PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

STRUCTURE NOTE

Crystal structure of a Cas6 paralogous protein from *Pyrococcus furiosus*

Hye-Mi Park,[1] Minsang Shin,[2] Jiali Sun,[1] Gwang Sik Kim,[2] Young Chul Lee,[2] Jeong-Hoh Park,[1] Bo Yeon Kim,[3] and Jeong-Sun Kim[1]*

[1] Department of Chemistry and Institute of Basic Sciences, Chonnam National University, 300, Yongbong-dong, Buk-gu, Gwangju 500-757, Korea

[2] Hormone Research Center, School of Biological Sciences and Technology, Chonnam National University, 300, Yongbong-dong, Buk-gu, Gwangju 500-757, Korea

[3] Korea Research Institute of Bioscience and Biotechnology, 685-2, Yangcheongri, Cheong-Won, Choongbuk 363-883, Korea

Key words: Cas6; CRISPR; crystal structure.

INTRODUCTION

Foreign genetic elements imported into the host cells are removed via suitable recognition and degradation systems in host organisms. In bacteria and archaea, clustered regularly interspaced short palindromic repeats (CRISPRs) within the host genomes and CRIPSR-associated (Cas) proteins confer adaptive and heritable immunity against genetic elements of foreign invaders.[1–5]

The CRISPR locus consists of two characteristic sequences; variable and conserved repeat sequences. The heritable adaptive immunity comes from variable sequences, which are derived from the past invaders.[1,3,6,7] After transcription of CRISPR locus, large transcripts are processed into small CRISPR (cr) RNAs that contain individual invader-targeting sequences and form a component of complexes that eliminate incoming foreign nucleic acids.[6,7] The CRISPR locus is accompanied by sets of *cas* genes encoding a series of Cas proteins.[8]

Generation of suitable crRNAs from a large CRISPR transcript is essential for the functioning of CRISPR/Cas defense system,[6] which is accomplished by Cas6 in *Pyrococcus furiosus* (pfCas6).[9,10] In *P. aeruginosa*, Csy4 processes a CRISPR RNA transcript into short crRNAs, which further forms a hairpin structure.[11,12] On the other hand, the processed crRNAs by pfCas6 have been suggested to be non-structured.[11] In the complex structure of pfCas6 with an RNA fragment of ~10 nucleotides corresponding to the first part of a conserved sequence, the RNA is located in the deep cleft of the globular pfCas6 structure.[13] The RNA cleavage site, where the second part of a conserved repeat sequence should be bound, is on the opposite side of the first RNA binding cleft in the globular pfCas6 surface,[9,10] implying the absence of a hairpin structure formed by the interaction between the first and second part of a conserved repeat sequence.

The pfCas6 is a metal-independent endoribonuclease that specifically binds to and cleaves the 3′ end of the repeat sequences of a *P. furiosus* CRISPR transcript.[9] Based on sequence conservation, structural analysis, and mutational studies, Tyr31, His46, and Lys52 have been suggested to constitute a nucleolytic triad.[9,10] However, several regions (red-dotted lines in Fig. 1) including loops around the nucleolytic core were disordered in pfCas6,[9,10,13] which made it difficult to define the nucleolytic mechanism of pfCas6. Here we report the crystal structure of a pfCas6 paralog in *P. furiosus* (pfCas6-3) and compare its structure to pfCas6. The structure of pfCas6-3 has clearly defined loop regions around the nucleolytic core but the specific residues proposed to be essential for catalysis are not conserved, which suggests that pfCas6-3 may not possess endonucleolytic activity. Indeed, pfCas6-3

Additional Supporting Information may be found in the online version of this article.
Grant sponsor: Basic Science Research Program through the National Research Foundation of Korea (Ministry of Education, Science and Technology); Grant number: 2010-0008560; Grant sponsor: World Class Institute (WCI) Program of the National Research Foundation of Korea (NRF) [Ministry of Education, Science and Technology of Korea (MEST)]; Grant sponsor: WCI 2009-002.
Hye-Mi Park and Minsang Shin contributed equally to this work.
*Correspondence to: Jeong-Sun Kim, Department of Chemistry, College of Natural Science, Chonnam National University, 300, Yongbong-dong, Buk-gu, Gwangju, 500-757, Korea. E-mail: jsunkim@chonnam.ac.kr
Received 11 November 2011; Revised 23 January 2012; Accepted 25 January 2012
Published online 22 February 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24061

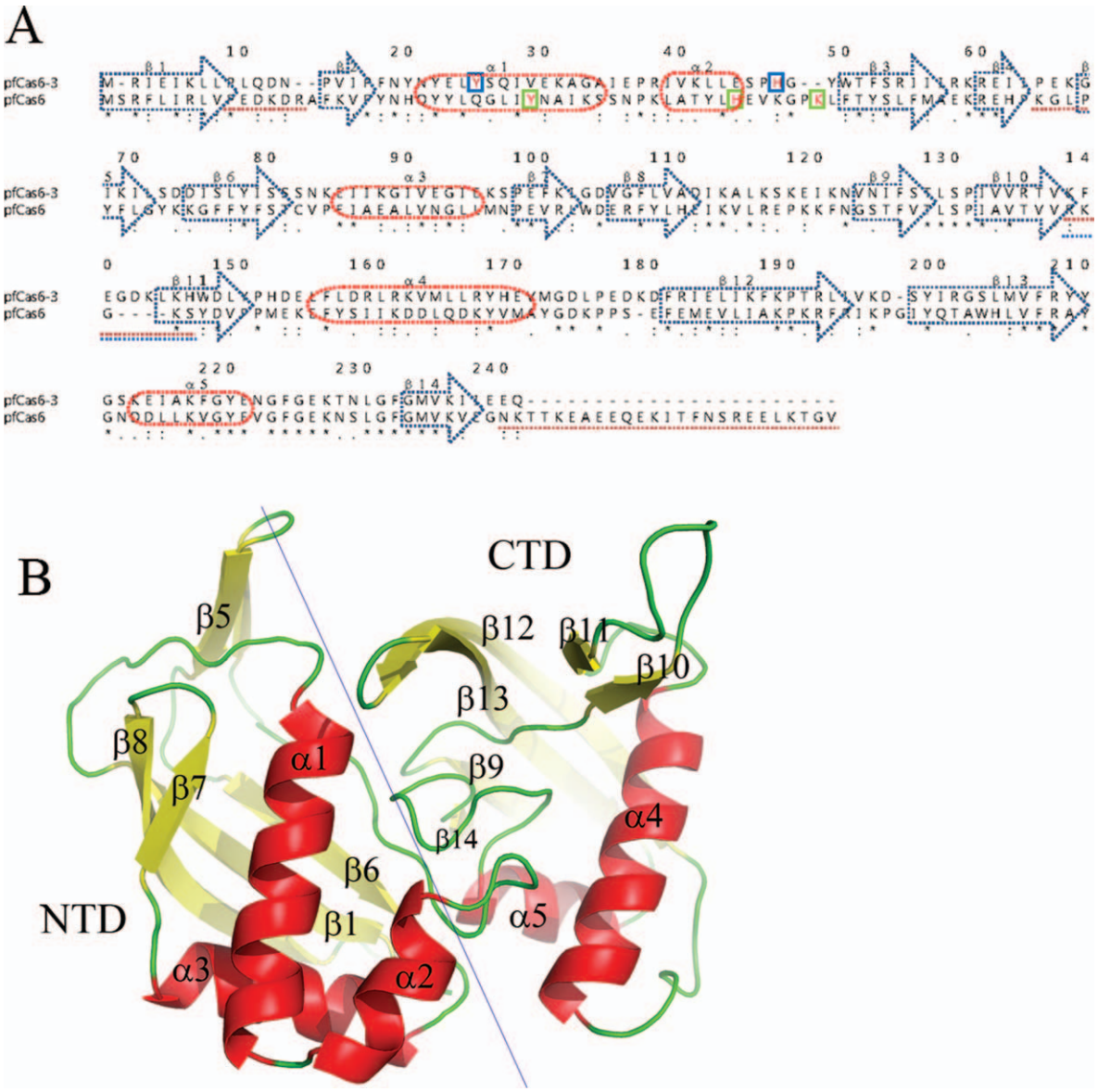

Figure 1

Structure of pfCas6-3. **A: Sequence comparison of Cas6s.** The ellipses (α) and arrows (β) stand for the helix and strand, respectively. The numbering scheme follows the amino acid sequence of pfCas6-3. The catalytic residues at the nucleolytic core are enclosed in green boxes for pfCas6. The disordered regions in pfCas6 are displayed by continuous red-dotted lines below the aligned sequences and regions with weak electron densities for the side chains in pfCas6-3 are presented by continuous blue-dotted lines. Identical residues in all the sequences are marked by an asterisk and conserved residues by a colon or a dot. **B: Crystal structure of pfCas6-3.** The monomeric globular structure consists of two similarly-folded domains (NTD and CTD). The α-helices, loops, and β-strands are represented by red, green, and yellow, respectively. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

does not have activity against the pfCas6 RNA substrate though it does bind to RNA.

MATERIALS AND METHODS

Cloning, expression, and purification of pfCas6-3

The *P. furiosus* gene coding for the pfCas6-3 (Met1∼Gln241) was amplified by a polymerase chain reaction (PCR) using *P. furiosus* chromosomal DNA as a template and primers designed for ligation-independent cloning (LIC).[14] The amplified PCR product was prepared for vector insertion by purification and treatment with T4 DNA polymerase (New England Biolabs, Beverley, MA) in the presence of 1 m*M* dTTP. The prepared insert was annealed into a derivative of pET21a (Novagen, Madison, WI) prepared in 1 m*M* dATP. This vector expresses the cloned gene fused with an N-terminal 6-His-Tobacco etch

virus (TEV) cleavage sequence. The expression construct was transformed into *E. coli* BL21(DE3) star strain and was grown in a 1 L LB medium containing ampicillin (100 μg/mL) at 310 K. After induction by adding 0.5 m*M* IPTG, the culture media was maintained for further 8 h at 310 K. The cell was harvested by centrifugation at 5000*g* at 277 K. Cell pellet was resuspended in the buffer A comprising of 20 m*M* Tris-HCl at pH 7.5 and 500 m*M* NaCl and then disrupted by ultrasonication. The cell debris was removed by centrifugation at 11,000 g for 1 h. The pfCas6-3 fusion protein was purified using a 5-mL Histrap chelating column (GE healthcare, Uppsala, Sweden). The column was extensively washed by buffer A and then the bound protein was eluted with a linear gradient from 0 to 500 m*M* imidazole in buffer A. The eluted sample was dialyzed against the buffer B comprising of 20 m*M* Tris-HCl at pH 7.5 to remove imidazole and 6-His residues were cleaved with TEV. Protein was further purified by ion exchange chromatography by applying the sample to a 5 mL HitrapQ anion exchange column (GE healthcare, Uppsala, Sweden). The column was washed by buffer B and then the bound protein was eluted with a linear gradient from 0 to 1000 m*M* NaCl in buffer B. The purified protein was >95% pure as judged by Coomassie Blue-stained SDS-PAGE.

Crystallization, data collection, and structure determination

The purified protein of pfCas6-3 aggregated heavily upon concentration prior to crystallization trials. This problem was solved by screening many buffers and additives using Optimum Solubility Screen protocol.[15] For concentration, the buffer was changed to 20 m*M* Tris-HCl at pH 8.5, 300 m*M* NaCl, and 5% (v/v) iso-propanol. Subsequently, the purified protein was concentrated up to 5 mg/mL in this buffer. The initial crystallization condition for the pfCas6-3 protein was obtained from Sparse Matrix Screening.[16] Suitable crystals for diffraction experiments were obtained using the hanging-drop vapor-diffusion method at 295 K within 3 days from the precipitant, 15% (w/v) polyethyleneglycol 400, 0.1 *M* $MgCl_2 \cdot H_2O$, and 0.1 *M* Na-HEPES at pH 7.5. For data collection, the crystals were briefly immersed into a precipitant solution containing 15% (v/v) glycerol and immediately placed in a 100 K nitrogen-gas stream. The diffraction data for the pfCas6-3 crystals was collected at the 8.3.1 beamline of the advanced light source (ALS) at a wavelength of 1.1159 Å. The data were then indexed, integrated, and scaled with the HKL-2000 suite.[17] The crystal belonged to the space group of $P3_121$ with unit-cell parameters, $a = b = 79.81$ Å, $c = 117.36$ Å, $\alpha = \beta = 90^\circ$, and $\gamma = 120^\circ$. The crystal structure of pfCas6-3 was solved at a resolution of 2.0 Å using the molecular replacement program, Phaser-1.3.[18] Further model building was performed manually using the programs, Win-Coot[19] and O,[20] and the refinement was performed with PHENIX[21] and CNS.[22] The statistics for the collected data and refinement are summarized in Table I. The coordinates and structure factors were deposited in the Protein Data Bank with an accession number of 3UFC for the pfCas6-3 structure.

Table I
Data Collection and Refinement Statistics

Parameters	pfCas6-3
Synchrotron	8.3.1, ALS
Wavelength (Å)	1.1159
Space group	$P3_121$
Cell parameters	$a = b = 79.81$ Å, $c = 117.36$ Å, $\alpha = \beta = 90^\circ$, and $\gamma = 120^\circ$
Resolution (Å)	50.0–2.0 (2.1–2.0)
Completeness (%)	97.3 (96.2)
R_{sym}[a] (%)	3.4 (31.0)
Reflections (total/unique)	237,711/31,200
R_{factor}[b] (%)/R_{free}[c] (%)	20.3/22.8
No. of atoms (Protein/Water)	1986/182
rmsds (Bonds Å/Angles °)	0.005/1.26
Geometry (%)	
Most favored	95.4
Generously allowed	4.2
Disallowed	0.4

Values in parentheses are for the highest-resolution shell.rmsds, root-mean-square deviations.
[a] $R_{sym} = \Sigma_{hkl} \Sigma_j |I_j - \langle I \rangle| / \Sigma_{hkl} \Sigma_j I_j$, where $\langle I \rangle$ is the mean intensity of reflection *hkl*.
[b] $R_{factor} = \Sigma_{hkl} ||Fobs| - |Fcalc|| / \Sigma_{hkl} |Fobs|$; where Fobs and Fcalc are, respectively, the observed and calculated structure factor amplitudes for the reflections *hkl* included in the refinement.
[c] R_{free} is the same as R_{factor} but is calculated over a randomly selected fraction (8%) of the reflection data not included in the refinement.

RNA binding and cleavage assay

To characterize the RNA binding and cleavage activity of pfCas6-3, two RNA fragments were synthesized as follows; 5′-GUUACAAUAAGACUAAAAUAGAAUUGAAAG-3′ (wild-type) and 5′-GAAUGUUUAAGACUAAAAUAGAAUUGAAAG-3′ (mutant). After purification from PAGE elution, they were end-labeled with [γ-^{32}P] ATP using T4 polynucleotide kinase (Roche, Indianapolis, IN). Reaction mixtures contained 0.1 pmol of end-labeled RNA in 20 m*M* Tris-acetate (pH 8.0), 15 m*M* $MgCl_2$, 300 m*M* K-glutamate, 1 m*M* DTT, and various concentrations of pfCas6-3. The reactions were allowed to proceed for 30 min at 70°C and the reaction mixture was separated by 8% (w/v) native PAGE at 100 V for 1.5 h. RNA cleavage was assessed by separation of RNAs on denaturing, 17% (w/v)-sequencing gel containing 8 *M* urea.

RESULTS AND DISCUSSION

Overall feature of pfCas6-3

The pfCas6-3 has ~30% of sequence identity and ~60% of homology to pfCas6 and contains several conserved regions including the Gly-rich loop (GRL, ^{223}GFG××××GFG^{233}) [Fig. 1(A)]. Interestingly, three

residues responsible for forming a nucleolytic triad in pfCas6 are not conserved in pfCas6-3.

The crystal structure of pfCas6-3 was determined by molecular replacement using the pfCas6 structure (PDB ID 3I4H) as a search model and refined to a resolution of 2.0 Å (Table I). One pfCas6-3 molecule exists in the asymmetric unit. All the 241 residues were modeled in pfCas6-3. Several loops disordered in the pfCas6 and/or RNA-pfCas6 structures [red-dotted lines in Fig.1(A)] were modeled in the present structure. Weak electron density was observed on side chains of the β10-β11 and β12-β13 loops [blue-dotted lines in Fig.1(A)]. The analysis of the modeled residues by PROCHECK[23] showed that all the modeled residues were found in the valid regions of the Ramachandran plot except Lys66 and Lys196 that have defined electron density (Table I). However, none of these two residues was involved in the formation of the putative nucleolytic core or the distal RNA binding cleft.

The traced 241 amino acids form 5 α-helices and 14 β-strands, which comprises of two ferredoxin-like domains that are connected by a loop (Ile112-Asn122) between the β8-strand of the N-terminal domain (NTD, Met1-Asp111) and the β9-strand of the C-terminal domain (CTD, Val123-Glu240). In NTD, a central β-sheet of five β-strands (β1, β3, β6, β7, and β8) is intertwined with one β-sheet of three strands (β2, β4, and β5) at one side and one α-helical bundle (α1, α2, and α3) at the other side [Fig. 1(B)]. In CTD, two long β-strands (β12 and β13) are surrounded by a couple of two short β-strands (β9 and β14, β10 and β11) and two α-helices (α4 and α5) [Fig. 1(B)].

Two domains assembled to form a single globular structure. Along the domain interface, there are two deep clefts, which are annotated as a nucleolytic core and a distal RNA binding site in pfCas6. At the putative nucleolytic core of pfCas6-3, the α2-β3 loop interacts with GRL and the α4-helix [Fig. 2(A)]. The α1-helix also interacts with the β12-strand and the β12-β13 loop [Fig. 1(B)]. At the distal RNA binding site, two β-strands (β3 and β6) make polar interactions with the GRL-containing α5-β14 loop and the β9-β10 loop. The last interaction between the two domains is observed at the long α5-β14 loop, which is packed into the open space formed between the α1 and α2 helices.

Comparison of pfCas6-3 with pfCas6

Even though the revealed pfCas6-3 structure is well superposed on pfCas6 with small root-mean-square-deviations (rmsds) of 1.8 Å [Fig. 2(A,B)], there is an intrinsic difference between these two paralogous proteins [Figs. 1(A) and 2(A)]. The most noticeable difference is observed at the C-termini of both the proteins [Fig. 1(A)]. The pfCas6 has an additional region of ~20 amino acids at the extreme C-terminus, which includes a number of charged residues and was disordered in the structures of pfCas6

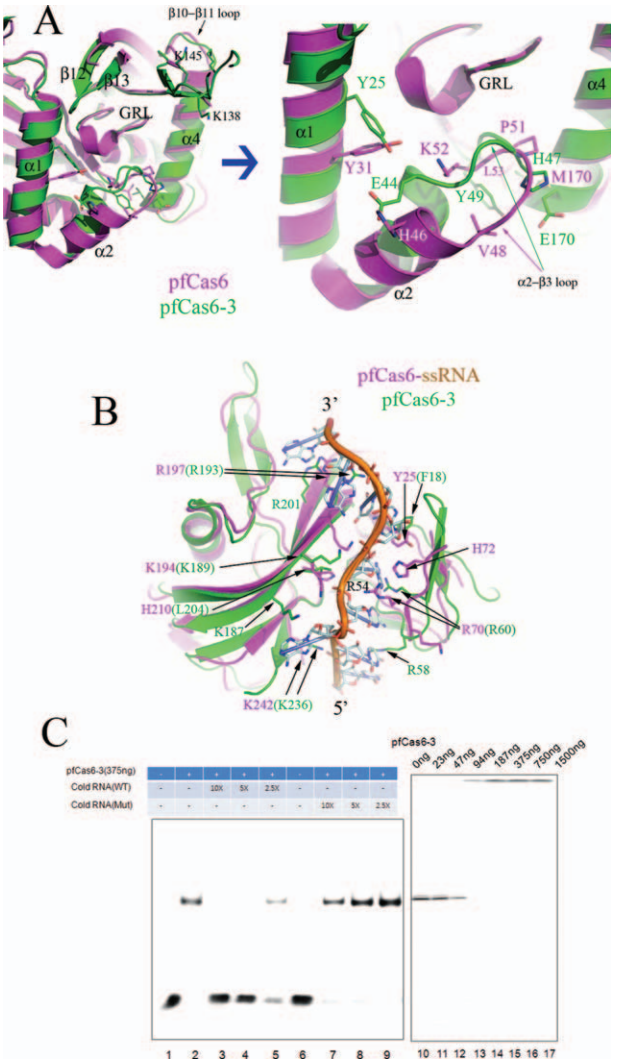

Figure 2

Comparison of two paralogous Cas6 proteins. **A: At the putative nucleolytic site.** The figure is oriented to clearly show the differences around the putative nucleolytic cores of two superposed structures. Two paralogous structures are displayed with Ca traces and residues responsible for nucleolytic activity in pfCas6 are drawn as stick models in magenta. Two residues that occupy similar spatial positions of pfCas6-3 to the nucleolytic residues of pfCas6 are displayed in green and their atoms are displayed in yellow and red for sulfur and oxygen atoms, respectively. Large deviations are observed at the α2-helix and the loop that connects the α2-helix to the β3-strand. A number of hydrophobic residues, which affect the environment of each putative nucleolytic core, are displayed by stick models and the nucleic acid binding motif, GRL, is marked. **B: At the distal RNA binding cleft.** The bound RNA fragment (ssRNA) into the cleft of pfCas6 13 and residues that interact with a RNA fragment are displayed with stick models. The conserved residues around this cleft in both the proteins have been displayed and labeled. **C: Gel mobility shift.** The wild-type synthetic RNAs were radioactively labeled and then used as a probe. The lanes 2–7 were loaded with the probe (0.1 pmole) mixed with pfCas6-3 (375 ng). On lane 3–5, a wild-type probe without end-labeling [Cold RNA (WT)] was loaded with the end-labeled wild-type probe as a competitor. On lane 7–9, a mutant probe without end-labeling [Cold RNA (Mut)] was also loaded with end-labeled wild-type probe as a competitor for a specific binding of pfCas6-3. The lanes 11–17 were loaded with the end-labeled wild-type probe mixed with increasing amounts of pfCas6-3.

alone and in its complex with the crRNA fragment.[9,13] However, this region is not conserved in pfCas6-3 as well as in its homologous proteins (Supporting Information Fig. S1). Furthermore, the traced C-terminus of pfCas6 is far from the nucleolytic core and the distal RNA binding site. Therefore, its role is obscure or might be negligible in exerting the catalytic activity of pfCas6.

Another difference is found in the middle of the aligned sequences, where some of the amino acids are inserted in one of the two proteins [Fig. 1(A)]. They were not traced in pfCas6 [red-dotted lines of Fig.1(A)] and correspond to loop regions (loops of β1-β2, α2-β3, β10-β11, and α4-β12) in pfCas6-3. Among them, the difference at the α2-β3 loop near the nucleolytic core of pfCas6 may discriminate the functional variation of pfCas6 and pfCas6-3. The last and most remarkable difference upon their sequence comparison is found in the residues that have been suggested to form a catalytic triad in pfCas6. These two latter features are described in the following section.

Distal RNA binding site

The complex structure of pfCas6 with an upstream crRNA fragment reveals the presence of a single-stranded RNA fragment in the cleft of the distal RNA binding site.[13] Several positively charged residues from both the domains are aligned along the distal RNA binding cleft and form ionic interactions with the bound crRNA fragment in the crRNA-pfCas6 complex structure.[13] For example, residues of Arg12 (Lys6 in pfCas6-3), Arg70 (Arg58 in pfCas6-3), Lys194 (Lys189 in pfCas6-3), Arg197 (Arg193 in pfCas6-3), and Lys242 (Lys238 in pfCas6-3) interact with phosphate moieties of the bound RNA fragment. The Tyr25 (Phe18 in pfCas6-3) also contributes to the RNA binding by providing an aromatic interaction with one base of the bound RNA fragment [Fig. 2(B)].

Most of the residues at the distal RNA binding site of pfCas6 are conserved in pfCas6-3 (residues in parentheses above) and occupy structurally equivalent positions to those of pfCas6 [Fig. 2(B)]. Further, positively charged residues are additionally observed along the cleft in pfCas6-3, for example, Arg58 and Arg201 [Fig. 2(B)]. Based on these structural features, we tested the RNA binding activity of pfCas6-3 with the repeated crRNA sequence used in the RNA binding assay of pfCas6. As expected, pfCas6-3 exhibited the RNA binding activity against the used crRNA [Fig. 2(C)], implicating that similar kinds of interactions are involved as those in the pfCas6-crRNA complex structure.

Putative nucleolytic core of pfCas6-3

The cleft formed on the opposite side of the distal RNA binding site is responsible for the nucleolytic activity of pfCas6, where Tyr31, His46, and Lys52 are located without any direct interaction amongst them. Based on this structural feature as well as their conservation among the compared Cas6 homologous proteins, these three residues were suggested to form a catalytic triad.[9] Successive mutational studies of these individual residues into the Ala residues have shown that all the three respective Ala mutations dramatically decreased RNA-degrading activity.[10]

However, three residues in pfCas6 are not strictly conserved in pfCas6-3. As mentioned above, the largest difference upon comparison of two structures is observed around the nucleolytic core, especially in the spatial positions of the α2-helix and the α2-β3 loop that contains two catalytic residues, His46 and Lys52 [Figs. 1(A) and 2(A)]. The α2-helix forms α-helical bundle with the α1-helix and interacts with the α4-helix [Fig. 2(A)]. In pfCas6, smaller hydrophobic residues in the α2-β3 loop (Val48, Gly50, Pro51, and Leu53) are entangled with Met170 and Ala171 on the α4-helix [Fig. 2(A)]. On the other hand, Pro51 and Leu53 in pfCas6 correspond to residues His47 and Tyr49 in pfCas6-3, respectively [Fig. 1(A)]. The insertion of two amino acids on the α2-β3 loop and smaller residues in pfCas6 mediate tight interaction with the α4-helix and locate the α2-helix away from the GRL and the α1-helix. This spatial arrangement confers sufficient space for the formation of a RNA binding cleft [Fig. 2(A)], which might be necessary for nucleolytic activity of pfCas6 family. On the other hand, short α2-β3 loop and large aromatic residues in pfCas6-3 position the α2-helix far away from the α4-helix and close to the α1-helix [Fig. 2(A)]. This minute difference invokes tight interaction with the α1-helix and the GRL in pfCas6-3, and not with the α4-helix.

At the opposite site of the α2-helix, there is a β10-β11 loop, which was not traced in pfCas6 structure.[9,13] On the other hand, four residues are present at that position in pfCas6-3 when compared with pfCas6 [Fig. 1(A)]. Nonetheless, positively charged residues (Lys138, Lys143, and Lys145 of pfCas6-3) are observed in both the proteins. Because this region is located around the catalytic triad and the substrate RNA molecule has a number of negatively charged moieties, it is proposed that they might be involved in crRNA binding at the nucleolytic core for the RNA cleavage process.

Collectively, the putative nucleolytic core is well-traced in pfCas6-3 structure with minute structural differences from pfCas6. Since several Cas6 homologous proteins do not possess the residues for forming a nucleolytic triad in pfCas6 (Supporting Information Fig. S2), we further tested for the presence of a crRNA cleavage activity of pfCas6-3 against the crRNA fragment used for pfCas6 degrading activity. However, no significant RNA-degrading activity was observed by employing this crRNA fragment [Fig. 2(C)]. Therefore, this pfCas6-3 might lack nucleolytic activity against the nucleotide sequence that was degraded by pfCas6. Nonetheless, we could not

totally ignore the aspect that this pfCas6-3 does not exhibit any nucleolytic activity.

ACKNOWLEDGMENTS

The authors wish to thank the staff scientists at ALS for their assistance with data collection.

REFERENCES

1. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology 2005;151:2551–2561.
2. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol 2005;60:174–182.
3. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. CRISPR provides acquired resistance against viruses in prokaryotes. Science 2007;315:1709–1712.
4. Sorek R, Kunin V, Hugenholtz P. CRISPR—a widespread system that provides acquired resistance against phages in bacteria and archaea. Nat Rev Microbiol 2008;6:181–186.
5. van der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJ. CRISPR-based adaptive and heritable immunity in prokaryotes. Trends Biochem Sci 2009;34:401–407.
6. Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, Dickman MJ, Makarova KS, Koonin EV, van der Oost J. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 2008;321:960–964.
7. Marraffini LA, Sontheimer EJ. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nat Rev Genet 2010;11:181–190.
8. Jansen R, Embden JD, Gaastra W, Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. Mol Microbiol 2002;43:1565–1575.
9. Carte J, Wang R, Li H, Terns RM, Terns MP. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. Genes Dev 2008;22:3489–3496.
10. Carte J, Pfister NT, Compton MM, Terns RM, Terns MP. Binding and cleavage of CRISPR RNA by Cas6. RNA 2010;16:2181–2188.
11. Kunin V, Sorek R, Hugenholtz P. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. Genome Biol 2007;8:R61.
12. Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Doudna JA. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. Science 2010;329:1355–1358.
13. Wang R, Preamplume G, Terns MP, Terns RM, Li H. Interaction of the Cas6 riboendonuclease with CRISPR RNAs: recognition and cleavage. Structure 2011;19:257–264.
14. Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res 1990;20:6069–6074.
15. Jancarik J, Pufan R, Hong C, Kim R, Kim SH. Optimum solubility (OS) screening: an efficient method to optimize buffer conditions for homogeneity and crystallization of proteins. Acta Crystallogr D Biol Crystallogr 2004;60:1670–1673.
16. Jancarik J, Kim SH. Sparse matrix sampling: a screening method for crystallization of proteins. J Appl Cryst 1991;23:409–411.
17. Otwinowski Z, Minor W. Procession of x-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
18. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ. Likelihood-enhanced fast translation functions. Acta Crystallogr D Biol Crystallogr 2005;61:458–464.
19. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 2004;60:2126–2132.
20. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr A 1991;47:110–117.
21. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 2010;66:213–221.
22. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998;54:905–921.
23. Laskowski R, MacArthur M, Hutchinson E, Thorton J. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.