The core Cas1 protein of CRISPR-Cas I-B in *Leptospira* shows metal-tunable nuclease activity

Bhuvan Dixit a, Aman Prakash a, Pankaj Kumar b, Prerana Gogoi a, Manish Kumar a, b, *

a Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati, Assam 781039, India
b Division of Livestock and Fisheries Management, ICAR Research Complex for Eastern Region, Patna, Bihar 800014, India

1. Introduction

CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats- CRISPR associated proteins) is an emerging microbial acquired immune defense system against the foreign genetic elements transferred horizontally from viruses and plasmids (Barrangou et al., 2007; Yosef et al., 2012; Sternberg et al., 2010; Makarova et al., 2015). AT-rich regions precede the first repeat (Makarova et al., 2015; Koonin et al., 2017). The architecture of the CRISPR-Cas immune system comprises of an array of palindromic direct repeats interrupted by the variable spacers and a set of associated cas genes (Haft et al., 2005; Pourcel et al., 2005; Garneau et al., 2010; Makarova et al., 2015). AT-rich regions precede the first direct repeat called a leader sequence (Yosef et al., 2012; Sternberg et al., 2016). CRISPR immunity is initiated by acquiring the foreign genetic material into CRISPR arrays (Westra et al., 2012). The two core proteins Cas1 and Cas2 obtains short foreign DNA fragments (protopacers) from the invasive genetic element and integrates them into the CRISPR array as new spacers between the leader end and the start of the first repeat (Makarova et al., 2015; Koonin et al., 2017).

CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats- CRISPR associated proteins) is an emerging microbial acquired immune defense system against the foreign genetic elements transferred horizontally from viruses and plasmids (Barrangou et al., 2007; Yosef et al., 2012; Sternberg et al., 2014). The CRISPR-Cas systems have been classified into six major types (Type I-VI). Each type is further subdivided into various subtypes based on signature cas genes and the locus arrangement (Makarova et al., 2015; Koonin et al., 2017). The architecture of the CRISPR-Cas immune system comprises of an array of palindromic direct repeats interrupted by the variable spacers and a set of associated cas genes (Haft et al., 2005; Pourcel et al., 2005; Garneau et al., 2010; Makarova et al., 2015). AT-rich regions precede the first direct repeat called a leader sequence (Yosef et al., 2012; Sternberg et al., 2016). CRISPR immunity is initiated by acquiring the foreign genetic material into CRISPR arrays (Westra et al., 2012). The two core proteins Cas1 and Cas2 obtains short foreign DNA fragments (protopacers) from the invasive genetic element and integrates them into the CRISPR array as new spacers between the leader end and the start of the first repeat (Makarova et al., 2015; Koonin et al., 2017).

Protopacer-adjacent motifs (PAMs) in the DNA of invasive genetic elements determine the acquisition efficiency and integration specificity during the spacer acquisition process (Bolotin et al., 2005; Deveau et al., 2008; Mojica et al., 2009). The Cas1-Cas2 complex processes the acquired protospacers by cleaving-off the PAM sequence to appropriate sizes before integrating it into the CRISPR array (Yosef et al., 2012; Nuñez et al., 2014; Levy et al., 2015; Nuñez et al., 2015; Wang et al., 2015; Xiao et al., 2017; He et al., 2018; Kieper et al., 2018; Rollie et al., 2018). Among diverse CRISPR-Cas systems, the PAM dependent spacer acquisition is observed only in type-I, -II and -V (Bolotin et al., 2005; Deveau et al., 2008; Mojica et al., 2009; Shah et al., 2013; Heler et al., 2014). Bioinformatics and various enzymatic assays describe Cas1 to function as a nuclease/integrase (Makarova et al., 2006). Cas1 is a metal-dependent nuclease that can cleave a range of different substrates, including single-stranded DNA (ss-DNA), double-stranded DNA (ds-DNA), branched DNA, and plasmid DNA (Wiedenheft et al., 2009; Babu et al., 2011; He et al., 2018). Two types of phosphodiester bonds, 5'- or 3'- of a scissile phosphate, can be cleaved in DNA by the nucleases through nucleophilic substitution. The metal-ion-dependent phosphodiester hydrolysis of DNA involves metal-ion in nucleophile activation, transition state stabilization, and leaving group protonation (Dupureur,
The previous sequence analysis of Cas1 described elsewhere (Makarova et al., 2002) stipulates that four strictly conserved residues (three carboxylates and one histidine) represent the signature motif of the Cas1 family. To date, Cas1 proteins from six different bacteria and one phage have been crystallized (Babu et al., 2011; Wiedenheft et al., 2012; Kim et al., 2013; Ka et al., 2016; Wilkinson et al., 2016; Yang et al., 2020). The crystal structures of Cas1 proteins show it to exist in the homo-dimeric state whose monomer is organised into an N-terminal β-strand domain (NTD) and a C-terminal α-helical domain (CTD) connected by a small linker. However, the catalytic residues responsible for the hydrolysis of nucleic acid are existing within the CTD, whereas the NTD interact with adjacent protomers to promote a stable Cas1 homo-dimer state (Wiedenheft et al., 2009; Kim et al., 2013; Nuñez et al., 2014; Ka et al., 2016; Wilkinson et al., 2016). Nevertheless, from the structural point of view, the Cas1 nuclease activity observed against different substrates has failed to provide significant insights into substrate specificity except that of *E. coli* Cas1-Cas2 (Wang et al., 2015). Therefore, further detailed enzymatic studies of Cas1 proteins from diverse organisms are required to delineate the functional versatility of this nuclease family.

The pathogenic and intermediate strains of *Leptospira* that cause leptospirosis disease in animals and human possess three subtypes (I-B, I-C, and I-E) of the CRISPR-Cas system whereas, type V CRISPR-Cas system has been recently reported in the saprophytic strain of *Leptospira* (*L. biflexa*) (Makarova et al., 2011; Xiao et al., 2019; Arbas et al., 2021). In a recent computational study, a weakly conserved PAM sequence (5′-TAC-3′) has been predicted for CRISPR-Cas I-B of *Leptospira* (Xiao et al., 2019). Bioinformatics and transcriptional analysis have revealed that CRISPR-Cas I-B exists in the *L. interrogans* genome, wherein a functional cas1 gene has previously been disclosed to be transcriptionally active (Dixit et al., 2016). As part of our ongoing investigation of the CRISPR-Cas I-B system in *Leptospira* (Dixit et al., 2021; Prakash and Kumar, 2021), in this study, we explored the nuclease activity of recombinant Cas1 protein (LinCas1) and its mutant variants. In this study, we intend to show that LinCas1 possesses metal-ion dependent DNase activity, which is inhibited upon substitution of metal-ion or its binding residues. The nuclease and immunoassay demonstrate that LinCas1 interacts with LinCas2 and the mixture of LinCas1-Cas2 shows restrained nuclease activity than the pure core Cas proteins.

![Fig. 1. LinCas1 nuclease activity on double-stranded DNA. In all panels, EI represents LinCas1 and M denotes Mn^{2+}. All the reactions were carried out for 1 h at 37 °C unless stated. (A) LinCas1 shows metal-ion dependent DNase activity. LinCas1 activity on plasmid DNA (pTZ57R/T; 2.9 kbp) in the presence or absence of metal-ion. EDTA was used as a metal-ion chelator in the absence of Mn^{2+} ion to rule out the presence of any metal-ion that may have co-purified. (B) Time-dependent DNase activity in LinCas1. Nuclease activity on plasmid DNA (pTZ57R/T) as substrate (1 μg) was performed for the different duration (0.5–3 h) where LinCas1 shows complete cleavage of DNA in 2.5 h. (C) LinCas1 DNase activity in the presence of different divalent-metal illustrates its preference for Mn^{2+}. (D) Nuclease activity in LinCas1 on ds-DNA at different pH. DNase activity in LinCas1 is exhibited at pH 6.0–11.0 and is best observed at pH 9.0–11.0. (E) Effect of different monovalent ions and its varying concentrations on LinCas1 ds-DNase activity. The DNA cleavage reaction shows that 150 mM of KCl or 50 mM of NH_4Cl are optimal for LinCas1 activity on DNA. WT indicates nuclease activity without any monovalent ions. Results presented were confirmed by two independent experiments.](image-url)
2. Results

2.1. Leptospira Cas1 nuclease activity on double-stranded DNA

The recombinant LinCas1 was purified using Ni-affinity chromatography with a yield of 0.3 mg/L (Fig. 1A and B). The yield and concentration of LinCas1 could not be further improvised even after performing multiple attempts of overexpression and purification of LinCas1 in numerous expression vectors. Nuclease assay performed with the given LinCas1 concentration (6 μM) manifested metal-ion (Mn^{2+}) dependent endo-deoxyribonuclease activity against circular ds-DNA (pTZ57R/T vector) where the substrate (1 μg) was completely cleaved in 2.5 h (Fig. 1A and B). The Cas1 has been reported to recognize the proximal PAM sequence in the protospacer bound by the Cas1-Cas2 complex and trims the protospacer at the PAM before its integration in the CRISPR loci (Yosef et al., 2012; Nunez et al., 2014,2015; Levy et al., 2015; Wang et al., 2015; Xiao et al., 2017; He et al., 2018; Kieper et al., 2018; Rollie et al., 2018). Based on this, we investigated if the cleavage of the circular ds-DNA by the LinCas1 is driven by the presence of a possible PAM motif in the pTZ57R/T vector. The vector sequence was derived from the SnapGene tool for screening the PAM motif. In the pTZ57R/T nucleotide sequence, around 60 PAM motifs (5′-TAC-3′) were determined, perhaps that may be recognized by the LinCas1 during cleavage. The determination of the number of PAM motifs (5′-TAC-3′) was based on a lead from a recent study about PAM in L. interrogans (Xiao et al., 2019). For studying metallo-nucleases, substituting metal-ions is a common practice to understand the role of metal-ions in the nuclease activity. On replacing Mn^{2+} with other divalent metals (Mg^{2+}, Ca^{2+}, and Ni^{2+}), we observed that LinCas1 prefers Mn^{2+} followed by Mg^{2+} for DNase activity. In the presence of Ni^{2+} or Ca^{2+}, the DNase activity of LinCas1 gets equally reduced (Fig. 1C). The endonuclease activity in LinCas1 is inhibited by EDTA due to the chelation of the divalent-metal (Fig. 1C). Upon heat inactivation of LinCas1 at 100 °C for 10 min, the DNase activity was observed to be abrogated entirely, and a shift in the mobility of the circular ds-DNA was noticed (Fig. 52).

Next, the effect of pH (pH 3.0–11.0) in LinCas1 nuclease activity on ds-DNA was studied as described previously for LinCas2 and LinCas4 (Dixit et al., 2016,2021). The nuclease activity of LinCas1 was inhibited below pH 6.0. (Fig. 1D). At pH 3.0, no smear of DNA cleavage products was evident, suggesting the abolition of nuclease activity. Moreover, at pH 3-5, due to nucleo-protein interaction, a shift in the migration of a fraction of DNA substrate was exhibited on agarose gel electrophoresis (Fig. 1D). LinCas1 exhibited maximum nuclease activity at a wide pH range (pH 6–11). Thereafter all the nuclease assays were preferred to be done in the range of pH 7-7.5. The effect of salt on endonuclease activity in LinCas1 in the presence of Mn^{2+} was analyzed by using various monovalent-ions (NaCl, KCl, and NH4Cl) at an increasing concentration (Fig. 1E). LinCas1 optimum endonuclease activity on ds-DNA was observed at 150 mM of KCl or 50 mM NH4Cl (Fig. 1E). However, in the presence of NaCl (50–150 mM), LinCas1 did not exhibit a considerable endonuclease activity (Fig. 1E). To understand the diversity of nuclease activity in LinCas1, substrates other than ds-DNA were used in nuclease assays. LinCas1 cleaves viral single-stranded DNA (linear M13mp18 and circular φx174) only in the presence of divalent-metal cofactor (Fig. 2A and B). Interestingly, LinCas1 did not demonstrate any cleavage activity on shorter DNA oligos (28 and 37-mer nucleotides) (Fig. 2C). The inability of the LinCas1 to cleave the short oligos may have been due to the absence of the probable PAM sequence (5′-TAC-3′) of L. interrogans type I-B CRISPR-Cas system or is substrate size-dependent.

2.2. In silico characterization of Cas1 protein

A homology search for LinCas1 revealed its closest homologs to be Cas1 protein from thermophilic organisms, including Thermotoga maritima (sequence identity: 28% and query coverage: 96%), Aquifex aeolicus (27 and 95%) and Archaeoglobus fulgidus (27 and 79%). Low sequence similarity was observed for Cas1 protein from other organisms with solved crystal structures such as Streptococcus pyogenes (23 and 55%) and Escherichia coli (36 and 35%). The least sequence similarity was perceived for Cas1 proteins from Pectobacterium atrosepticum (67 and 6%) and Pseudomonas aeruginosa (67 and 2%). A multiple sequence alignment (MSA) was performed to identify the conserved residues in LinCas1 with the crystallized Cas1 proteins from different organisms. Though Cas1 orthologs showed a more extensive range of sequence similarity, the amino acid residues involved in metal-ion binding was highly conserved in all the Cas1 orthologs (Fig. 3). Notably, the LinCas1 (Lin) also possesses all the conserved metal-ion interacting residues (Glu108, His176, and Glu191) located at the CTD (Fig. 3). Moreover, the residues Asn145 and Asp188 present in the vicinity of the metal-ion binding pocket of LinCas1 are also found to be conserved among Cas1 orthologs (Fig. 3). The chemical niche of the LinCas1 metal-ion binding site is similar to that of the Cas1 orthologs, which ascertainment a familiar role in diverse CRISPR-Cas subfamilies (Wiedenhoff et al., 2009; Yosef et al., 2012; Kim et al., 2013; Wilkinson et al., 2016).

The closest structural homolog of LinCas1 in the Protein Data Bank (PDB) shows Cas1 from T. maritima (TmaCas1) with an RMSD (root mean square deviation) of 6.7% for the TmaCas1 and LinCas1 (Dixit et al., 2016,2021). The nuclease activity of LinCas1 was inhibited below pH 6.0. At pH 3.0, no smear of DNA cleavage products was evident, suggesting the abolition of nuclease activity. Moreover, at pH 3-5, due to nucleo-protein interaction, a shift in the migration of a fraction of DNA substrate was exhibited on agarose gel electrophoresis. The closest structural homolog of LinCas1 in the Protein Data Bank (PDB) shows Cas1 from T. maritima (TmaCas1) with an RMSD (root mean square deviation) of 6.7% for the TmaCas1 and LinCas1 (Dixit et al., 2016,2021).
mean square deviation) of 0.6 Å (Table 2). The tertiary structure of LinCas1 monomer, modelled using TmaCas1 as the template, revealed the presence of two structurally distinct domains composed of $\alpha_1$-$\beta_1$-$\beta_2$ (NTD) and $\alpha_2$-$\alpha_3$-$\alpha_4$-$\alpha_5$-$\eta_1$-$\alpha_6$-$\alpha_7$-$\beta_3$-$\beta_4$-$\alpha_8$-$\beta_5$-$\beta_6$-$\alpha_9$ (CTD) secondary folds connected by a linker (Fig. 4 A). The modelled CTD [amino acid (aa) 32–284] of LinCas1 protomer consists of eight $\alpha$-helices (grey; $\alpha_2$–$\alpha_9$), one 310 helix (grey) and four $\beta$-strands (orange; $\beta_3$–$\beta_6$) (Fig. 4 A). Although the overall structure of the CTD in LinCas1 from various organisms is similar, the number of $\alpha$-helices and $\beta$-strands slightly varies. For instance, the CTD (aa 95–305; $\alpha_2$–$\alpha_8$) of EcoCas1 (Babu et al., 2011) and PaeCas1 (aa 113–324; $\alpha_3$–$\alpha_{10}$) (Wiedenheft et al., 2009) is an all-helical structure while in the case of LinCas1 four additional $\beta$-strands (orange) are present (Fig. 4 A and B). Such an additional four $\beta$-strands in CTD have also been reported in AfuCas1 (Kim et al., 2013) and AaeCas1 (PDB ID: 2YZS). However, these structural variations do not significantly affect the enzymatic activity of Cas1 as they are located far apart from the putative nucleolytic core (Kim et al., 2013). The metal-ion binding pocket in LinCas1 is located between $\alpha_5$ and $\alpha_7$ helices of the CTD (Fig. 4 A) and agrees with the reported structure of AfuCas1 (Kim et al., 2013).

Incredibly, the NTD (cyan and blue) of LinCas1 is much shorter than that of other Cas1 orthologs (Fig. 4 A and B) (Wiedenheft et al., 2009; Babu et al., 2011; Kim et al., 2013). The NTD of LinCas1 comprises of two $\beta$-strands (blue) and one $\alpha$-helix (cyan), whereas, EcoCas1 possesses one $\alpha$-helix (pink), two $\beta_3$ helices and eight $\beta$-strands (magenta) (Fig. 4 A and B) (Babu et al., 2011). The NTD of PaeCas1 is composed of ten $\beta$-strands and two $\alpha$-helices whereas that of AfuCas1 contains one $\alpha$-helix and eight short $\beta$-strands (Wiedenheft et al., 2009; Kim et al., 2013). Variations in LinCas1 secondary structure prompted us to model the dimeric form of LinCas1 as NTD of Cas1 has been described to play a significant role in the formation of homo-dimer in an asymmetric fashion through hydrogen bond and salt bridges (Wiedenheft et al., 2009). In general, the homo-dimer of the Cas1 protein from all organisms attains a butterfly-like shape (Fig. 4 C and D). Modelling of homo-dimer of LinCas1 using the available crystal structures shows a similar shape.

Also, a selective Cas1 evolutionary and comparative analysis with LinCas1 binding regions (metal-ion) was determined. The molecular phylogeny of LinCas1, along with other Cas1 orthologs, reveals LinCas1 closeness to thermophilic organisms (Fig. S3A). It is observed that at the secondary structure level, due to extra $\beta$-strands (CTD), LinCas1 is relatively closer to AfuCas1 than EcoCas1 and PaeCas1. Also, the
recently published molecular phylogeny of another core protein, Lin-Cas2 of CRISPR-Cas in *Leptospira*, shows its lineage closer to Cas2 of *S. solfataricus* than *B. halodurans*, which is another thermophilic archaeon (Dixit et al., 2016). Under the MSA of Cas1 orthologs (Fig. 3), the metal-ion coordinating residues typically found among nucleases are conserved at the tertiary structure of the modelled LinCas1. More precisely, the LinCas1 residues (Glu108, His176 and Glu191) and PaeCas1 residues (Glu190, His254 and Asp268) coordinate the divalent-metal (most probably Mn$^{2+}$) ion. Comparative analysis of metal-ion binding pocket at the tertiary level of LinCas1 (cyan) with PaeCas1 (pale yellow) shows a preference for Mn$^{2+}$ than Mg$^{2+}$ (Fig. S3B).

2.3. Mutation of residues involved in coordinating metal-ion in LinCas1 compromises its activity

Based on multiple sequence alignment of Cas1 proteins, site-directed mutagenesis of the predicted metal-ion binding residues (Glu108, His176, and Glu191) of LinCas1 was performed. The mutant LinCas1 variants (LinCas1E108A, LinCas1H176A, and LinCas1E191A) were overexpressed and purified with a similar yield (Fig. S1C and D). The nuclease activity of mutants LinCas1 (E108A), (H176A), and (E191A) was on expected lines, as there occurs dramatic loss in the endonuclease activity on circular ss-DNA (Fig. 5). This leads us to conclude that Lin-Cas1 possesses metal-ion dependent DNase activity as previously reported in other Cas1 orthologs (Wiedenheft et al., 2009; Babu et al., 2011; He et al., 2018).

2.4. In vitro characterization of Cas1-Cas2 protein interaction

Nuclease assays of core LinCas proteins (LinCas1 and LinCas2) in this study and other reported studies by Dixit et al. (2016), reveals that these core LinCas proteins possess metal-dependent DNase activity. In *E. coli* (CRISPR type I-E) and *E. faecalis* (CRISPR type II-A), the Cas1-Cas2 complex has been reported to form an asymmetric heterohexameric complex consisting of two Cas1 dimers that sandwich one Cas2 dimer (Nunez et al., 2015; Xiao et al., 2017). It was thus interesting to address whether the nuclease activity is altered when core LinCas proteins are mixed. In addition, the NTD of LinCas1 responsible for interacting with LinCas2 is shorter than its Cas1 orthologs. Thus, nuclease assay with a mixture of LinCas proteins may address the impact of short NTD of LinCas1 over LinCas2. Based on this, a complex of LinCas1-Cas2 was generated under in vitro condition by incubating pure LinCas1 and
LinCas2 in a 2:1 molar ratio, as suggested elsewhere (Nutiez et al., 2015). The activity of the LinCas1-Cas2 mixture on the DNA substrate (linearized ds-DNA pTZ57R/T) shows a reduction in DNA cleavage and a shift in the mobility of DNA during agarose gel electrophoresis (Fig. 6A). However, the pure LinCas1 at an equivalent molar concentration shows intense nuclease activity on ds-DNA (Fig. 6A).

Due to the constraints in purifying LinCas1 at higher concentrations (Fig. S1), a lower amount of LinCas2 was incubated (3 µM) with LinCas1 (6 µM) to generate the mixture (2:1) for nuclease activity on ds-DNA (Fig. 6A). Nuclease activity was not detected at a lower concentration of pure LinCas1-Cas2 (3 µM); however, LinCas2 at higher concentration (20 µM) exhibits nuclease activity (Fig. 6A), as reported previously (Dixit et al., 2016).

Next, in the LinCas1-Cas2 mixture, we addressed whether mutating the residues (involved in metal-ion binding) of LinCas1 will affect its affinity for DNA. Therefore, LinCas2 was mixed with one of the mutants LinCas1(E108A). The nuclease assay revealed a loss of the cleavage activity in LinCas1(E108A) and LinCas2 mixture on circular ds-DNA similar to that of LinCas1-Cas2 mixture (Fig. 6B). However, LinCas1 (E108A) shows a change in the conformation of circular ds-DNA from supercoiled to relaxed form, possibly due to a moderate nick (Fig. 6B). Nevertheless, the alteration in the activity of the LinCas1 (E108A)-Cas2 mixture for ds-DNA was in agreement with LinCas1-Cas2 mixture, where LinCas2 promotes the binding affinity of LinCas1 (E108A) for ds-DNA (Fig. 6C). Such findings agree with the recorded information about the EcoCas2 dimer. In E. coli, the EcoCas2 dimer acts as an adaptor protein, bringing two EcoCas1 dimers together to stabilize and scale the length of the protospacer DNA (Wang et al., 2015). The Cas1-Cas2 complex selects the protospacer based on the flanking protospacer adjacent motif (PAM) sequence and processes it into appropriate sizes while also cleaving off the PAM sequence, and integrates them between the first CRISPR repeat and leader sequence upstream of the CRISPR array (Barrangou et al., 2007; Yosef et al., 2012; Wei et al., 2015). Recently, through a computational approach, a weakly conserved PAM sequence (5′-TAC-3′) has been identified (Xiao et al., 2019). In the LinCas1 and LinCas2 interaction study, the substrate (ds-DNA and linearized pTZ57R/T) used has about 60 such PAM motifs. These PAM motifs (5′-TAC-3′) may direct the LinCas1-Cas2 mixture to recognize protospacers in the substrate pTZ57R/T for the adaptation process. Nevertheless, further in vitro and in vivo studies are required to confirm this hypothesis and understand the adaptation process of the CRISPR-Cas system in Leptospira.

Additionally, the interaction between the core LinCas proteins (LinCas1 and LinCas2) was addressed by performing an enzyme-linked immunosorbent assay (ELISA) using antibodies generated against LinCas2 (anti-LinCas2). Microtitr plate was coated with LinCas1 (100 ng) or bovine serum albumin and was overlaid with LinCas2 at an increasing concentration (5 ng–12.8 µg). At 450 nm of absorbance, specific interaction was demonstrated between LinCas2 and LinCas1 that could achieve a saturation point (Fig. 6D). The ELISA data supports our previous finding that LinCas1 and LinCas2 mixture interact to form a functionally regulated Cas1-Cas2 complex. The dissociation constant (Kd) of the LinCas2 to LinCas1 interaction through ELISA was 12.86 ± 1.73 nM with a stoichiometry of 0.88. Moreover, in support of ELISA, an immunoblot was performed to demonstrate recombinant LinCas2 existence in monomer and homo-dimer form on SDS-PAGE (Fig. 6E). Monomers and dimer forms of native LinCas2 could be detected through immunoblot in the lysate of L. interrogans serovar Copenhageni (Fig. 6F). Faint bands in immunoblot obtained from the lysate (5 × 10^6 cells in 50 ml) of spirochete suggest that the expression level of cas2 in Leptospira is minimal (Fig. 6F).

The activity of LinCas1-Cas2 mixture was also investigated on the ss-DNA oligonucleotide (37 mer) sequence with no PAM motif (5′-TAC-3′). The LinCas1-Cas2 did not show oligonucleotide (37 mer) cleavage on denaturing PAGE (Fig. S4A). However, on native PAGE, a shift in the mobility of a small proportion of oligonucleotides was detected when incubated with pure LinCas1 or LinCas1-Cas2 mixture (Fig. S4B). The shift in the mobility of the oligonucleotide in the presence of pure LinCas1 or LinCas1-Cas2 mixture appears to be similar, which suggests that LinCas1 interacts with the ss-DNA oligo (Fig. S4B).

This study witnessed an alteration in LinCas1 nuclease activity similar to E. coli and E. faecalis Cas1-Cas2 (Nutiez et al., 2015; Xiao et al., 2017). In addition to this, several other attempts were made to understand the association between LinCas1 and LinCas2. Owing to the high isoelectric point (pI >9.5) of LinCas1 and LinCas2, the protein complexes on native PAGE carried a strong positive charge and thus deterred its migration on the gel. Reversing the direction of the electrode assembly in the native PAGE also failed to resolve the complex. The size exclusion chromatography could not be accomplished to study the interaction between LinCas1 and LinCas2 due to the low yield and concentration of purified LinCas1.

3. Discussion

Biochemical analysis of the LinCas1 reveals that it possesses endonuclease activity in the presence of specific divalent-metal, particularly Mn^{2+}, which stands in agreement with the reported preference of divalent metal-ion for ds-DNAse activity in the Cas1 orthologs of Pseudomonas aeruginosa (PaeCas1) and Riemerella anatipestifer (RanCas1) (Wiedenheft et al., 2009; He et al., 2018). In contrast, the DNAS enzyme activity in A. fulgidus of Archaeoglobus fulgidus was observed in the presence of Ca^{2+} and Mg^{2+} metal-ion (Kim et al., 2013). Thus, the specificity of divalent-metal for nuclease activity varies among various characterized Cas1 proteins, which probably depends on the negatively charged residues interacting with the divalent-metal-ion. The LinCas1 (Glu108 and Glu191) and PaeCas1 (Glu190 and Asp268) attract positively charged metal-ions such as Mn^{2+} and Mg^{2+}. However, the presence of histidine residue in LinCas1 (His176) and PaeCas1 (His254) along with negatively charged residues shifts its preference towards Mn^{2+}, as discussed elsewhere (Khristaleva, 2014). Furthermore, the preference for mono- and divalent ions for LinCas1 nuclease activity also agrees with the earlier reported activity in PaeCas1 (Wiedenheft et al., 2009). The nuclease activity of LinCas1 on viral DNA substrates is consistent with the nuclease activity reported in PaeCas1 (Wiedenheft et al., 2009). However, LinCas1 inability to cleave short oligos is contentious to the activity reported in EcoCas1, where it cleaves short ss-DNAs as well.
The cleavage of ss-DNA is essential in the processing of the protospacer bound by the Cas1-Cas2 complex prior to its integration in the CRISPR loci (Yosef et al., 2012; Nuñez et al., 2014, 2015; Levy et al., 2015; Wang et al., 2015; Xiao et al., 2017; He et al., 2018; Kieper et al., 2018; Rollie et al., 2018). The LinCas1 disability to act on short ss-DNAs, perhaps can be compensated by the activity of the LinCas4 in L. interrogans, which has been recently reported to cleave short oligos under in vitro condition (Dixit et al., 2021). In numerous prokaryotes harboring different CRISPR-Cas subtypes, Cas4 orthologs has been established to be essential for spacer acquisition with high fidelity in association with Cas1 and Cas2 including Bacillus halodurans (type I-C) (Lee et al., 2018), Sulfolobus islandicus (type I-A) (Liu et al., 2017), Geobacter sulfurreducens (type I-U) (Almendros et al., 2019), Synechocystis sp. 6803 (Kieper et al., 2018), and Haloarcula hispanica (type I-B) (Li et al., 2014).

The substitution of Glu108, His176, and Glu191 residues with Ala in LinCas1 showed a drastic reduction in its DNase activity, similar to the loss of the DNase activity reported in PaeCas1 (D268A), EcoCas1 (H208A and D221A), and RanCas1 (E149A, H206A, and E221A), respectively (Wiedenheft et al., 2009; Babu et al., 2011; He et al., 2018). The metal-ion interacting residues (Glu108, His176, and Glu191) in the LinCas1 correspond to the similar residues in different CRISPR-Cas subtypes, including I-A (S. solfataricus), I-F (P. atrosepticum), and I-E (E. coli) (Yosef et al., 2012; Wilkinson et al., 2016; Rollie et al., 2018).
The metal-ion interacting residues in these Cas1 orthologs are illustrated to be essential for integrase activity. Mutagenesis studies suggest that the interacting metal residues of the LinCas1 may conceivably be crucial for the CRISPR acquisition in the L. interrogans.

The existence of a shorter NTD in LinCas1 compared to its orthologs did not hinder its interaction with LinCas2, which was confirmed through ELISA demonstrating a high affinity of LinCas2 to LinCas1 interaction ($K_d = 12.86 \pm 1.73$ nM). In contrast, the Cas1 affinity of E. coli (Núñez et al., 2014) and S. pyogenes (Ka et al., 2018) to its cognate Cas2 by isothermal titration calorimetry was weaker ($K_d = 290$ and 240 nM, respectively). Also, the nuclease assay of the in vitro generated LinCas1-Cas2 complex demonstrates alteration in the nuclease activity of the pure-LinCas1. Notably, a previous study reported that the protospacer binding to the EcoCas1-Cas2 complex triggers large structural rearrangements within the core Cas protein (Wang et al., 2015). Thus we speculate that LinCas1, upon forming a complex with LinCas2, undergoes structural rearrangements leading to a change in the nuclease activity.

Interestingly, bacterial cells (E. coli) grew normally after LinCas1 overexpression, which suggests that the genomic DNA of E. coli was not much influenced by the LinCas1 nuclease activity during the 20 min of generation time. Based on the in vitro nuclease assays of the LinCas1, it is possible that the host genomic DNA remains intact due to the lack of required optimum concentrations of divalent and monovalent ions within the E. coli. Similarly, on overexpressing LinCas1 and LinCas2 together in E. coli, there occurred no deleterious effect on the host genomic DNA.

Plasmids such as pMaOri (around 5 kbp) and pMaOri.dCas9 (around 10 kbp) in L. interrogans has been reported to be maintained for several passages (Fernandes et al., 2021). Based on the nuclease activity of LinCas1 on circular ds-DNA (pTZ5R/T plasmid), which possesses numerous (n = 60) PAM motifs (5′-TAC-3′), we believe that under in vitro conditions, LinCas1 may also be able to cleave pMaOri and pMaOri.dCas9 by recognizing its PAM motifs (5′-TAC-3′). On the contrary, under in vivo conditions, the expression level of the native Cas1 may be minimal similar to that of LinCas2. Moreover, the influence of different divalent and monovalent ions within the Lepotospira may modulate its nuclease activity on plasmids. The other plausible reason for the stability of the exotic plasmids (pMaOri and pMaOri.dCas9) in Lepotospira spp. could be due to the occurrence of only primed adaptation as has been reported in type I-B system of Haloarcula hispanica (Li et al., 2014). In primed adaptation, aside from the adaptation proteins (Cas1, Cas2, and Cas4), a pre-existing spacer is a prerequisite to initiate CRISPR adaptation (Li et al., 2014).

In L. interrogans serovar Copenhageni, the adaptation modules of I-B (cas1, cas2, and cas4) involved in adaptation have been reported to be actively transcribed under in vitro growth conditions (Dixit et al., 2016). In this study, native LinCas2 protein has been shown to be expressed in L. interrogans which is contrary to the previous finding where quantitative proteome analysis of L. interrogans did not reveal cas2 gene expression (Malmström et al., 2009). This discrepancy can be reasonably explained due to the low differential expression level of the cas2 gene.

4. Conclusion

The results of this study demonstrate that LinCas1 is biochemically active like LinCas2 and LinCas4 of the CRISPR-Cas I-B in Lepotospira. The data presented here ascertain that the LinCas1 has versatile nuclease activity and can interact with LinCas2 despite its short NTD. The diverse DNA substrates that can be cleaved by LinCas1 enable us to determine the substrate types that can be used as a possible source of protospacers for CRISPR adaptation in Lepotospira. The DNase activity for LinCas1 is consistent with a role for this protein in the recognition/cleavage of foreign DNA and integration as a new spacer into the CRISPR array (Núñez et al., 2015; Rollie et al., 2018). Further detailed analysis of the role of LinCas1 in the CRISPR adaptation mechanism within Lepotospira requires using a natural experimental model as established in E. coli CRISPR spacer integration.

5. Materials and methods

5.1. Bacterial strains

The bacterial strain of L. interrogans serovar Copenhageni Fiocruz L1-130 was procured from the Indian Council of Medical Research (ICMR), Regional Medical Research Centre, Port Blair, Andaman and Nicobar Island, India. High-passaged Lepotospira were cultured in EMJH (Ellinghausen-McCullough-Johnson-Harris, Difco) medium at 29 °C and subcultured at 7-days intervals to isolate genomic DNA. Bacterial strains of E. coli DH5α and BL21 (DE3) were grown at 37 °C in LB (Luria Bertani) medium or agar for cloning, transformation, and expression.

5.2. Recombinant DNA techniques and nucleic acid isolation

Genomic DNA of L. interrogans serovar Copenhageni was isolated from a 7-day old 10 ml culture using a bacterial genomic DNA purification kit (Qiagen), as per the manufacturer’s instruction. Plasmids were isolated from 5 ml of an overnight culture of E. coli using a plasmid purification kit from New England Biolabs (NEB). Standard procedures were used for the generation of the recombinant plasmid. The QIAquick gel extraction kit (Qiagen) was used to isolate DNA fragments from agarose gels. Substrate ss-DNA (M13mp18, 4x174 genome) and all enzymes used for genetic engineering were sourced through NEB or Fermentas. The confirmed clones were outsourced to Eurofins Genomics India Pvt. Ltd, Bengaluru for DNA sequencing before purification and characterization of proteins.

5.3. Nuclease activity assays

Nuclease activity in LinCas1 was investigated on various DNA substrates, as described elsewhere (Beloglazova et al., 2015). The different nucleic acid substrates (1.0 µg each) used for evaluating the nuclease activity of LinCas1 were circular double-stranded (ds) DNA plasmid (pTZ5R/T), linear ss-DNA (M13mp18), circular ss-DNA (4x174 genome), and short oligos of 28 & 37 mer sequence with no PAM (5′-TAC-3′) motif (Xcelris Genomic India). These substrates (1.0 µg) were incubated with the purified rLinCas1 (active) or its heat-inactivated (100 °C for 10 min) form (6 µM each) in a 25 µl reaction buffer (25 mM HEPES pH 7.0, 100 mM KCl, and 2.5 mM MnSO4 at 37 °C for 1 h unless stated. To chelate the metals in the nuclease reaction EDTA (10 mM) was used at the specified reaction condition. Preference for divalent-metal for nuclease activity was determined by substituting Mn²⁺ with various other metal-ion like Mg²⁺, Ca²⁺, and Ni²⁺ ions. The optimal pH was determined by substituting the buffer with either 25 mM sodium citrate (pH 3.0 to 5.0), MES (pH 6.0), HEPES (pH 7.0), Tris-HCl (pH 8.0), or CAPS (pH 9.0 to 11.0). The effect of monovalent-ions on LinCas1 nuclease activity was studied using NaCl, KCl, and NH₄Cl, at varying concentrations (50–150 mM). The time-bound nuclease assay was performed to determine the duration required to cleave the substrate completely. All the above resulting reaction products were separated on 0.8% (w/v) agarose gel electrophoresis except the short oligo (ss-DNA) reaction products which were separated on an 18% (w/v) 8 M denaturing Urea PAGE, stained with ethidium bromide and visualized in gel documentation instrument (Biorad, Hercules, CA, USA).

5.4. Site-directed mutagenesis

The potential metal-ion binding amino acids in the LinCas1 were mutated using the Q5 site-directed mutagenesis kit (NEB) and pET28a-LinCas1 plasmid construct. Primers were designed (Table 1) using the NEBaseChanger tool program to mutate residues Glu108/191 (codon GAA) to Ala (GCA) in LinCas1. Similarly, His176 (CAC) was substituted.
with Ala (GCC). All the derived mutants were confirmed by sequencing before purification of the proteins using Ni-NTA chromatography.

5.5. Biochemical and immunoassay for Cas1-Cas2 interaction

Generation of LinCas1-Cas2 complex under in vitro condition was done by incubating LinCas1 (10 µM) and LinCas2 (5 µM) at 4 °C in 2:1 molar ratio for 1 h. A nuclease reaction was set up at 37 °C for 1 h using the LinCas1-Cas2 complex (6:3 µM) and 1 µg of linear ds-DNA (pTZ57R/-T) or ss-DNA oligonucleotide (37 mer) as the substrate in a nuclease buffer (25 mM Tris-Cl, 100 mM KCl and 10 mM MgCl2). A similar nuclease reaction was set up containing an equal molar concentration of pure LinCas1 (6 µM) or LinCas2 (3 µM) independently along with linear ds-DNA. The resulting reaction products were resolved on 0.8% agarose gel, stained with ethidium bromide, and visualized in a gel documentation instrument.

Antibodies against recombinant LinCas2 proteins were generated in rabbits by outsourcing the purified proteins to Abgenex, Bhubaneswar, India. Microtiter plates were coated with 50 µl of recombinant LinCas1 proteins or bovine serum albumin (BSA) as control (100 ng/well) overnight at 4 °C. Thereafter, the unbound LinCas1 protein was discarded from the wells, and each well was blocked with 3% bovine serum albumin (BSA) solution for 2 h at 37 °C and then was overlaid by LinCas2 at an increasing concentration ranging from 5 ng to 12.8 µg per well for 2 h at room temperature. After three washes of each well with 0.05% phosphate buffer saline-Tween 20 (PBS-T) buffer, anti-LinCas2 at a 1:1000 dilutions was used to find the interaction between LinCas2 and LinCas1 and detected using HRP-conjugated anti-rabbit secondary antibodies (1:5000) with TMB (Trimethylbenzidine, Thermo scientific) as substrate. The absorbance was measured at 450 nm after terminating the reactions with 1 M H2SO4 as per the manufacturer’s instruction. The results are indicative of the average of two independent experiments. The dissociation constant and stoichiometry of LinCas2 to LinCas1 interaction was calculated by nonlinear curve fitting using the Hill equation.

5.6. Multiple sequence alignment of Cas1 orthologs and structure predictions

In a search for the conserved metal-ion coordinating residues, a multiple sequence alignment (MSA) of Cas1 protein sequences from L. interrogans, Streptococcus pyogenes, Archeoglobus fulgidus, Aquifex aeolicus, Thermotoga maritima, Pseudomonas aeruginosa and Pectobacterium atrosepticum was performed using the program Clustal Omega with a default set of parameters (Sievers and Higgins, 2014). The aligned sequences were further rendered using the online web server ESPript (Easy Sequencing in PostScript) (Gouet et al., 2003). The amino acid sequences of all the Cas1 proteins from various organisms were downloaded from the UniProtKB database (https://www.uniprot.org/). Structural homologs of LinCas1 were obtained using the Dali server (Holm and Rosenstrom, 2010).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

There are no conflicts to declare.

Table 1

| Primer sequences used in this work. |
|-------------------------------------|
| 5' to 3' Sequence | Purpose |
| CTAGTCAGCTTGTATGTTTAATTTTTATTGTGAAGAG | Forward primer LinCas1 ( LIC10942 ) NheI |
| CGCTGAGTTCTCTCAACCTGCCGGTGTCA | Reverse primer LinCas1 ( LIC10942 ) XhoI |
| ATTAGGGATCGcAGTTGGCG | Forward primer for the LinCas1(E108A) mutation |
| AATTTGACCTGGGAATCCGG | Reverse primer for the LinCas1(E108A) mutation |
| CGGTATTTTAFeGCACCTCAAGATCTTTC | Forward primer for the LinCas1(H176A) mutation |
| AAACCTTTGCTGAGCCACAC | Reverse primer for the LinCas1(E191A) mutation |
| AGATTTAATGcGeACTTTCGCTGTAAG | 28 mer oligo |
| AAAAAGAGAGGGTGGGG | 37 mer oligo |

*Higher Z-scores correspond to structures which agree more closely in architectural detail.

Acknowledgments

The authors would like to acknowledge ICMR, Port Blair India for providing the Leptospira strains and Dr. Shankar Prasad Kanaujia for giving feedback while writing this manuscript. The authors are also grateful to the members of the laboratory for their indirect help while pursuing this work.

Credit author statement

MK conceived and supervised the study; MK designed experiments; BD and AP performed experiments; PG performed modelling and structural studies; MK, BD, PK and AP analysed data; MK, BD, PK and PG wrote the manuscript.

Funding

The Department of Biotechnology, Government of India, Ministry of Science and Technology bearing project number BT/PR683/GET/119/361/2020 financially supported the present work.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.crmicr.2021.100059.

References

Almendros, C., Nobrega, F.L., McKenzie, R.E., 2019. Brouns SJJ. Cas4–Cas1 fusions drive efficient PAM selection and control CRISPR adaptation. Nucleic Acids Res. 47 (10), 5223–5230.
Arbas, S.M., Narayanasamy, S., Herold, M., Lebrun, L.A., Hoogman, M.R., Li, S., Lam, T. J., Kunath, B.J., Hicks, N.D., Liu, C.M., 2021. Roles of bacteriophages, plasmid and CRISPR immunity in microbial community dynamics revealed using time-series integrated metatransomics. Nat. Microbiol. 6 (1), 123–135.

Babu, M., Beloglazova, N., Flick, R., Graham, C., Skarina, T., Nocek, B., Gagarinova, A., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Fernandes, L.G.V., Hornsby, R., ALTOd, N., Nally, J., 2021. Genetic manipulation of CRISPR-Cas systems in bacterial and archeal genomes. Mol. Microbiol. 79 (2), 484–502.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., Horvath, P., 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science 318 (5851), 1709–1712

Beloglazova, N., Lemak, S., Yakunin, A.F., 2015. Analysis of nuclease activity of Cas1 protein against complex DNA substrates. CRISPR 1311, 251–264.

Bolotin, A., Quinquis, B., Sorokin, A., Ehrlich, S.D., 2005. Clustered regularly interspaced short palindromic repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology 151 (8), 2551–2561.

Deveau, H., Barrangou, R., Garneau, J.E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., Moineau, S., 2008. Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J. Bacteriol. 190 (4), 1390–1400.

Dixit, B., Anand, V., Hussain, M.S., Kumar, M., 2021. The CRISPR-associated Cas4 protein from Leptospira interrogans demonstrates versatile nuclease activity.Curr. Res. Microbio. Sci. 2, 100040.

Dixit, B., Ghosh, K.K., Fernandes, G., Kumar, P., Gogoi, P., Kumar, M., 2016. Dual nuclease activity of a Cas2 protein in CRISPR-Cas subtype I-B of Leptospira interrogans. FEBS Lett. 590 (7), 1002–1016.

Dupureur, CM., 2008. Roles of metal ions in nucleases. Curr. Opin. Chem. Biol. 12 (2), 250–255.

Fernandes, L.G.V., Hornsby, R., Altitud, N., Nally, J., 2021. Genetic manipulation of pathogenic Lepthea: CRISPR interference (CRISPRi)-mediated gene silencing and rapid mutant recovery at 37°C. Sci. Rep. 11 (1), 1–12.

Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A.H., Moineau, S., 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468 (7320), 67.

Gouet, P., Robert, X., Courcelle, E., 2003. ESpIPT/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res. 31 (13), 3320–3323.

Haft, D.H., Selengut, J., Wong, C., Egan, D., 2005. A guild of 45 CRISPR-Cas systems in prokaryotic genomes. PLoS Comput. Biol. 1 (6), e60.

He, Y., Wang, M., Liu, M., Huang, L., Liu, C., Zhang, X., Yi, H., Cheng, A., Zhu, D., Wang, J., Li, J., Zhao, H., Sheng, G., Wang, M., Yin, M., Wang, Y., 2015. Structural and functional analysis of Cas4-dependent prespacer processing and Cas10-mediated programming of Cas1 in vivo. Mol. Cell 61 (6), 797–809.

Holm LRosenstrom P, 2010. Dali server: conservation mapping in 3D. Nucleic Acids Res. 38 (7), W545–W549.

Heler, R., Marraffini, L.A., Bikard, D., 2014. Adapting to new threats: the generation of memory by CRISPR-Cas immunes systems. Mol. Microbiol. 93 (1), 1–9.

Holm LBrömmelström P. 2010. Dal server: conservation mapping in 3D. Nucleic Acids Res. 38, W545–W549. Web Server issue.

Yang, J., Li, J., Wang, J., Sheng, G., Wang, M., Zhao, H., Yang, Y., Wang, Y., 2020. Crystal structure of Cas1 in complex with branched DNA. Sci. China Life Sci. 63 (4), 516–528.

Ko, D., Jung, D.M., Han, B.W., Bae, E., 2018. Molecular organization of the type II-A CRISPR adaptation module and its interaction with Cas9 via Csn2. Nucleic Acids Res. 46 (18), 9805–9815.

Ko, D., Lee, H., Jung, Y.D., Kim, K., Seok, C., Suh, N., Bae, E., 2016. Crystal structure of Streptococcus pyogenes Cas1 and its interaction with Cas2 in the type II-C CRISPR-Cas system. Structure 24 (1), 70–79.

Khristaleva, T.A., 2012. Secondary structure preferences of mur (+) binding sites in bacterial proteins. Adv. Bioinform. 2014, 501641.

Kieper, S.N., Almenndorff, C., Behler, J., McKenzie, R.E., Nobrega, F.L., Haagena, A.C., Vink, J.N., Hess, W.R., Brouns, J.S., 2018. Cas4 facilitates PAM-compatible spacer selection during CRISPR adaption. Cell Rep. 22 (13), 3377–3384.

Kim, T.Y., Shin, M., Hynnh Thi Yen, L., Kim, J.S., 2013. Crystal structure of Cas1 from Arthromonas lutea and its potential involvement in immunity. Biochem. Biophys. Res. Commun. 441 (4), 720–725.

Koonin, E.V., Makarova, K.S., 2017. Diversity, classification and evolution of CRISPR-Cas systems. Curr. Opin. Microbiol. 37, 67–78.

Kumar, M., Kaur, S., Kaur, J., Xiong, X., Bonis, L., Anderson, J.F., Pal, U., 2011. Borrelia burgdorferi BBA25 is a potential target for transmission blocking Lyme disease vaccine. Vaccine 29 (48), 9012–9019.

Lee, H., Zhou, Y., Taylor, D.W., Sashital, D.G., 2018. Cas4-dependent prespacer processing enhances high-fidelity programming of CRISPR arrays. Mol. Cell 70 (1), 48–59 e5.

Lev, A., Goren, M.G., Yosef, I., Auster, O., Manor, M., Amitai, G., Edgar, R., Qimron, U., Sorek, R., 2015. CRISPR adaptation biases explain prefernce for acquisition of foreign DNA. Nature 520 (750), 509–510.

Li, M., Wang, R., Zhao, D., Xiang, H., 2014. Adaptation of the Haloarcula hispanica CRISPR-Cas system to a purified virus strictly requires a priming process. Nucleic Acids Res. 42 (4), 2483–2492.

Liu, T., Liu, Z., Ye, Q., Pan, S., Wang, X., Li, Y., Peng, W., Liang, Y., She, Q., Peng, N., 2017. Coupling transcriptional activation of CRISPR-Cas system and DNA repair genes by Cas3a in sulfolobus islandicus. Nucleic Acids Res. 45 (15), 9878–9892.

Makarova, K.S., Aravin, L., Grishin, N.V., Bogozin, L.B., Koonin, E.V., 2002. A dual function of the CRISPR-Cas system in bacterial anti-virus and DNA repair. Mol. Microbiol. 79 (2), 484–502.

Makarova, K.S., Beloglazova, N., Lemak, S., Flick, R., Yakunin, A.F., 2015. Analysis of nuclease activity of Cas1 protein against complex DNA substrates. CRISPR 1311, 251–264.

Boeva, V., Beloglazova, N., Beloglazova, N., Lemak, S., Yakunin, A.F., 2015. Analysis of nuclease activity of Cas1 protein against complex DNA substrates. CRISPR 1311, 251–264.

Cowan, A.J., Koonin, E.V., Yakunin, A.F., 2011. Dual function of the CRISPR-Cas system in bacterial anti-virus and DNA repair. Mol. Microbiol. 79 (2), 484–502.

Roucas, J., Beloglazova, N., Beloglazova, N., Lemak, S., Yakunin, A.F., 2011. Dual function of the CRISPR-Cas system in bacterial anti-virus and DNA repair. Mol. Microbiol. 79 (2), 484–502.

Romero, D.A., Horvath, P., Moineau, S., 2008. Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J. Bacteriol. 190 (4), 1390–1400.

Makarova, K.S., Aravin, L., Grishin, N.V., Bogozin, L.B., Koonin, E.V., 2002. A dual function of the CRISPR-Cas system in bacterial anti-virus and DNA repair. Mol. Microbiol. 79 (2), 484–502.