FEBS Letters 587 (2013) 562–568

journal homepage: www.FEBSLetters.org

Crystal structure of Cmr5 from *Pyrococcus furiosus* and its functional implications

Jeong-Hoh Park [1], Jiali Sun [1], Suk-Youl Park, Hyo-Jeong Hwang, Mi-Young Park, Minsang Shin, Jeong-Sun Kim *

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

ARTICLE INFO

Article history:
Received 5 November 2012
Revised 8 January 2013
Accepted 9 January 2013
Available online 28 January 2013

Edited by Stuart Ferguson

Keywords:
Cmr5
Clustered regularly interspaced short palindromic repeats
Crystal structure

ABSTRACT

The bacterial acquired immune system consists of clustered regularly interspaced short palindromic repeats (CRISPRs) and CRIPSR-associated (Cas) genes, which include Cas-module repeat-associated mysterious proteins (Cmr). The six Cmr proteins of *Pyrococcus furiosus* (pfCmr1–pfCmr6) form a Cmr effector complex that functions against exogenous nucleic acid. Among the Cmr proteins, the role of pfCmr5 and its involvement in the complex's cleavage activity have been obscure. The elucidated pfCmr5 structure has two inserted α-helices compared with the other trimeric Cmr5 structure. However, pfCmr5 exists as a monomeric protein both in the crystalline state and in solution. In vitro assays indicate that pfCmr5 interacts with pfCmr4. These structural and biophysical data might help in understanding the complicated and ill-characterized Cmr effector complex.

Structured summary of protein interactions:
pfCmr4 and **pfCmr5** bind by molecular sieving (View interaction)
pfCmr4 and **pfCmr4** bind by molecular sieving (View interaction)
pfCmr5 and **pfCmr4** bind by ion exchange chromatography (View interaction)

1. Introduction

In bacteria and archaea, clustered regularly interspaced short palindromic repeats (CRISPRs) and CRIPSR-associated (Cas) proteins confer heritable and adaptive immunity against incoming genetic elements [1–5]. Among the two characteristic sequences in the CRISPR locus, variable and conserved repeat sequences, the information about the heritable and adaptive immunity is stored in the variable sequences, which are derived from past encounters and can be added from new exogenous genes [1,3,6,7].

In *Pyrococcus furiosus*, small crRNAs and the repeat-associated mysterious proteins (RAMPs) module-containing Cas proteins (Cmr) form a Cmr effector complex, which cleaves complementary exogenous target RNAs at a fixed distance from the 3′-end of the integral crRNAs. The isolated Cmr effector complex from the crude cell fraction of *P. furiosus* has been shown to contain six Cmr proteins (pfCmr1–pfCmr6). However, the elimination of pfCmr5 in the recombined effector complex leads to no observable effect on the cleavage activity against added RNA [8]. The Cmr effector complex in *Sulfolobus solfataricus* has also been confirmed to have six proteins (ssCmr1–ssCmr6). Furthermore, the complex includes another protein, ssCmr7 [9]. These data invoke questions concerning the composition of Cmr proteins and their roles in the effector complex.

Detailed information on the Cmr5 protein has been obtained from the recent crystal structure of Cmr5 from *Thermus thermophilus* HB8 (ttCmr5). The protein forms a single globular structure of six α-helices. Seven hydrophobic residues from the two neighboring molecules form a hydrophobic interaction at the interfaces, which has been suggested as a main force facilitating the assembly of a trimeric ttCmr3 structure. Furthermore, its molecular size in solution is close to the trimeric protein. In addition, ttCmr5 has been suggested to be a nucleic acid-binding protein based on the

Abbreviations: afCmr5, Cmr5 from *Archaeoglobus fulgidus*; ASU, asymmetric unit; Cas, CRIPSR-associated gene; Cmr, Cas-module RAMP; CRISPR, clustered regularly interspaced short palindromic repeats; M.W., molecular weight; PCR, polymerase chain reaction; PDB, Protein Data Bank; pI, isoelectric point; pfCmr5, Cmr5 from *P. furiosus*; RAMP, repeat-associated mysterious proteins; rmsd, root-mean-square deviation; SEC, size exclusion chromatography; SSRF, Shanghai Synchrotron Radiation Facility; TEV, tobacco etch virus; ttCmr5, Cmr5 from *Thermus thermophilus* HB8

* Corresponding author. Address: Department of Chemistry, College of Natural Science, Chonnam National University, 300 Yongbong-dong, Buk-gu, Gwangju 500-757, Republic of Korea. Fax: +82 62 530 3389.
E-mail address: jsunkim@chonnam.ac.kr (J.-S. Kim).

[1] These authors contributed equally to this work.

0014-5793/$36.00
http://dx.doi.org/10.1016/j.febslet.2013.01.029

Table 1
Data collection and refinement statistics.

Parameters	pfCmr5
Synchrotron	BL17U1, SSRF
Wavelength (Å)	0.9762
Space group	$P4_32_12$
Cell parameters	$a = b = 53.94$ Å, $c = 235.99$ Å, $\alpha = \beta = \gamma = 90°$
Resolution (Å)	50.0–2.0 (2.03–2.0)
Completeness (%)	93.2 (85.4)
R_{sym}[a] (%)	5.7 (46.5)
Reflections, total/unique	102, 154/23 080
R_{factor}[b] (%)/R_{free}[c] (%)	20.8/26.0
No. of atoms, protein/water	2, 423/183
rmsds, bonds Å/angles o	0.002/0.69
Geometry (%)	
Favored	99.65
Allowed	0.35
Outliers	0

Values in parentheses are for the highest-resolution shell.
rmsds: root-mean-square deviations.
[a] $R_{sym} = \sum_{hkl}\sum_j |I_j - \langle I\rangle| / \sum_{hkl}\sum_j I_j$, where $\langle I\rangle$ is the mean intensity of reflection *hkl*.
[b] $R_{factor} = \sum_{hkl} ||Fobs| - |Fcalc|| / \sum_{hkl} |Fobs|$, 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.

conserved RAMP module structure and localized basic patch at the subunit interface [10]. On the other hand, the crystal structure of Cmr5 from *Archaeoglobus fulgidus* (afCmr5), which has been deposited in the Protein Data Bank (PDB) and publicly released (PDB ID 2oeb at http://www.rcsb.org/pdb), contains a single molecule in the asymmetric unit (ASU). However, a trimeric molecule cannot be generated by applying the crystallographic H32 symmetry. These data indicate that the biophysical properties of Cmr5 and its role in the formation of the effector complex are still controversial and should be investigated.

We elucidated the crystal structure of pfCmr5. A comparison with ttCmr5 revealed the insertion of two α-helices in the middle of the aligned sequence. Two molecules in the ASU cannot generate a trimeric structure that has been shown in ttCmr5. Furthermore, pfCmr5 behaved as a monomer in solution. Our in vitro binding assays exhibited the presence of a direct interaction between the pfCmr5 and pfCmr4 proteins.

2. Materials and methods

2.1. Cloning, expression, and purification of pfCmr5 and pfCmr4

The *P. furiosus* genes coding for pfCmr5 (Met1-Ser169, M.W. 19.7 kDa) and pfCmr4 (Met1-Lys295, M.W. 32.6 kDa) were

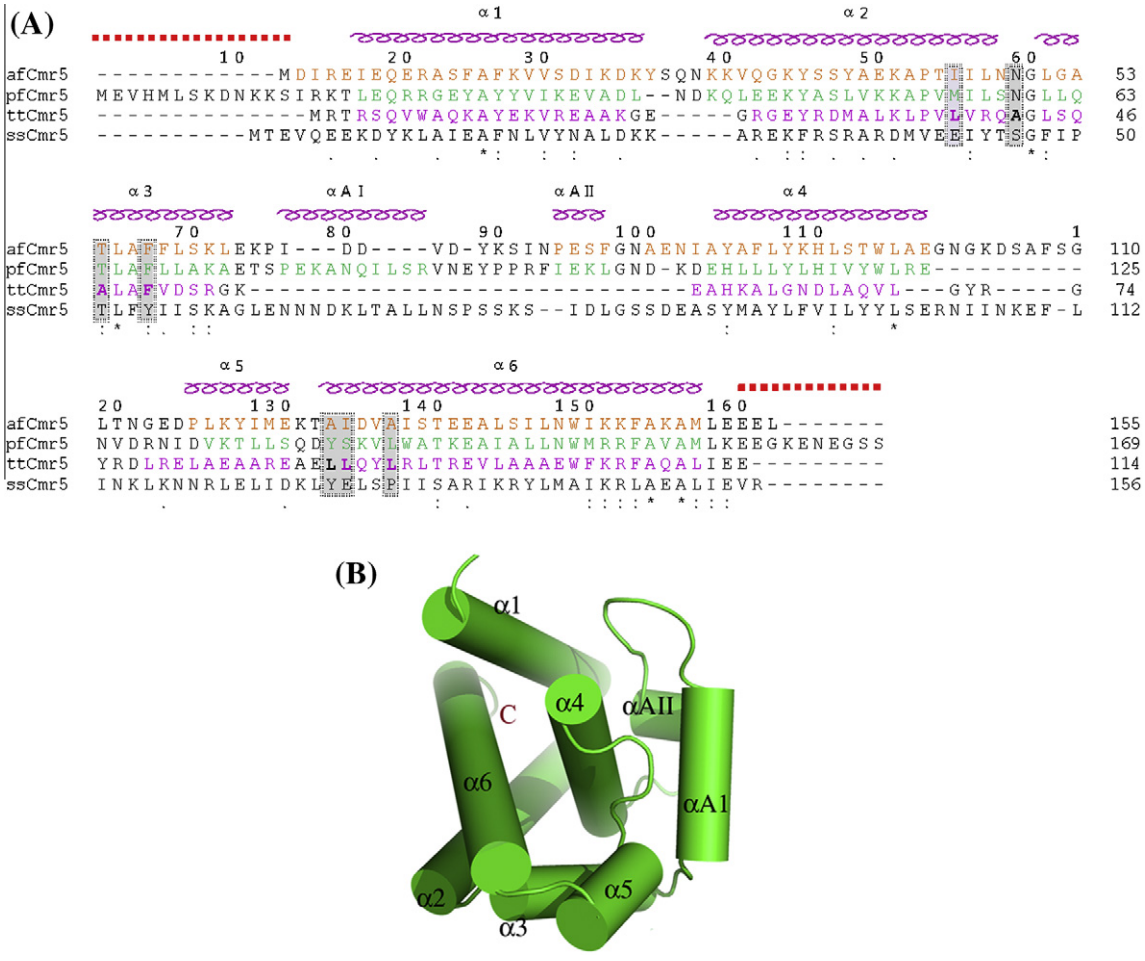

Fig. 1. The overall features of pfCmr5. (A) Structure-based sequence alignment of Cmr5s. The amino acid sequences are shown for Cmr5 from *P. furiosus* (pfCmr5), Cmr5 from *T. thermophilus* HB8 (ttCmr5), Cmr5 from *A. fulgidus* (afCmr5), and Cmr5 from *S. solfataricus* (ssCmr5). The coils (α) stand for the α-helix. The numbering scheme follows the amino acid sequence of pfCmr5. The continuous red-dotted lines above the aligned sequences display the extreme N-terminal and C-terminal regions that were not traced in the present structure. Identical residues are marked by "∗". Conserved residues are indicated by ":" or ".", respectively. The residues in the α-helices are colored. The residues forming hydrophobic interactions at the interfaces of two subunits are indicated by bold in the dotted boxes of grey background. The sequence alignment was prepared using the ClustalW2 program of the European Bioinformatics Institute and based on primary amino acid sequences [19]. (B) The overall shape of pfCmr5. The α-helices are represented as cylinders.