FEBS Letters 586 (2012) 3956–3961

FEBS Letters

journal homepage: www.FEBSLetters.org

Nucleic acid binding surface and dimer interface revealed by CRISPR-associated CasB protein structures

Ki Hyun Nam [a], Qingqiu Huang [a,b], Ailong Ke [a,*]

[a] Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14850, USA
[b] Macromolecular Diffraction Facility at CHESS, Cornell University, Ithaca, NY 14850, USA

ARTICLE INFO

Article history:
Received 8 September 2012
Revised 26 September 2012
Accepted 26 September 2012
Available online 16 October 2012

Edited by Richard Cogdell

Keywords:
CRISPR
Cas
CasB
Nucleic acid binding
Cascade

ABSTRACT

The CRISPR system is an adaptive RNA-based microbial immune system against invasive genetic elements. CasB is an essential protein component in Type I-E Cascade. Here, we characterize CasB proteins from three different organisms as non-specific nucleic acid binding proteins. The *Thermobifida fusca* CasB crystal structure reveals conserved positive surface charges, which we show are important for its nucleic acid binding function. EM docking reveals that CasB dimerization aligns individual nucleic acid binding surfaces into a curved, elongated binding surface inside Type I-E Cascade, consistent with the putative functions of CasB in ds-DNA recruitment and crRNA–DNA duplex formation steps.

Structured summary of protein interactions:
TthCasB and **TthCasB** bind by x-ray crystallography (View interaction)
TfuCasB1 and **TfuCasB1** bind by molecular sieving (View Interaction: 1, 2)

1. Introduction

CRISPR drives adaptation to harmful invading nucleic acids – such as conjugative plasmids, transposable elements and phages – using an RNA-mediated interference mechanism with fundamental similarities to our innate and adaptive immune responses [1,2]. CRISPR and *cas* (CRISPR-associated) genes provide resistance against the invading nucleic acids through three molecular events. (1) CRISPR adaptation upon exposure to new foreign genetic elements through the insertion of a new spacer into the genomic CRISPR array; (2) transcription of the CRISPR array (pre-crRNA) and the endonucleolytic processing of it into the mature form (crRNA); and (3) crRNA-mediating degradation of the foreign nucleic acids (in most cases DNA, sometimes RNA) complementary to the spacer sequences in crRNA [1–3].

One interesting observation is the existence of multiple subtypes among CRISPR-Cas systems [3]. A recent study classified CRISPR-Cas systems into Type I, II, and III based on the presence of mutually exclusive signature proteins Cas3, Csn1(Cas9), and Cmr2(Cas10), respectively [3]. CRISPR interference mechanisms are quite different among these three types [2,3]. Each Type can be further divided into several subtypes. For example, the most wide-spread Type I system can be further classified into six different subtypes (I-A to I-F), each encoding a unique set of subtype-specific genes [3].

In the most well-studied Type I-E CRISPR system, the processed crRNA guides the assembly of five Cas proteins (CasA-E) into a multi-subunit ribonucleoprotein complex named Cascade (CRISPR-associated complex for anti-viral defense). The *Escherichia coli* Cascade complex is a 405 kDa ribonucleoprotein complex assembled from crRNA and five functionally essential Cse proteins: one CasA(Cse1) protein, two copies of CasB(Cse2), one CasE(Cse3), six copies of CasC(Cse4), and one CasD(Cas5e) [4]. Electron microscopy (EM) reconstruction show that the *E. coli* Cascade adopts a seahorse-shaped architecture with a helical backbone [5]. It functions by promoting invasion and selective base-pairing between crRNA and target DNA, forming a so-called "R-loop" structure [6,7]. Following target recognition by Cascade, it is speculated that Cas3, possessing HD nuclease and helicase activities, interacts with Cascade and further degrades the DNA target [6].

The CasB protein is only found in the Type I-E CRISPR system [3]. It is an essential component of the Type I-E Cascade and deletion of *E. coli* CasB leads to the loss of resistance against phage λ

Abbreviations: CRISPR, clustered regularly interspaced short palindromic repeats; Cas, CRISPR-associated; Cascade, CRISPR-associated complex for antiviral defense; TfuCasB$_1$, *Thermobifida fusca* CasB$_1$; TfuCasB$_2$, *Thermobifida fusca* CasB$_2$; TthCasB, *Thermus thermophilus* CasB; EcoCasB, *Escherichia coli* CasB; EMSA, electrophoretic gel mobility shift assay

* Corresponding author. Fax: +1 607 255 6249.
E-mail address: ak425@cornell.edu (A. Ke).

0014-5793/$36.00
http://dx.doi.org/10.1016/j.febslet.2012.09.041

infection [4,5]. Its strategic location along the crRNA-binding "spine" of the *E. coli* Cascade, and the observation that a conformational change occurs upon target single stranded (ss)-DNA binding, suggest that the function of CasB may be to provide non-specific nucleic acid interacting surfaces to facilitate the invasion of crRNA into the target double stranded (ds)-DNA. Although the crystal structure of *Thermus thermophilus* CasB has been reported, the detailed molecular mechanism about how CasB dimerizes inside Cascade and how this leads to presumably more efficient binding of nucleic acid remain elusive [8].

Here we characterize four CasB proteins from three species as non-specific nucleic acid binding proteins, and determine the crystal structures of *Thermobifida fusca* $CasB_2$ and *T. thermophilus* CasB (from an alternative crystal form [8]). Both structures adopt an α-helical bundle scaffold and display a positive surface for nucleic acid binding. Mutagenesis of residues on this conserved surface in *T. fusca* $CasB_2$ impairs its nucleic acid binding function. Docking of CasB coordinates into the EM envelope of *E. coli* Cascade provides insight into how CasB dimerizes to present an elongated nucleic acid binding surface in the Cascade complex. Our data contributes to the understanding of CasB function inside Cascade in Tyep I-E/*Ecoli* CRISPR system.

2. Materials and methods

2.1. Cloning and protein purification

The full-length coding region of *T. fusca* $CasB_1$ (Accession No.: Q47PJ2) and $CasB_2$ (Q47PI5), *T. thermophilus* CasB (Q53VY0) and *E. coli* CasB (P76632) were PCR-amplified from the their genomic DNA. The PCR product of $TfuCasB_1$, EcoCasB and TthCasB were cloned into the pQE-80 vector (Qiagne). The PCR product of $TfuCasB_2$ was cloned into a modified pSUMO vector. Recombinant proteins were expressed in *E. coli* BL21(DE3) star cells by inducing at OD_{600} of 0.8 using 0.5 mM IPTG for 18 h at 18 °C. In the purification of $TfuCasB_1$, EcoCasB and TthCasB proteins, cells were harvested and sonicated in cold buffer A (50 mM Tris–HCl pH8.5, 150 mM NaCl, 10 mM β-mercaptoethanol, and 20 mM imidazole). The supernatant was loaded onto the Ni–NTA column, washed with buffer A and eluted with the same buffer supplemented with 300 mM imidazole. Elution samples were then loaded onto a Heparin column (GE Healthcare) pre-equilibrated with 20 mM Tris–HCl pH 8.0, 50 mM NaCl and 2 mM DTT, and eluted with a NaCl gradient from 0.1 to 1.0 M in the same equilibration buffer. Fractions containing pure CasB protein were pooled, concentrated, and loaded onto a Superdex 200 10/300 column (GE Healthcare) equilibrated in 10 mM Tris–HCl pH 8.0, 200 mM NaCl and 2 mM DTT. The same purification procedures were followed for $TfuCasB_2$, except that an additional Ni–NTA back-absorption step was included after SUMO protease digestion to remove the cleaved SUMO tag.

2.2. Oligomeric state determination

Purified CasB proteins (~10 mg/ml) were analyzed in buffer containing 10 mM Tris–HCl pH8.0, 200 mM NaCl, and 2 mM DTT using an analytical Superdex 200 HR 10/30 column (GE Healthcare) at 4 °C at 0.4 ml/min. Protein standards from the gel filtration standard calibration kit (Bio-RAD) were separated under the identical condition. The protein elution peak was monitored by the absorbance measurement at 280 nm.

2.3. Electrophoretic mobility shift assay (EMSA)

The ds-DNA (291 bp) and ss-RNA (634 nt) substrates were obtained by PCR-methods and in vitro transcription, respectively (Supplemental Table 1). The purified CasB proteins were incubated with ds-DNA (1.1–2.8 μM) and ss-RNA (0.5 μM) in 10 μL reaction volume containing 20 mM HEPES pH 7.5, 50 mM NaCl, and 1 mM DTT at 25 °C for 20 min. EMSA was performed on a 1.2% (w/v) agarose gel in 1X TAE buffer. The ds-DNA and ss-RNA substrates were visualized by ethidium bromide staining and analyzed using a Doc XR System (BIO-RAD).

2.4. Crystallization, data collection and structure determination

The CasB proteins were concentrated to 20 mg/ml in 10 mM Tris–HCl pH8.0, 200 mM NaCl, and 2 mM DTT for hanging-drop vapor-diffusion crystallization at 18 °C. The $TfuCasB_2$ crystals were grown from a well solution containing 0.1 M Tris–HCl, pH 8.0–8.5, 2% (v/v) ethanol. The TthCasB crystals were obtained from a well solution containing 0.1 M MES, pH 5.7–5.9, 50 mM $MgSO_4$, and 5% (w/v) PEG 3350, respectively. The crystals were cryoprotected by the addition of 30% (w/v) ethylene glycol. The X-ray diffraction data sets were collected at beamline A1 at MacCHESS, and processed using the software, HKL2000 [9]. The MOLREP program [10] within the CCP4 program suite was used for molecular replacement using the *T .thermophlius* CasB as the search model (PDB code 2ZCA) [8]. Crystallographic refinement was carried out using Refmac [11]. Several rounds of restrained refinement were alternated with rebuilding in Coot [12]. The geometry of the final model was checked using MolProbity [13]. The data collection and structural refinement statistics are shown in Supplemental Table 2. The atomic coordinates and structure factors (PDB code 4H79 for $TfuCasB_2$ and 4H7A for TthCasB) have been deposited in the Protein Data Bank.

2.5. Docking of the CasB into cryo-EM map of Cascade complex

The coordinates of CasB were docked into the cryo-electron microscopy (Cryo-EM) density maps of the *E. coli* Cascade complex (electron microscopy data bank under accession numbers of 5314 and 5315). The *E. coli* CasB dimer corresponding densities were selected, and the CasB crystal structure was docked sequentially into each CasB density using the automated EM docking procedure in program Chimera [14]. The docked coordinates were further docked and refined using the program Sculptor [15] following the feature extraction procedure.

2.6. Structural analysis

The structure-based sequence alignment was carried out using ESPRIPT [16]. Surface conservation within the $TfuCasB_1$ was calculated and illustrated using the Consurf server [17]. Dimer interfaces of CasB proteins were analyzed using PISA protein-interface webserver [18]. Molecular graphic figures were generated using PyMOL (http://www.pymol.org).

3. Results

3.1. CasB proteins interacted with nucleic acids non-specifically

To better understand the molecular function of CasB, we characterized the biochemical behavior of four CasB proteins from *T. fusca* (containing two CasB denoted as $TfuCasB_1$ and $TfuCasB_2$, one in each of the two CRISPR operons), *T. thermophilus* and *E. coli*. The oligomerization state of these CasB homologs varied significantly as judged by the size exclusion chromatography (SEC) (Fig. 1A and B). $TfuCasB_1$ (27.4 kDa, 244aa, pI 5.94) and $TfuCasB_2$ (24.5 kDa, 214aa, pI 9.55) proteins migrated as a dimer/aggregates mixture and dimer, respectively. TthCasB (19.4 kDa, 169 aa, pI 9.18) migrates as a monomer, whereas EcoCasB (18.7 kDa, 160

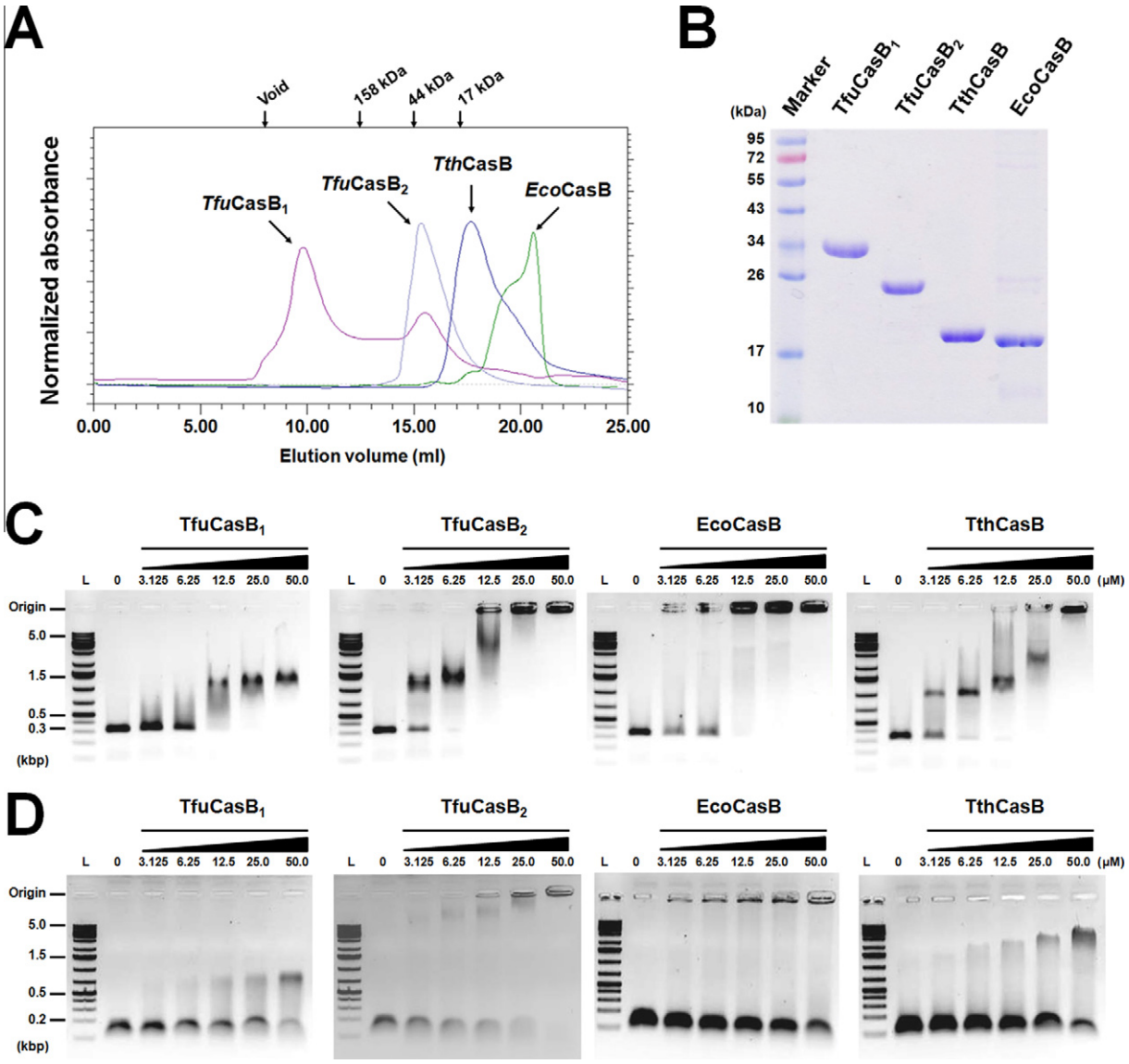

Fig. 1. CRISPR-associate CasB proteins bind nucleic acid non-specifically. (A) Oligomeric state of CasB proteins as revealed by SEC. (B) SDS–PAGE of the purified CasB proteins. EMSA of CasB proteins binding to (C) dsDNA and (D) ssRNA. The purified CasB proteins were incubated with dsDNA (2.8 μM) or ssRNA (0.5 μM) at 25 °C for 30 min.

aa, pI 9.68) had a rather unusually slow SEC migration behavior uncharacteristic of its molecular weight. Next, the nucleic acid binding behavior of these four proteins were compared in electrophoretic mobility shift assays (EMSA), using generic ds-DNA (291 bp) and ss-RNA (634 nt) as substrates (Fig. 1C and D). All four CasB proteins were able to bind ds-DNA and ss-RNA. The lack of distinct banding pattern suggests that these four CasB proteins are best characterized as non-specific nucleic acid binding proteins. In each case, CasB's affinity for ds-DNA appeared to be higher than ss-RNA, implying that CasB may have a preference for duplexed nucleic acid.

3.2. Crystal structures of CasB revealed a conserved positive surface patch

Next, we determined the $TfuCasB_2$ crystal structure from crystals in the hexagonal, $P6_1$, space group with unit-cell dimensions of $a = b = 87.5$ Å, $c = 69.7$ Å and one molecule occupying the asymmetric unit. The electron density of $TfuCasB_2$ allowed unambiguous tracing of from Ser11 to Arg196 (except Glu51-Pro56), whereas the non-conserved N- and C-termini (Met1-Ile10, and Pro197-Ser214) were disordered (Fig. S1). The refined $TfuCasB_2$ model contains a 10 α-helix bundle of size $45 \times 45 \times 30$ Å in dimension (Fig. 2A). α1–α2 and α5–α10 helices form a compact core and connect to the flanking α3 and α4 helices. The interior of the α-helical bundle is stabilized by numerous hydrophobic interactions (Fig. 2B). In parallel, the TthCasB structure was determined from crystals in the $P4_32_12$ space group, with unit-cell dimensions of $a = b = 104.8$ Å, $c = 83.1$ Å and two molecules in the asymmetric unit. The electron density was well defined for Met1-Ala156 residues, whereas C-terminus was disordered (Gly157-Ala169) (Fig. S1). This crystal structure was largely identical to the previously reported TthCasB structure from the $P2_1$ crystal form [8], with an r.m.s. deviation of only 0.17 Å for all Cα atoms. However, it was found that a crystal packing interface in our TthCasB structure was coincidental with the CasB dimer interface inside the Cascade complex (see below). Despite moderate sequence homology (~16%), superimposition of $TfuCasB_2$ and TthCasB revealed an identical α-helical bundle topology and a moderate r.m.s. deviation of 2.2 Å for all Cα atoms in the helical

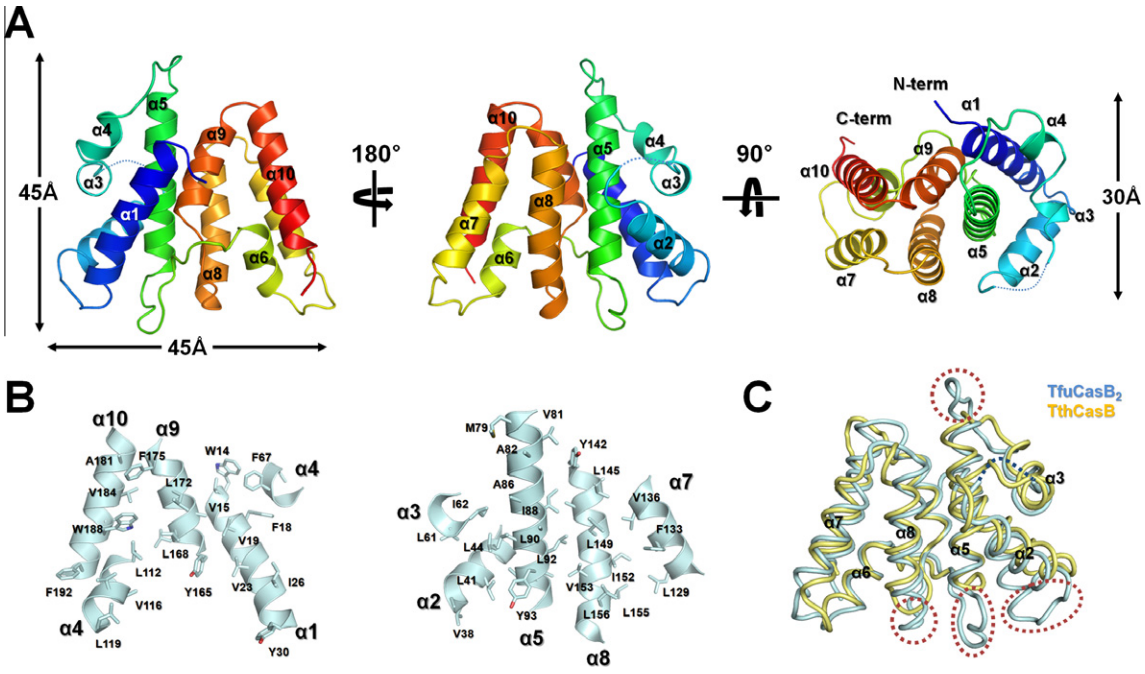

Fig. 2. Crystal structure of $TfuCasB_2$. (A) The α-helical bundle architecture in the crystal structure of $TfuCasB_2$ (residues 11–196). (B) Dissection of $TfuCasB_2$ in half to reveal that the hydrophobic core is stabilized by van der Waals contacts. (C) Structure superposition of $TfuCasB_2$ (in cyan) and TthCasB (in yellow) reveals an identical core with different structural features in the peripheral regions (Cα r.m.s deviation of 2.2 Å).

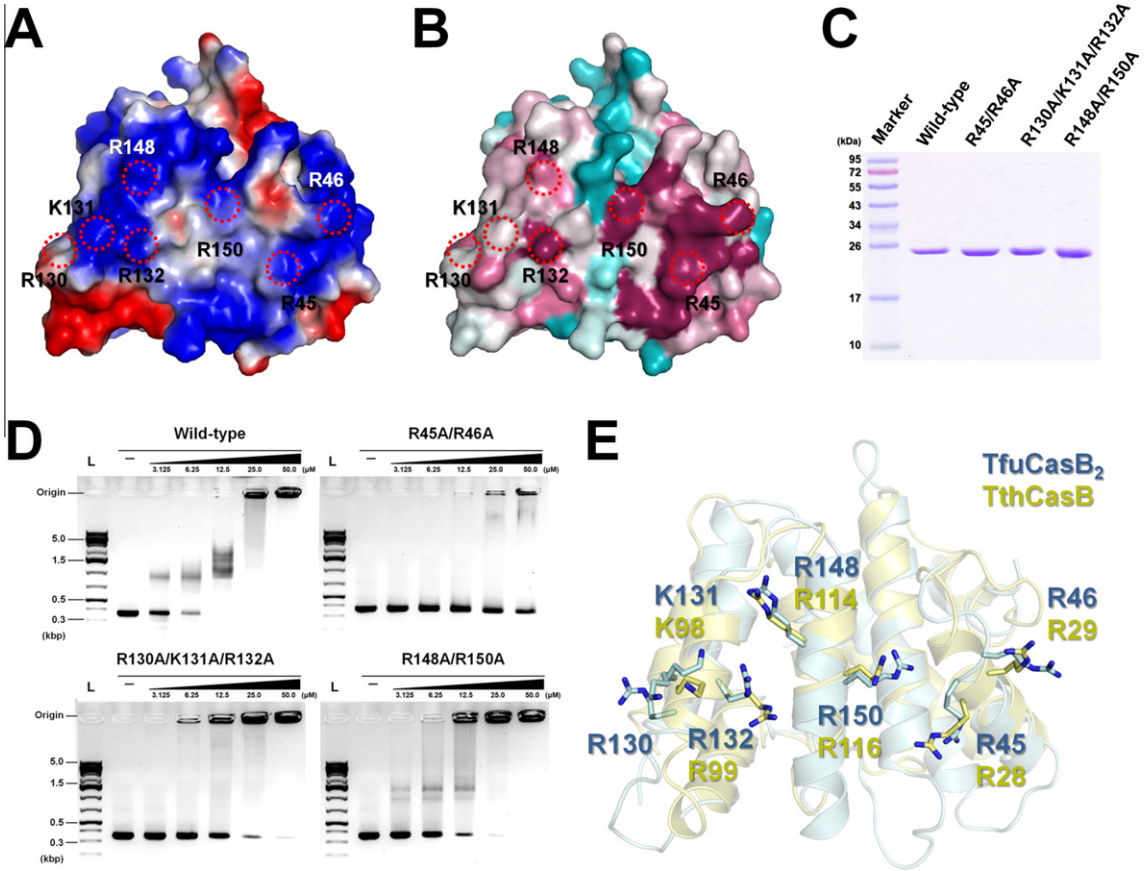

Fig. 3. A conserved positive surface patch important for nucleic acid binding in $TfuCasB_2$. (A) Electrostatic surface and (B) surface conservation in the $TfuCasB_2$. Residues are colored from magenta to cyan with descending order of conservation. (C) SDS–PAGE of the purified wild-type and mutant $TfuCasB_2$ proteins. (D) EMSA for the binding of wild-type and mutant $TfuCasB_2$ proteins to the dsDNA in Fig. 1D. A 2-fold titration of CasB proteins from 3.125 to 50 μM were incubated with 1.1 μM ds-DNA at 25 °C for 30 min. (E) Superimposition of $TfuCasB_2$ (cyan) with TthCasB (yellow) showing that locations of the positively charged residues on the putative nucleic acid binding surface are highly conserved.

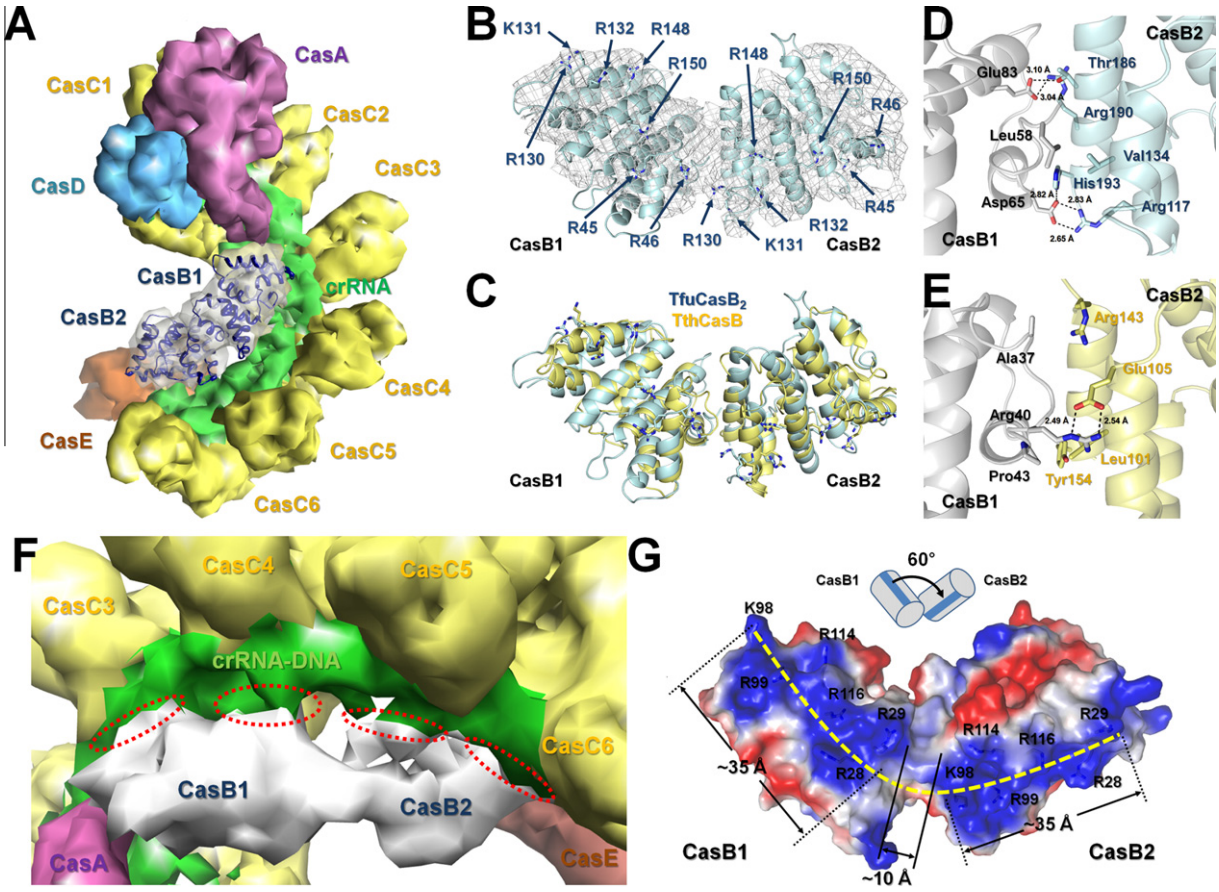

Fig. 4. Asymmetric CasB dimer formation from cryo-EM docking and crystal packing analysis. (A) Overall architecture of the *E. coli* Cascade complex (EMDB code 3514). CasA to E are colored in magenta, grey, yellow, cyan, and orange, respectively. The crRNA–ssDNA duplex is in green. An asymmetric CasB dimer is positioned along the helical spine composed of crRNA and six copies of CasC, connecting the head (CasA) and tail (CasE) of the *E. coli* Cascade. (B) Close-up view of two $TfuCasB_2$ molecules docked into the *E. coli* CasB envelope. (C) Superimposition of the EM-docked $TfuCasB_2$ dimer with a TthCasB dimer formed through crystal contacts. The dimer interface in (D) $TfuCasB_2$ and (E) TthCasB involves hydrogen bonds and salt-bridges, but residues involved are not conserved. (F) Close-up view of the cryo-EM density in *E. coli* Cascade (EMDB code 3515), showing that the CasB dimer makes multiple contacts (red-dotted circled) over a long distance on the crRNA/ss-DNA duplex. (G) Electrostatistic surface representation of the EM-docked TthCasB dimer showing that the nucleic acid binding surface of each CasB monomer aligns into an extended, curved binding site upon dimer formation in Cascade.

bundle part of the proteins (Fig. 2C). Variations were mostly found in the size and conformation of the distal loops (Fig. 2C).

3.3. Identification of a nucleic acid binding site of CasB

Comparison of the surface electrostatic potential of the $TfuCasB_2$ and TthCasB structures allowed the identification of a conserved, elongated basic patch of approximately 35 Å in length on the surface of the α-helical bundle (Fig. 3A and Fig. S2). This surface involves residues in α2- (Arg45 and Arg46, following the $TfuCasB_2$ numbering), α7- (Arg130, Lys131 and Arg132) and α8-helices (Arg148 and Arg150). Among them, Arg45 and Arg46 are especially conserved among the CasB family (Fig. 3B and Fig. S2). The equivalent residues in TthCasB are R28, R29, K98, R99, R114 and R116 (Fig. 3E). To investigate whether these residues may be important for the non-specific nucleic acid binding function of CasB, an alanine mutagenesis scan was carried out, which included R45A/R46A (mutant 1), R130A/K131A/R132A (mutant 2) and R148A/R150A (mutant 3). The ds-DNA binding affinity of $TfuCasB_2$ mutants 1, 2 and 3 were reduced by approximately 32-fold (relative activity: ~3.13%), 8-fold (~13%) and 16-fold (~6.25%), respectively, as compared to the wild-type protein (Fig. 3C and D). Taken together, mutagenesis validated the electrostatic surface potential analysis in suggesting that the positive surface on $TfuCasB_2$ contributes to the non-specific binding of nucleic acids, presumably through contacts with the sugar-phosphate backbone.

3.4. Docking of CasB into the E. coli Cascade EM map

To better understand the molecular function of CasB inside the Cascade complex, the coordinates of $TfuCasB_2$ were docked into the EM density map of the *E. coli* Cascade (Fig. 4A and B) [5] where an asymmetric CasB dimer was located along the Cascade spine formed by the helical oligomerization of six CasC proteins and the crRNA. Since rigid-body conformational changes have been detected in the CasB dimer when Cascade binds the target ss-DNA [5], it is expected that the dimer interface within the CasB dimer may not be very extensive.

No information is available about the detailed molecular interactions that lead to the asymmetric CasB dimer formation in the Cascade complex. Interestingly, we found that the orientation and packing of the $TfuCasB_2$ dimer as the result of EM docking is highly similar to a crystal packing interaction in our TthCasB structure, with an r.m.s deviation of 2.57 Å (Fig. 4C). This finding allowed an opportunity to approximate the CasB dimerization interactions in the Cascade complex. In the crystal structure, the buried surface between two CasB molecules of TthCasB was approximately 335 $Å^2$. Surprisingly, the amino acid conservation

at this interface is not significant. Distinct hydrogen bond and salt-bridge interactions were observed in the TthCasB and $TfuCasB_2$ dimer interface (Fig. 4D and E). Lack of extensive contacts and poor conservation at the dimer interface suggest that the CasB dimer formation may be transient in solution, and is stabilized upon crRNA-guided Cascade complex formation.

An interesting observation from the EM docking was that the nucleic acid binding surface in each CasB molecule became aligned to form a much more extensive, curved interface in the CasB dimer (Fig. 4F and G). The positive patch in each TthCasB monomer measured approximately 35 Å in length. Patches were separated by ~10 Å in successive monomers with an ~ 60° kink (Fig. 4G). The curvature of this continuous positive patch is complementary to the helical CasC–crRNA spine in the Cascade complex (Fig. 4G). Although, little is known about the exact function of CasB inside Cascade, it is imaginable that this extensive nucleic acid binding surface, which spans more than two turns of nucleic acid duplex, could be used to provide non-specific binding affinity for Cascade to scan through ds-DNA, and/or assisting the unwinding of the ds-DNA substrate to form the R-loop intermediate.

4. Discussion

In the Type I-E CRISPR-Cas system, five subtype-specific Cas proteins (CasA-E) form the Cascade complex to recognize and unwind ds-DNA target, prompting duplex formation between crRNA and the target strand ss-DNA while looping out the non-target DNA strand, forming an R-loop intermediate for subsequent DNA degradation by the Cas3 protein [4]. An asymmetric CasB dimer spans the crRNA-binding groove, connecting the 'head'(CasE) and 'tail'(CasA) of the seahorse-shaped Cascade.

Interestingly, SEC characterization of four CasB proteins shows that CasB is not necessarily a dimer outside the Cascade complex (Fig. 1). The published TthCasB [8] and our $TfuCasB_2$ crystal structures did not reveal any dimer formation. Serendipitously, the crystal packing interaction in our TthCasB structure, in an alternative space group, agrees well with the dimer interface inside the *E. coli* Cascade. This small and non-conserved CasB dimer interface would necessarily require additional contacts from nearby Cascade proteins and RNA components for stability.

Based on the structural analysis of $TfuCasB_2$ and structure-guided mutagenesis, we show that a conserved positive patch contributes to the non-specific nucleic acid binding function of CasB. EM docking further reveals that CasB dimer formation inside the *E. coli* Cascade leads to an elongated positive patch spanning a distance of more than 70 Å along the crRNA-binding "spine" of Cascade. This non-specific nucleic acid binding surface at the specific place in Cascade suggests that CasB may play one of the following functions in Cascade-assisted CRISPR interference: (1) assisting the recruiting and scanning of ds-DNA; (2) stabilizing the crRNA/target ss-DNA duplex in the R-loop structure; or (3) interacting with the non-complementary DNA strand to prevent the collapse of the R-loop structure (Fig. S3). In addition to the Type I-E Cascade, Cascade-like complexes have recently been reported in Type I-F (*Pseudomonas aeruginosa*) and I-C (*Bacillus halodurans*) CRISPR-Cas systems [19,20]. While the biochemical behavior and function of these ribonucleoprotein complexes are likely to be similar, the subunit composition and sequence conservation are quite different. In particular, CasB homologs have not been identified in the newly reported Cascade-like complexes, therefore, the equivalent function of CasB has to be carried out by different subunit/domains in Type I-F and I-C systems.

Acknowledgments

We thank the staff at MacCHESS for assistance in data collection, and Jason C. Grigg for helpful discussions. This work was supported, in whole or in part, by the National Institutes of Health Grant GM-086766 (to A.K.), and by the National Research Foundation of Korea Grant NRF-2010-357-C00106 from the Korean Ministry of Education, Science and Technology (to K.H.N.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2012.09.041.

References

[1] Terns, M.P. and Terns, R.M. (2011) CRISPR-based adaptive immune systems. Curr. Opin. Microbiol. 14, 321–327.
[2] Wiedenheft, B., Sternberg, S.H. and Doudna, J.A. (2012) RNA-guided genetic silencing systems in bacteria and archaea. Nature 482, 331–338.
[3] Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J., Wolf, Y.I., Yakunin, A.F., van der Oost, J. and Koonin, E.V. (2011) Evolution and classification of the CRISPR-Cas systems. Nat. Rev. Microbiol. 9, 467–477.
[4] Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V. and van der Oost, J. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 321, 960–964.
[5] Wiedenheft, B., Lander, G.C., Zhou, K., Jore, M.M., Brouns, S.J., van der Oost, J., Doudna, J.A. and Nogales, E. (2011) Structures of the RNA-guided surveillance complex from a bacterial immune system. Nature 477, 486–489.
[6] Westra, E.R., van Erp, P.B., Kunne, T., Wong, S.P., Staals, R.H., Seegers, C.L., Bollen, S., Jore, M.M., Semenova, E., Severinov, K., de Vos, W.M., Dame, R.T., de Vries, R., Brouns, S.J. and van der Oost, J. (2012) CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. Mol. Cell. 46, 595–605.
[7] Jore, M.M., Lundgren, M., van Duijn, E., Bultema, J.B., Westra, E.R., Waghmare, S.P., Wiedenheft, B., Pul, U., Wurm, R., Wagner, R., Beijer, M.R., Barendregt, A., Zhou, K., Snijders, A.P., Dickman, M.J., Doudna, J.A., Boekema, E.J., Heck, A.J., van der Oost, J. and Brouns, S.J. (2011) Structural basis for CRISPR RNA-guided DNA recognition by Cascade. Nat. Struct. Mol. Biol. 18, 529–536.
[8] Agari, Y., Yokoyama, S., Kuramitsu, S. and Shinkai, A. (2008) X-ray crystal structure of a CRISPR-associated protein, Cse2, from *Thermus thermophilus* HB8. Proteins 73, 1063–1067.
[9] Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. Macromol. Crystallogr. Part A 276, 307–326.
[10] Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. J. Appl. Crystallogr. 30, 1022–1025.
[11] Winn, M.D., Murshudov, G.N. and Papiz, M.Z. (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. Methods Enzymol. 374, 300–321.
[12] Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D: Biol. Crystallogr. 60, 2126–2132.
[13] Chen, V.B., Arendall IIIrd., 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.
[14] Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF chimera – a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.
[15] Birmanns, S., Rusu, M. and Wriggers, W. (2011) Using sculptor and situs for simultaneous assembly of atomic components into low-resolution shapes. J. Struct. Biol. 173, 428–435.
[16] Gouet, P., Courcelle, E., Stuart, D.I. and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15, 305–308.
[17] Ashkenazy, H., Erez, E., Martz, E., Pupko, T. and Ben-Tal, N. (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res. 38, W529–W533.
[18] Krissinel, E. and Henrick, K. (2005) Detection of protein assemblies in crystals. LNCS 3695, 163–174.
[19] Wiedenheft, B., van Duijn, E., Bultema, J., Waghmare, S., Zhou, K.H., Barendregt, A., Westphal, W., Heck, A., Boekema, E., Dickman, M. and Doudna, J.A. (2011) RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. Proc. Natl. Acad. Sci. USA 108, 10092–10097.
[20] Nam, K.H., Haitjema, C., Liu, X., Ding, F., Wang, H., Delisa, M.P. and Ke, A. (2012) Cas5d protein processes pre-crRNA and assembles into a Cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system. Structure 20, 1574–1584.