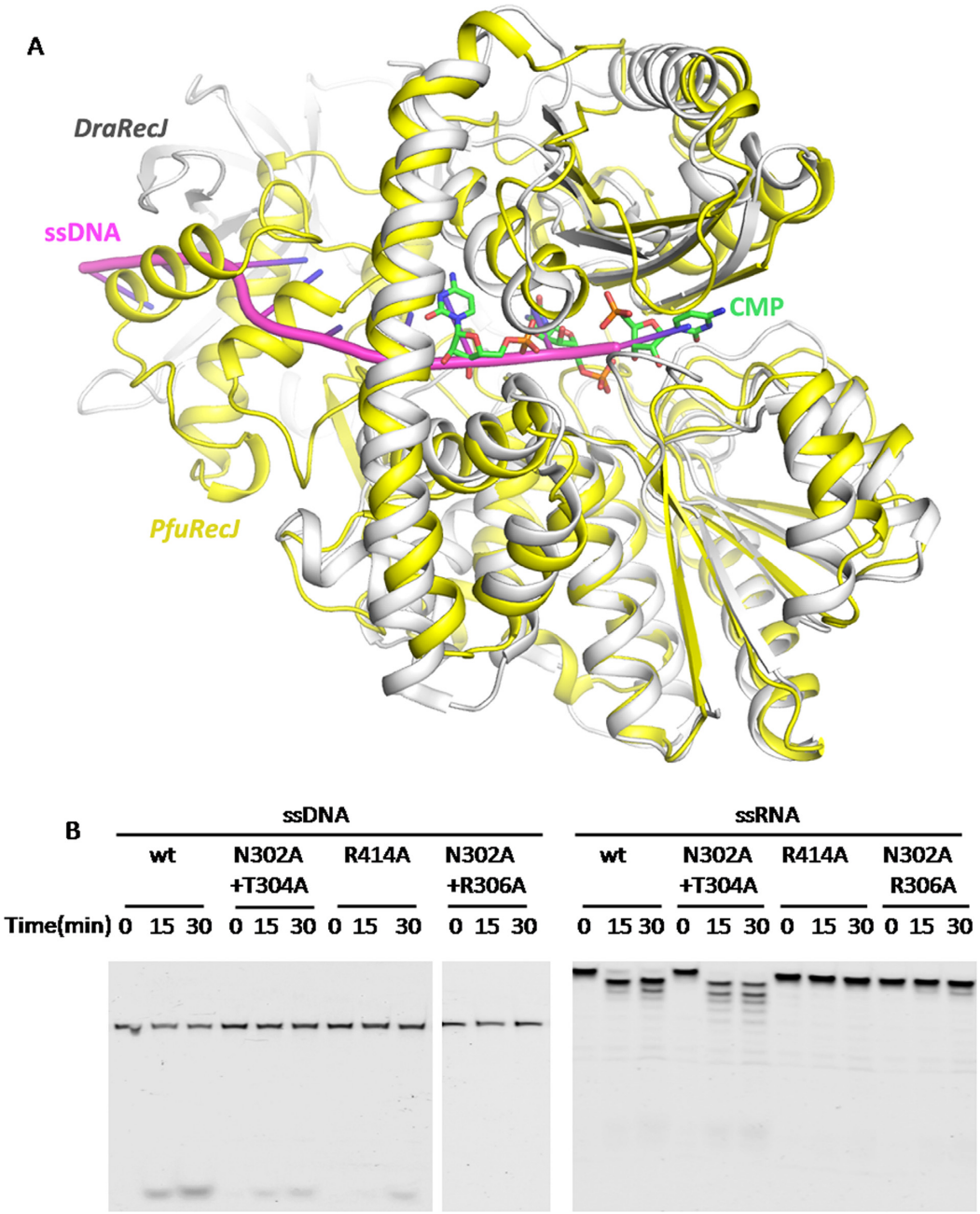

Figure 4. Oligonucleotides -binding model and potential hydrolysis mechanism. (**A**) The comparison of the ssDNA-binding pocket of archaeal PfuRecJ and bacterial DrRecJ. (**B**) The exonuclease activity of wt and mutant PfuRecJs on ssDNA and ssRNA. The activities were determined with 42 nt ssDNA and 16 nt ssRNA as substrates in a buffer consisting of 20 mM Tris–HCl (pH 7.5), 30 mM NaCl, 10 mM KCl, 1.0 mM Mn^{2+}, 1 mM DTT, 100 ng/μl BSA and 4 U Rnsin. Site-directed mutants include mutations to three groups of ssDNA-binding residues. Substrates (50 nM) were incubated with wt or mutated PfuRecJ (50 nM) at 50°C for 0, 15 and 30 min.

DISCUSSION

Function and evolution of RecJ and Cdc45

Different from the bacterial RecJ, archaeal RecJ is the only member of RecJ family with an additional CID between motif IV and V of DHH domain (Figure 1). Although the CID of archaeal RecJ occupies a similar spatial position to the OB-fold domain of bacterial RecJ (Figure 1A, C and D), its function on nuclease activity is very different from that of the OB-fold. The OB-fold of bacterial RecJ is a domain that strongly promotes the nuclease activity via enhancing binding of ssDNA to RecJ (14). Deletion of the OB-fold increases the K_m value of TthRecJ by a factor of ∼500 (11). In contrast, the CID of archaeal RecJ is an inhibitory domain of nuclease activity. Removal of CID decreases the K_m of PfuRecJ by approximately 4–5 times (Table 2). Since the promotion of nuclease by GINS is dependent on the CID domain, here we propose a functional model of the

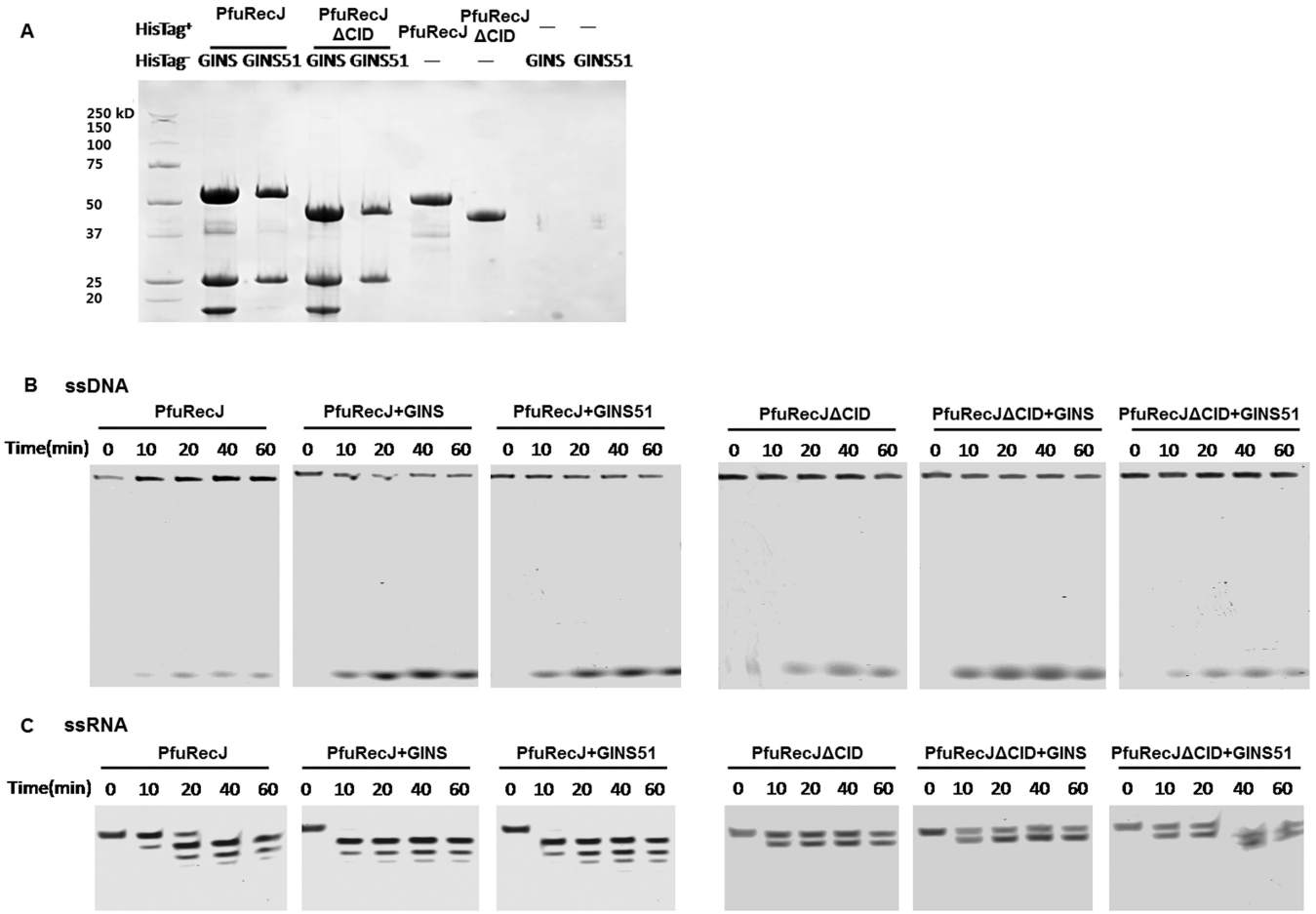

Figure 5. Domain CID is not responsible for associating with GINS but involved in promoting nuclease. (**A**) Pulldown of GINS/GINS51 by HisTag PfuRecJ or PfuRecJΔCID. The pulldowns were performed in the presence of excess GINS or GINS51. Purification of HisTag proteins PfuRecJ and PfuRecJΔCID and HisTag-free proteins GINS and GINS51, which were produced by inserting the *gins*/*gins*51 genes just downstream the first start cordon of expression frame of vector, were performed alone as controls. The exonuclease activity of PfuRecJ/PfuRecJΔCID and their complexes with GINS/GINS51 on 42 nt ssDNA (**B**) and 16 nt ssRNA (**C**) was measured. Activities were determined in a buffer consisting of 20 mM Tris–HCl (pH 7.5), 30 mM NaCl, 10 mM KCl, 1.0 mM $MnCl_2$, 1 mM DTT, 100 ng/μL BSA and 4 U Rnsin. Substrates (50 nM) were incubated with 50 nM enzymes at 50°C for 0, 10, 20, 40 and 60 min.

Table 2. Dynamitic parameters of PfuRecJ and PfuRecJΔCID

Substrates	Proteins	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m (min^{-1}•μM^{-1})
ssDNA	PfuRecJ	0.21 ± 0.03	0.48 ± 0.04	2.3 ± 0.3
	PfuRecJΔCID	0.043 ± 0.003	0.49 ± 0.04	11.4 ± 0.07
	PfuRecJ-GINS	0.13 ± 0.02	0.51 ± 0.02	3.9 ± 0.3
	PfuRecJΔCID-GINS	0.041 ± 0.05	0.46 ± 0.04	11.2 ± 0.6
	PfuRecJ-GINS51	0.13 ± 0.02	0.47 ± 0.03	3.6 ± 0.3
	PfuRecJΔCID-GINS51	0.042 ± 0.03	0.48 ± 0.04	11.4 ± 0.8
	PfuRecJ-GINS51_B	0.20 ± 0.03	0.48 ± 0.04	2.4 ± 0.3
	PfuRecJΔCID-GINS51_B	0.042 ± 0.003	0.47 ± 0.04	11.2 ± 0.07
ssRNA	PfuRecJ	0.47 ± 0.04	0.23 ± 0.02	0.49 ± 0.05
	PfuRecJΔCID	0.12 ± 0.01	0.23 ± 0.02	1.9 ± 0.18
	PfuRecJ-GINS	0.44 ± 0.03	0.25 ± 0.02	0.57 ± 0.05
	PfuRecJΔCID-GINS	0.13 ± 0.05	0.24 ± 0.03	1.8 ± 0.12
	PfuRecJ-GINS51	0.46 ± 0.03	0.23 ± 0.02	0.50 ± 0.04
	PfuRecJΔCID-GINS51	0.12 ± 0.06	0.22 ± 0.03	1.8 ± 0.15
	PfuRecJ-GINS51_B	0.45 ± 0.04	0.23 ± 0.02	0.51 ± 0.05
	PfuRecJΔCID-GINS51_B	0.12 ± 0.01	0.23 ± 0.02	1.9 ± 0.2

K_m and k_{cat} were calculated by double reciprocal plotting using the initial reaction rates of ssDNA and ssRNA at various substrate concentrations (0.02, 0.05, 0.1, 0.2 and 0.5 μM). The initial rates were presented as the percentage of degraded substrate per min. Experiments were performed in triplicate, and the results are presented as an averaged value with the mean standard error from three independent experiments.

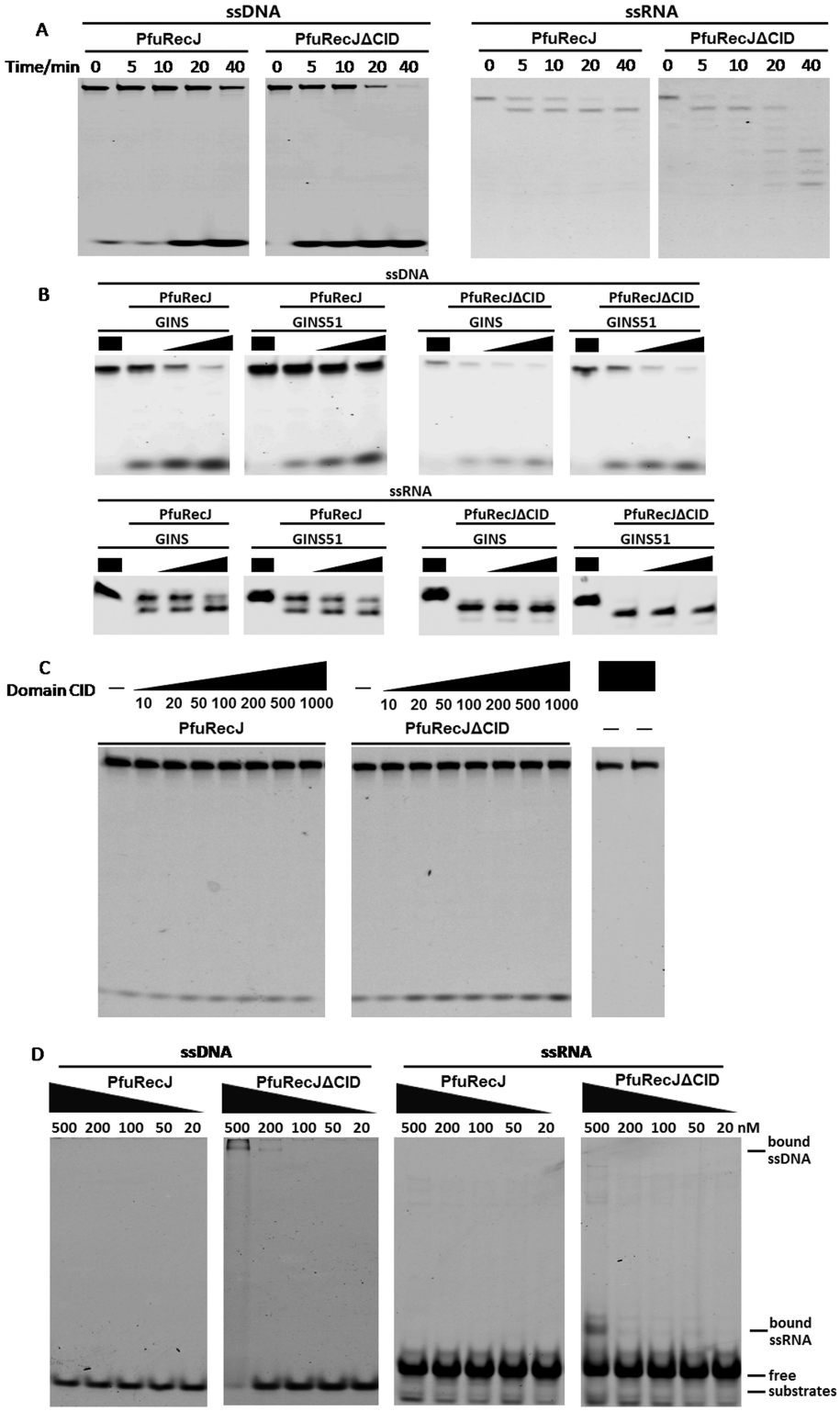

Figure 6. Domain CID functions as a negative regulator for nuclease activity. (**A**) The exonuclease activities of wt PfuRecJ and PfuRecJΔCID on 42 nt ssDNA and 16 nt ssRNA. The activities were determined in a buffer consisting of 20 mM Tris–HCl (pH 7.5), 30 mM NaCl, 10 mM KCl, 1.0 mM Mn^{2+}, 1 mM DTT, 100 ng/μl BSA and 4 U Rnsin. Substrates (50 nM) were incubated with wt or truncated PfuRecJ (50 nM) at 50°C for 0, 5, 10, 20 and 40 min. (**B**) Determination of the promotion of nuclease activity by GINS or itsGINS51 subunit. PfuRecJ (20 nM) or PfuRecJΔCID (20 nM) were incubated with 50 nM ssDNA or ssRNA at 50 °C for 20 min in the absence/presence of GINS or GINS51 subunit. The concentrations of GINS or GINS51 are 20 and 50 nM, respectively. **(C)** The effect of recombinant CID on the activity of PfuRecJ and PfuRecJΔCID. The activity was determined in the presence of increasing amounts of recombinant CID. (**D**) The EMSA of PfuRecJ and PfuRecJΔCID. Increasing amounts of PfuRecJ and PfuRecJΔCID were incubated with 50 nM ssDNA or ssRNA at 50°C for 15 min in the same buffer as the activity assay, except that Mn^{2+} was replaced by Ca^{2+}, which supports binding but not catalysis for PfuRecJ.

PfuRecJ-GINS complex to interpret the mechanism of promoting nuclease activity by the GINS complex. After binding to PfuRecJ, GINS or GINS51 will twist the CID aside. This shifting of CID causes PfuRecJ to take a conformation similar to PfuRecJΔCID and hydrolyze the ssDNA at a higher efficiency. The PfuRecJΔCID-GINS and PfuRecJ-GINS complex have almost equal values of K_m and k_{cat}/K_m to those of PfuRecJΔCID, providing indirect evidence supporting our functional model.

Both archaeal RecJ and eukaryotic Cdc45 bind ssDNA and ssRNA (15,19), but the latter protein loses the exonuclease activity and only functions as a wedge to unwind dsDNA in chromosome replication (20). Since the DNA-binding groove is similar in both archaeal RecJ and Cdc45 (27,29), the mutations of conserved motifs, which is responsible for binding the divalent metal ions in archaeal RecJ, are a possible reason for Cdc45 lacking nuclease activity. In comparison with archaeal RecJ, Cdc45 has an additional polypeptide sequence inserted between motif III and IIIa (19,27), which is an additional possible factor for abrogating the nuclease activity, in a similar manner to the down-regulation of exonuclease activity of PfuRecJ by CID. Since human Cdc45 exhibits a structural fold more similar to archaeal RecJ than bacterial RecJ (13,19), we speculate that Cdc45 originates from archaeal RecJ by the insertion of another domain (the light brown line in Supplementary Figure S1) between motifs III and IIIa. Furthermore Cdc45 might possess a nuclease activity for a long time during the evolution of eukaryotic CMG (36).

Similar to eukaryotic Cdc45, unlike euryarchaeal RecJs, the RecJdbh proteins from *Sulfolobus* genus are no nuclease activity and also referred as archaeal Cdc45 (17,37). In *Sulfolobus* GINS complex forms the CMG complex, via interacting with archaeal Cdc45 and MCM helicase, which might function as a replicative DNA helicase (37). Similar to hCdc45, both archaeal Cdc45 and TkoGAN interact with the B domain of GINS51 through their DHH domain (29,37). Since the PfuRecJ and TkoGAN have a sequence similarity higher than 70% (Supplementary Figure S1), they might share a similar interaction mechanism with GINS.

Archaeal CMG might participate in both DNA replication and repair

Recent works on *T. kodakarensis* demonstrated that GAN can be deleted with no discernable effects on viability and growth, indicating that it is not essential to the archaeal MCM replicative helicase (38). Like the *recj* gene in *T. Kodakarensis*, the two *recj* genes are also non-essential in *Haloferax volcanii* (39). However, it is not clear whether all four Haloferax RecJ proteins are non-essential. Because deleting both *recj* and *fen1* genes is impossible in one *T. Kodakarensis* cell, it was proposed that TkoGAN participate in RNA primer removal during Okazaki fragment maturation coordinated with the Fen1 nuclease (38). Similar to TkoGAN, PfuRecJ might remove the RNA primer by its 5′-exonuclease on the flapped RNA section of Okazaki fragment. However, the function of *recJdbh* gene is yet to be confirmed in Crenarchaea, because of lacking of the *recjdbh*-knockout mutant Strain (37).

The two RecJs from *M. jannaschii* can complement the function of the deleted *recj* gene during DNA recombination repair in *E. coli* (2), suggesting that archaeal RecJ function as 5′-3′ exonuclease during DNA recombination repair. Therefore, archaeal RecJ, at least in euryarchaeota, might participate in resecting the dsDNA end via coordinating with MCM helicase during DNA recombination repair, similar to the coordination of RecJ and RecQ helicase in bacteria (40). Considering that there are several different DNA resection pathways in prokaryotes (41), the archaeal CMG might be an alternative resecting process similar to the Mre11-Rad50 pathway that occurs during recombination repair of dsDNA breaks (42). Given that the archaeal CMG is a simplified counterpart of eukaryotic CMG (3,16,17), it is plausible that the archaeal CMG takes a quaternary structure similar to eukaryotic CMG (43), but with some specific interaction surfaces and partners that are, respectively, different.

Hydrolysis direction of PfuRecJ

In addition to its 5′-3′ exonuclease activity on ssDNA, PfuRecJ can also hydrolyze ssRNA in the 3′ to 5′ direction. Our results show that the residues mainly responsible for binding ssDNA are conserved between bacterial and archaeal RecJ. Residues N302, R306 and R414 are key for binding ssDNA. Among these, the residue N302 is particularly interesting. Its mutation decreases the activity on ssDNA, but increases the activity on ssRNA, suggesting that N302 crucially determines the hydrolysis direction and catalytic mechanisms of the two substrates. Because there is no crystal structure of PfuRecJ bound to ssDNA/ssRNA, it is difficult to clearly identify the key residues or motifs that are responsible for the contrasting hydrolysis directions of ssDNA and ssRNA by PfuRecJ. In future research, it is important to obtain the structure of PfuRecJ co-crystallized with ssDNA or ssRNA to interpret its catalytic mechanism.

DATA AVAILABILITY

The atomic coordinates and structure factors have been deposited in the Protein Data Bank (http://www.rcsb.org/pdb) with PDB ID codes 5X4H (Apo wt PfuRecJ), 5X4I (PfuRecJ D83A and Mn^{2+}), 5X4J (PfuRecJ D83A and Zn^{2+}) and 5X4J (PfuRecJ D83A and CMP).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Author contributions: X.P.L., M.J.L., G.S.Y. and J.H.H. designed the study. G.S.Y., M.J.L., X.P.L. and F.Y. performed experiments. X.P.L. wrote the paper. X.P.L., G.S.Y., M.J.L., F.Y. and J.H.H. improved the paper. All authors analyzed the data and approved the final version of the manuscript.

FUNDING

National Natural Science Foundation of China [31371260, 41530967, J1210047]; China Ocean Mineral Resources

R&D Association [DY125-22-04]. Funding for open access charge: National Natural Science Foundation of China.
Conflict of interest statement. None declared.

REFERENCES

1. Yang,W. (2011) Nucleases: diversity of structure, function and mechanism. *Q. Rev. Biophys.*, **44**, 1–93.
2. Rajman,L.A. and Lovett,S.T. (2000) A thermostable single-strand DNase from Methanococcus jannaschii related to the RecJ recombination and repair exonuclease from Escherichia coli. *J. Bacteriol.*, **182**, 607–612.
3. Li,Z., Pan,M., Santangelo,T.J., Chemnitz,W., Yuan,W., Edwards,J.L., Hurwitz,J., Reeve,J.N. and Kelman,Z. (2011) A novel DNA nuclease is stimulated by association with the GINS complex. *Nucleic Acids Res.*, **39**, 6114–6123.
4. Wakamatsu,T., Kim,K., Uemura,Y., Nakagawa,N., Kuramitsu,S. and Masui,R. (2011) Role of RecJ-like protein with 5′-3′ exonuclease activity in oligo(deoxy)nucleotide degradation. *J. Biol. Chem.*, **286**, 2807–2816.
5. Sanchez-Pulido,L. and Ponting,C.P. (2011) Cdc45: the missing RecJ ortholog in eukaryotes? *Bioinformatics*, **27**, 1885–1888.
6. Han,E.S., Cooper,D.L., Persky,N.S., Sutera,V.A.J., Whitaker,R.D., Montello,M.L. and Lovett,S.T. (2006) RecJ exonuclease: substrates, products and interaction with SSB. *Nucleic Acids Res.*, **34**, 1084–1091.
7. Dianov,G., Sedgwick,B., Daly,G., Olsson,M., Lovett,S. and Lindahl,T. (1994) Release of 5′-terminal deoxyribose-phosphate residues from incised abasic sites in DNA by the Escherichia coli RecJ protein. *Nucleic Acids Res.*, **22**, 993–998.
8. Thoms,B., Borchers,I. and Wackernagel,W. (2008) Effects of single-strand DNases ExoI, RecJ, ExoVII, and SbcCD on homologous recombination of recBCD+ strains of Escherichia coli and roles of SbcB15 and XonA2 ExoI mutant enzymes. *J. Bacteriol.*, **190**, 179–192.
9. Burdett,V., Baitinger,C., Viswanathan,M., Lovett,S.T. and Modrich,P. (2001) In vivo requirement for RecJ, ExoVII, ExoI, and ExoX in methyl-directed mismatch repair. *Proc. Natl Acad. Sci. U.S.A.*, **98**, 6765–6770.
10. Dianov,G. and Lindahl,T. (1994) Reconstitution of the DNA base excision-repair pathway. *Curr Biol.*, **4**, 1069–1076.
11. Wakamatsu,T., Kitamura,Y., Kotera,Y., Nakagawa,N., Kuramitsu,S. and Masui,R. (2010) Structure of RecJ exonuclease defines its specificity for single-stranded DNA. *J. Biol. Chem.*, **285**, 9762–9769.
12. Yamagata,A., Kakuta,Y., Masui,R. and Fukuyama,K. (2002) The crystal structure of exonuclease RecJ bound to Mn2+ ion suggests how its characteristic motifs are involved in exonuclease activity. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 5908–5912.
13. Cheng,K., Zhao,Y., Chen,X., Li,T., Wang,L., Xu,H., Tian,B. and Hua,Y. (2015) A Novel C-Terminal Domain of RecJ is Critical for Interaction with HerA in Deinococcus radiodurans. *Front. Microbiol.*, **6**, 1302.
14. Cheng,K., Xu,H., Chen,X., Wang,L., Tian,B., Zhao,Y. and Hua,Y. (2016) Structural basis for DNA 5′-end resection by RecJ. *Elife*, **5**, e14294.
15. Yuan,H., Liu,X.P., Han,Z., Allers,T., Hou,J.L. and Liu,J.H. (2013) RecJ-like protein from Pyrococcus furiosus has 3′-5′ exonuclease activity on RNA: implication of its proofreading capacity on 3′-mismatched RNA primer in DNA replication. *Nucleic Acids Res.*, **41**, 5817–5826.
16. Li,Z., Santangelo,T.J., Cuboňová,L., Reeve,J.N. and Kelman,Z. (2010) Affinity purification of an archaeal DNA replication protein network. *MBio.*, **1**, doi:10.1128/mBio.00221-10.
17. Marinsek,N., Barry,E.R., Makarova,K.S., Dionne,I., Koonin,E.V. and Bell,S.D. (2006) GINS, a central nexus in the archaeal DNA replication fork. *EMBO Rep.*, **7**, 539–545.
18. Fenwick,A.L., Kliszczak,M., Cooper,F., Murray,J., Sanchez-Pulido,L., Twigg,S.R., Goriely,A., McGowan,S.J., Miller,K.A., Taylor,I.B. *et al.* (2016) Mutations in CDC45, encoding an essential component of the pre-initiation complex, cause Meier-Gorlin syndrome and craniosynostosis. *Am. J. Hum. Genet.*, **99**, 125–138.
19. Krastanova,I., Sannino,V., Amenitsch,H., Gileadi,O., Pisani,F.M. and Onesti,S. (2012) Structural and functional insights into the DNA replication factor Cdc45 reveal an evolutionary relationship to the DHH family of phosphoesterases. *J. Biol. Chem.*, **287**, 4121–4128.
20. Szambowska,A., Tessmer,I., Kursula,P., Usskilat,C., Prus,P., Pospiech,H. and Grosse,F. (2014) DNA binding properties of human Cdc45 suggest a function as molecular wedge for DNA unwinding. *Nucleic Acids Res.*, **42**, 2308–2319.
21. Moyer,S.E., Lewis,P.W. and Botchan,M.R. (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.*, **103**, 10236–10241.
22. Aparicio,T., Guillou,E., Coloma,J., Montoya,G. and Méndez,J. (2009) The human GINS complex associates with Cdc45 and MCM and is essential for DNA replication. *Nucleic Acids Res.*, **37**, 2087–2095.
23. Gambus,A., Jones,R.C., Sanchez-Diaz,A., Kanemaki,M., van Deursen,F., EMBDondson,R.D. and Labib,K. (2006) GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat. Cell Biol.*, **8**, 358–366.
24. Pacek,M., Tutter,A.V., Kubota,Y., Takisawa,H. and Walter,J.C. (2006) Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Mol. Cell*, **21**, 581–587.
25. Costa,A., Ilves,I., Tamberg,N., Petojevic,T., Nogales,E., Botchan,M.R. and Berger,J.M. (2011) The structural basis for MCM2-7 helicase activation by GINS and Cdc45. *Nat. Struct. Mol. Biol.*, **18**, 471–477.
26. Pospiech,H., Grosse,F. and Pisani,F.M. (2010) The initiation step of eukaryotic DNA replication. *Subcell. Biochem.*, **50**, 79–104.
27. Simon,A.C., Sannino,V., Costanzo,V. and Pellegrini,L. (2016) Structure of human Cdc45 and implications for CMG helicase function. *Nat. Commun.*, **7**, 11638.
28. Abid,A.F., Renault,L., Gannon,J., Gahlon,H.L., Kotecha,A., Zhou,J.C., Rueda,D. and Costa,A. (2016) Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate. *Nat. Commun.*, **7**, 10708.
29. Oyama,T., Ishino,S., Shirai,T., Yamagami,T., Nagata,M., Ogino,H., Kusunoki,M. and Ishino,Y. (2016) Atomic structure of an archaeal GAN suggests its dual roles as an exonuclease in DNA repair and a CMG component in DNA replication. *Nucleic Acids Res.*, **44**, 9505–9517.
30. Kabsch,W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.
31. Winn,M.D., Ballard,C.C., Cowtan,K.D., Dodson,E.J., Emsley,P., Evans,P.R., Keegan,R.M., Krissinel,E.B. and Leslie,A.G. (2011) McCoy, Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 235–242.
32. Bricogne,G., Vonrhein,C., Flensburg,C., Schiltz,M. and Paciorek,W. (2003) Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D Biol. Crystallogr.*, **59**, 2023–2030.
33. Adams,P.D., Afonine,P.V., Bunkóczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010 PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
34. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
35. Chen,V.B., Arendall,W.B., Headd,J.J., Keedy,D.A., Immormino,R.M., Kapral,G.J., Murray,L.W., Richardson,J.S. and Richardson,D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 12–21.
36. Pellegrini,L. (2016) Structural insights into Cdc45 function: was there a nuclease at the heart of the ancestral replisome? *Biophys Chem.*, **225**, 10–14.
37. Xu,Y., Gristwood,T., Hodgson,B., Trinidad,J.C., Albers,S.V. and Bell,S.D. (2016) Archaeal orthologs of Cdc45 and GINS form a stable complex that stimulates the helicase activity of MCM. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 13390–13395.
38. Burkhart,B.W., Cubonova,L., Heider,M.R., Kelman,Z., Reeve,J.N. and Santangelo,T.J. (2017) The GAN exonuclease, or the flap endonuclease Fen1 and RNase HII are necessary for viability of

Thermococcus kodakarensis. *J. Bacteriol.*, **199**, doi:10.1128/JB.00141-17.
39. Giroux,X. and MacNeill,S.A. (2015) Molecular Genetic Methods to Study DNA Replication Protein Function in Haloferax volcanii, A Model Archaeal Organism. *Methods Mol. Biol.*, **1300**, 187–218.
40. Morimatsu,K. and Kowalczykowski,S.C. (2014) RecQ helicase and RecJ nuclease provide complementary functions to resect DNA for homologous recombination. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E5133–E5142.
41. Wigley,D.B. (2013) Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB. *Nat. Rev. Microbiol.*, **11**, 9–13.
42. Hopkins,B.B. and Paull,T.T. (2008) The P. furiosus mre11/rad50 complex promotes 5′ strand resection at a DNA double-strand break. *Cell*, **135**, 250–260.
43. Petojevic,T., Pesavento,J.J., Costa,A., Liang,J., Wang,Z., Berger,J.M. and Botchan,M.R. (2015) Cdc45 (cell division cycle protein 45) guards the gate of the Eukaryote Replisome helicase stabilizing leading strand engagement. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E249–E258.