Clustered regularly interspaced short palindromic repeats (CRISPR) and their associated protein genes (cas genes) are widespread in bacteria and archaea. They form a line of RNA-based immunity to eradicate invading bacteriophages and malicious plasmids. A key molecular event during this process is the acquisition of new spacers into the CRISPR loci to guide the selective degradation of the matching foreign genetic elements. Csn2 is a Nmeni subtype-specific cas gene required for new spacer acquisition. Here we characterize the Enterococcus faecalis Csn2 protein as a double-stranded (ds-) DNA-binding protein and report its 2.7 Å tetrameric ring structure. The inner circle of the Csn2 tetrameric ring is ∼26 Å wide and populated with conserved lysine residues poised for nonspecific interactions with ds-DNA. Each Csn2 protomer contains an α/β domain and an α-helical domain; significant hinge motion was observed between these two domains. Ca2+ was located at strategic positions in the oligomerization interface. We further showed that removal of Ca2+ ions altered the oligomerization state of Csn2, which in turn severely decreased its affinity for ds-DNA. In summary, our results provided the first insight into the function of the Csn2 protein in CRISPR adaptation by revealing that it is a ds-DNA-binding protein functioning at the quaternary structure level and regulated by Ca2+ ions.

Clustered regularly interspaced short palindromic repeats (CRISPR)2 drives the adaptation to harmful invading nucleic acids, such as conjugative plasmids, transposable elements, and phages, using an RNA-mediated defense mechanism with fundamental similarities to our innate and adaptive immune responses (1–7). Although the details of this defense mechanism remain to be determined, two distinct stages have been recognized: (i) adaptation upon first exposure to the foreign nucleic acid, whereby some combination of CRISPR-associated (Cas) proteins extracts recognizable features from the genomes of viruses (bacteriophages) and plasmids as proto-spacers that are subsequently incorporated as spacers at the 5’ end of the CRISPR loci; and (ii) interference upon re-exposure to the same nucleic acid whereby a ribonucleoprotein complex comprised of small guide RNAs derived from the genomic CRISPRs (crRNAs) and different Cas proteins targets foreign nucleic acids for destruction (8–16). CRISPR-Cas defense systems have been identified in 83% of archaeal genomes and 45% of bacterial genomes thus far sequenced, including important human pathogens such as Campylobacter jejuni, Clostridium botulinum, Escherichia coli, Listeria monocytogenes, Mycobacterium tuberculosis, and Yersinia pestis (8, 17, 18). The significance of this pathway for human health is best illustrated in the human pathogen Staphylococcus epidermidis, where horizontal gene transfer through conjugation and plasmid transformation is prevented by CRISPR-Cas (19).

Despite strong interest in understanding the CRISPR adaptation process, its detailed molecular mechanisms remain to be elucidated. It was shown that new spacers are integrated at the 5’-end (leader end) of the CRISPR cluster (10, 20, 21). Coupled with a new integration event, loss of repeats elsewhere has been frequently observed, suggesting the occurrence of spontaneous recombination (2, 3, 22). Two of the most conserved core cas genes, cas1 and cas2, have been implicated in the new spacer acquisition process (2, 5). Cas1 has been predicted to act as an integrase in new spacer acquisition (9, 11). Recently, Pseudomonas aeruginosa Cas1 protein has been characterized as a metal-dependent double-stranded (ds-) DNA endonuclease (24), whereas E. coli Cas1 possesses nuclease activity against single-stranded (ss-) and branched DNAs (25). Cas2 genes could be further divided into subgroups in different CRISPR subtypes. Sulfolobus solfataricus Cas2 protein was characterized as a metal-dependent endoribonuclease with sequence preference for U-rich ss-RNA (26). Bacillus halodurans Cas2, however, contains metal-dependent ds-DNA endonuclease
activity in our hands. Although the nuclease activities of Cas1 and Cas2 could be involved in the new spacer integration in the CRISPR adaptation stage, a convincing biochemical reconstitution of this process has not been demonstrated (3). A less conserved core cas gene, cas4, bears resemblance to the RecB family of exonucleases and was suggested to play a role in new spacer acquisition (8, 9).

Genetic screens further identified subtype-specific cas genes involved in new spacer acquisition. For example, in Streptococcus thermophilus, a Nmeni Cas subtype organism, the cas operon contains only two additional cas genes (csn1 and csn2) besides cas1 and cas2. Although csn1 is required for crRNA-mediated silencing, csn2 was shown by a genetic screen to be required for CRISPR adaptation (10). Although structural models and biochemical characterizations are available for the Cas1 and Cas2 proteins, little is known about the structure and function of the Csn2 protein. Here we show that the Enterococcus faecalis Csn2 binds to double-stranded (ds)-DNA and describe its 2.7 Å crystal structure. We conclude that the Csn2 protein functions at the quaternary structure level, by adopting its final shape through tetrameric ring formation. Tetramerization leads to a conserved set of lysine residues being presented toward the inner circle of the ring for interactions with the ds-DNA. The observation of tightly bound Ca$^{2+}$ ions in the Csn2 structure led to further investigations that demonstrated that Ca$^{2+}$ regulates the Csn2 function by affecting its oligomerization state and enabling DNA binding. These results provide the first insight into the role of csn2 in the CRISPR adaptation in the Nmeni subtype organisms.

EXPERIMENTAL PROCEDURES

Cloning, Expression, and Purification—Full-length csn2 gene (accession number: C7UDU4) from E. faecalis was cloned into a modified pSUMO vector fused with a His$_6$-tagged N-terminal SUMO protein. Recombinant protein was expressed from E. coli BL21 star (Novagen) cell after the cell density reached $A_{600}$ of 0.8 by the addition of 0.5 mM isopropyl-1-thio-$\beta$-d-galactopyranoside at 18 °C for 12 h. The harvested cells were resuspended in lysis buffer containing 50 mM Tris-HCl, pH 8.0, and 0.3 mM NaCl. After sonication and centrifugation, the supernatant was loaded onto a 5-ml Ni-NTA column (Qiagen) equilibrated with lysis buffer plus 2 mM 2-mercaptoethanol and eluted with the same buffer plus 300 mM imidazole. After a dialysis to remove imidazole, the N-terminal SUMO tag was cleaved by incubating with the SUMO protease and removed by passing through a second Ni-NTA column. Resulting Csn2 proteins were concentrated and further purified on a Superdex 200 column (GE Healthcare) equilibrated with sizing column buffer containing 50 mM Tris-HCl, pH 8.0, 0.2 mM NaCl, and 2 mM DTT. To remove bound nucleic acids, the Csn2 fractions were pooled and further purified on a Mono Q column in a NaCl gradient (GE Healthcare). The Csn2-bound nucleic acids were then desalted and concentrated using a Centriprep filter (molecular weight cutoff of 10,000) and cloned into the pJET blunt-end cloning vector (Fermentas).

Oligomeric State Analysis—The oligomeric state of the Csn2 protein was analyzed at 4 °C using the analytical Superdex 200 column equilibrated in the sizing column buffer. In the Ca$^{2+}$ dependence analysis, the protein and the elution buffer (10 mM Tris-HCl, pH 8.0, 200 mM NaCl, and 2 mM DTT) were supplemented with 20 mM CaCl$_2$, EDTA, or EGTA, respectively.

Analysis of the Interaction between Csn2 and ds-DNA—Co-purifying nucleic acids were separated from the Csn2 protein on the Mono Q column, concentrated, and analyzed on a 1% (w/v) agarose gel stained with ethidium bromide. To reveal their identity, DNase I (0.1 µg/ml) and RNase A(0.1 µg/ml) digestions were carried out in a buffer containing 20 mM HEPES (pH 7.5) for 30 min at 25 °C. Alkaline hydrolysis was carried out in 100 mM sodium carbonate, pH 8.8, and 2 mM EDTA at 75 °C for 5 min. The reaction products were visualized using a Gel Doc XR System (Bio-Rad).

Electrophoretic Mobile Shift Assay (EMSA)—The ds-DNA-Csn2 interaction was assayed using 100 ng of ds-DNA containing the cloned Csn2-bound nucleic acid number 3 PCR-amplified from the pJET cloning vector. Metal-dependent ds-DNA binding of Csn2 (0–160 µM) was measured in 20 mM HEPES, pH 7.5, and 50 mM NaCl and supplemented with 20 mM CaCl$_2$, EDTA, or EGTA. After incubation, the reaction mixture in the Ca$^{2+}$ condition was separated on a 6% native Tris-glycine gel, and the EDTA- and EGTA-containing samples were separated on a 6% native Tris-borate-EDTA gel to avoid incompatibility with the Tris-glycine buffer. Gels were visualized using ethidium bromide staining and analyzed using a Gel Doc XR system (Bio-Rad).

Crystallization, Data Collection, and Structure Determination—The Csn2 crystals were grown at 18 °C using the hanging drop vapor diffusion method by mixing 5 mg/ml protein at a 1:1 ratio with the well solution containing 0.1 M MES, pH 6.0, 0.1 M calcium acetate, and 6–14% (w/v) PEG 6000. Suitable crystals for x-ray diffraction grew within 5–7 days. Both the native and the selenium-methionine derivatized crystals were cryo-protected by soaking the crystals in the well solution supplemented with 30% (v/v) ethylene glycol. The native data set was collected at 100 K at the Macromolecular Diffraction at Cornell High Energy Synchrotron Source (MACCHESS) beam line A1. The selenium-methionine single wavelength anomalous dispersion (SAD) dataset was collected through the mail-in crystallography service at the NE-CAT beam line 24ID-C at the Advanced Photon Source (APS). Diffraction data sets were indexed, integrated, and scaled using HKL2000 (27). Initial sets of phases were obtained from a selenium-methionine SAD data set using the direct method in SHELXC/D/E (28). Refinement of the heavy metal sites and automated model building were carried out using the PHENIX software suite (29). Structure building and refinement were carried out using the programs COOT (30) and REFMAC (31), respectively. The final structure model was refined using the PHENIX software suite (29). Simulated annealing omit maps were systematically generated to check the quality of the model. We further checked the quality of the model using MOLPROBITY (32). The coordinates and structure factor have been deposited in the Protein Data Bank with the accession code 3SSU.
Structural Analysis—The structure-based sequence alignment was carried out using ClustalW and ESPRINT (33, 34). The three-dimensional structural similarity between Csn2 and other proteins was identified using the DALI server (35). Molecular contacts, buried surface areas, and temperature B-factor distribution were analyzed using CCP4 and CNS(36). Surface conservation within the Csn2 family of proteins was calculated and illustrated using the Consurf server (37) and Chimera (23). Figure illustrations were generated using PyMOL (38).

RESULTS

*E. faecalis* Csn2 Binds ds-DNA—*E. faecalis* Csn2 protein recombinantly expressed from *E. coli* was found to assemble into a large oligomeric state and interact with nucleic acid strongly, displaying higher absorbance at UV\textsubscript{260} rather than UV\textsubscript{280} after Ni-NTA and size exclusion chromatography (Fig. 1A). In the presence of the co-purifying nucleic acids, Csn2 migrated as a large oligomeric species on size exclusion chromatography with an estimated molecular mass of over 660 KDa. Removing the co-purifying nucleic acids reduced the average

FIGURE 1. Nucleic acid binding and oligomerization in Csn2. A, elution profile of the Csn2 protein with (peak 1) or without (peak 2) the co-purifying nucleic acids on an analytical Superdex 200 10/300 size exclusion column. Csn2 in peak 2 showed an average size of a pentamer to hexamer formation. Further changes in oligomerization state upon Ca\textsuperscript{2+} binding are shown in Fig. 5D. B, separation of the co-purifying nucleic acids from the Csn2 protein on the Mono Q column. C, analysis of the fractions in panel B using the Coomassie Blue-stained SDS-PAGE gel (upper panel) or ethidium bromide-stained agarose gel (lower panel). Fraction numbers are consistent with those shown in panel B. D, identity of the Csn2-bound nucleic acids examined by incubating with DNase I, RNase A, or the alkaline hydroxyl buffer.
molecular mass of the Csn2 to in a broad range of a pentamer or hexamer (Fig. 1A). As shown later, the oligomerization and ds-DNA binding properties of Csn2 were further regulated by the presence of Ca\(^{2+}\). The co-purifying nucleic acids could be extracted from the Csn2 protein using anion exchange chromatography (Fig. 1, B and C) and were shown to be ~100–500 bp in size. They were sensitive to DNase I digestion, but not RNase A digestion nor alkaline hydrolysis treatment that selectively degrades RNA (Fig. 1D), suggesting that the Csn2-bound nucleic acids are likely the E. coli endogenous ds-DNA, but not RNA. Consistent with this observation, we showed that the Csn2-bound nucleic acids are likely the E. coli endogenous ds-DNA, but not RNA. However, a more detailed biochemical study will be required to investigate whether certain short DNA sequences are preferentially bound by Csn2. A further study using EMSA revealed strong interactions between Csn2 and a ds-DNA substrate PCR-amplified to contain one of the sequences of the co-purifying DNA (supplemental Table S1, row 3), confirming that Csn2 interacts strongly with ds-DNA (see below). By contrast, EMSA done at similar conditions using ss-DNA, ss-RNA, and ds-RNA substrates did not show appreciable interactions with the Csn2 protein (data not shown). Analysis of these nucleic acids after Csn2 incubation on a sequencing gel did not reveal nuclease or ribonuclease activity in Csn2 either (data not shown).

Overall Structure of Csn2—The crystal structure of Csn2 from E. faecalis was solved by single SAD using the selenium-methionine derivatized protein (Table 1). The asymmetric unit of the orthorhombic P2\(_1\)2\(_1\)2\(_1\) space group contains two non-crystallographic symmetry-related tetrameric rings (supplemental Fig. S1). The two Csn2 tetramers adopt slightly different conformations with an r.m.s. deviation of 1.3 Å for all Ca atoms (supplemental Fig. S2). Each diamond-shaped tetrameric ring measures 70 × 70 Å in width and 50 Å in height (Fig. 2A). Each Csn2 protomer contains a globular α/β domain and an extended α-helical domain extruded from the middle of the α/β domain. Extensive interactions between the two α/β domains lead to Csn2 dimer formation. Further hydrophobic interactions between two such Csn2 dimers at the extended α-helical domain lead to the tetrameric ring formation (Fig. 2A). Four tightly bound Ca\(^{2+}\)-ions were found at this interface. The inner diameter of the ring measuring 26 Å at the narrowest region agrees well with accommodating the binding of a ds-DNA substrate through the center. Electrostatic analysis revealed the presence of a set of positively charged lysine residues populating the inner surface of the ring (Fig. 2B). Both features are consistent with this region being involved in the binding of ds-DNA in a sequence-nonspecific fashion. As the Csn2 family of proteins are highly conserved (Fig. 3A, supplemental Fig. 3), the observed structural features described here are likely shared by all Csn2 proteins.

### TABLE 1

| Data collection statistics | Native | SAD |
|---------------------------|--------|-----|
| Beamline                   | MACCHESS | APS |
| Wavelength                 | 0.9770 | 0.97917 |
| Space group                | P2\(_1\)2\(_1\)2\(_1\), P2\(_1\)2\(_1\)2\(_1\) | P2\(_1\)2\(_1\)2\(_1\), P2\(_1\)2\(_1\)2\(_1\) |
| Unit cell parameters (Å)   | a 104.11, b 140.09, c 148.43 | a 104.05, b 138.73, c 148.29 |
| Resolution (Å)             | 20.0-2.70 (2.75-2.70) | 30.0-3.10 (3.21-3.10) |
| Completeness (%)           | 98.8 (95.5) | 99.8 (99.6) |
| Redundancy                 | 5.5 (3.6) | 2.6 (2.5) |
| L/σ(I)                     | 25.26 (2.76) | 14.63 (1.57) |
| R\(_{merge}\) (%)           | 8.6 (29.5) | 3.9 (34.8) |

**Refinement statistics**

| Resolution (Å)             | 20.0-2.70 |
|----------------------------| 21.90/27.5 |
| B-factor (Averaged)         | 51.75 |
| Protein                    | 49.43 |
| Ca\(^{2+}\)                 | 44.35 |
| Water                      | 0.016 |
| Bond angles(°)              | 1.697 |
| Ramachandran plot(%)       | 96.3 |
| favored                    | 3.7 |
| allowed                    | |

\( R_{merge} = \sum h i \geq \sum I(h,i) - (1/I(h,i)) \sum h j \geq \sum I(h,j), \) where I(h,i) is the intensity of the standard measurement of reflection h and (1/I(h,i)) is the mean value of I(h,i) for all i measurements.

\( R = \sum || F_{o} || - || F_{c} ||/ \sum || F_{o} ||, \) where F\(_{o}\) and F\(_{c}\) are the observed and calculated structure-factor amplitudes, respectively. R\(_{merge}\) was calculated as R\(_{work}\) using a randomly selected subset (10%) of unique reflections not used for structure refinement.

Categories were defined by MOLPROBITY.
The angle between the two domains, however, varied by as much as 5° among eight Csn2 protomers in the asymmetric unit, suggesting the presence of a hinge region between these two domains (Fig. 3). Nevertheless, the two connecting loops in this hinge region between H9252 and H9251 (Thr-63–Ser-72) and also between H9251 and H9255 (Leu-133–Thr-143) display similar temperature B-factors as the rest of the structure due to the binding of Ca2+ ions to stabilize the hinge region (see below).

The α/β globular domain in Csn2 bears structural similarity to a family of ATP-binding proteins, although the sequence homology is hardly detectable (<11% sequence identity). Structural homologs include the Clostridium perfringens cobalt import ATP-binding protein, the enterobacterial phage T7 DNA primase/helicase, the Pyrococcus furiosus DNA double-strand break repair RAD20 ATPase, and the M. tuberculosis RecA (PDB accession codes: 3GFO, 1CR1, 3QKU, and 1MO4, respectively; supplemental Fig. S4). Closer structural analysis, however, suggested that ATP binding would unlikely be the native function of the Csn2 protein as neither the backbone geometry nor the key contacting residues at the ATP-binding site are conserved in Csn2. Indeed, we were not able to detect interactions between ATP and Csn2 using isothermal titration calorimetry analysis (data not shown). The distinct topology in the α/β domain and the presence of the α-helical domain insertion set the Csn2 protein apart from other ATP-binding α/β domain proteins (supplemental Fig. S4).

Two Dimerization Interfaces Lead to Csn2 Tetrameric Ring Formation—The tetrameric ring formation is best described as two sequential dimerization events, first leading to the dimer formation between molecules A-C and B-D (interface A-C; Fig. 4A) and then the dimerization in molecules A-B and C-D (interface A-B; Fig. 4B). Superimposition of each dimer (molecules A-B or A-C) yielded similar r.m.s. deviation of 0.5–1.1 Å in Cα alignment (supplemental Fig. S2). Both interfaces involve highly conserved residues among the Csn2 family of proteins (Fig. 3D; supplemental Fig. S3). The A-C interface involves symmetric interactions to the side of the α1-helix and β-sheet and the top of the α5 and α6 helices between two α/β domains, burying a surface area of ~2100 Å² (~16% of the total surface in each protomer; Fig. 4A and supplemental Fig. S5). Among the interface residues of molecule A-C, 33% are hydrophobic, 66% are hydrophilic (i.e. H-bonds between Tyr-36–Asp-64, Asp-64–Thr-171, and Asn-168–Tyr-172), and 16% are charged.

Interface A-B involves reciprocal interactions at the α2–α4 helices and hinge loop region of molecules A and B, burying a
FIGURE 3. Structural analysis of the Csn2 protomer. A, sequence alignment among the Csn2 family of proteins from E. faecalis (accession code: C7UDU4), Enterococcus faecium (D4W167), L. monocytogenes (E3YJP9), Streptococcus anginosus (E6J3Q7), S. bovis (E0PEL0), Streptococcus pyogenes serotype M1 (Q99ZV9), and S. thermophilus (Q03JI9). The absolutely conserved residues are boxed in red, and the highly conserved ones are in unfilled boxes and red letters. Residues involved in the binding of Ca²⁺ and ds-DNA (putative) are marked with red and blue asterisks, respectively. B, two views of the monomeric structure of the E. faecalis Csn2 protein. Csn2 consists of an α/β-domain and an α-helical domain. Hinge loops between the two domains are colored magenta. A disordered distal loop in the α/β-domain is represented by the magenta dotted line. The two Csn2-bound Ca²⁺ ions are represented in yellow spheres. The two β-sheets and a 3₁₀-helix inside the α/β-domain are colored in yellow and green, respectively. C, superimposition of the eight Csn2 monomers in the asymmetric unit along the α₁-helix (in blue). Hinge motion at the hinge region in magenta leads to ∼5° variation in the orientation of the α-helical domain. Large deviations in equivalent atom positions are marked. D, surface conservation in the Csn2 protomer. Residues are colored from magenta to cyan with descending order of conservation.

Ca²⁺-dependent ds-DNA Binding in Csn2
surface area of 4200 Å² (~32% of each subunit; Fig. 4B and supplemental Fig. S5). The interface residues are 51% hydrophobic, 49% hydrophilic, and 31% charged. The hydrophobic interactions include the two anti-parallel leucine/isoleucine zipper helices and the contacts from α3 to α2 and α4. Additional contacts include salt bridges between Lys-90 and Glu-114 and between Glu-116 and Arg-156.

Potential ds-DNA-binding Site Inside the Tetrameric Ring—Electrostatic surface potential analysis revealed a clear segregation of positive and negative charges on the inner and outer surfaces of the Csn2 tetrameric ring, respectively (Fig. 2B). The charge distribution is conserved among the Csn2 family of proteins (Fig. 3A) and is in line with its putative ds-DNA binding function. The lysine-rich basic patch is particularly interesting because it may mediate nonspecific interactions with the sugar-phosphate backbone of the ds-DNA (Figs. 2B and 4C). This patch is composed of at least seven highly conserved lysine residues (Lys-52, Lys-55, Lys-77, Lys-131, Lys-160, Lys-161, and Lys-162; Fig. 4D). The presence of a positive change is consistent in the place of Lys-55, Lys-77, Lys-160, Lys-161, and Lys-162, whereas glutamine is occasionally found in the place of Lys-52 and Lys-131. In the crystal structure, Lys-77 and Lys-131 are located in the α-helical domain, and the rest of the Lys residues are located in the α/β domain. Due to the presence of a hinge region between these two domains, the exact positions of the lysine residues differ among the Csn2 protomers. The versatility in lysine positions and the flexibility in their side chain conformations are consistent with this lysine-rich patch inside the tetrameric ring contacting the ds-DNA.

\[ \text{Ca}^{2+}-\text{dependent ds-DNA Binding in Csn2} \]
Ca\textsuperscript{2+} ions. EGTA chelation of Ca\textsuperscript{2+} from Csn2 crystals prior to data collection resulted in severe degradation of x-ray diffraction resolution (data not shown). Incubation of equimolar amounts of Mn\textsuperscript{2+} and Ca\textsuperscript{2+} with Csn2 crystals, followed by data collection at the anomalous edge of Mn\textsuperscript{2+}, did not show strong anomalous difference peaks indicative of competitive binding of Mn\textsuperscript{2+} to the Ca\textsuperscript{2+} sites (data not shown). These two circumstantial evidences further support the presence of Ca\textsuperscript{2+} in the Csn2 structure.

Both Ca\textsuperscript{2+}-binding sites are located at the interface A-B and mediate interactions between Csn2 protomers by shielding the charge repulsion among coordinating functional groups. The two Ca1 sites are located in the middle of the interface A-B near the α4-helix. Ca\textsuperscript{2+} is octahedrally coordinated by O\textsubscript{6} of Asp-122 (average distance of 2.42 Å), the main chain carbonyl of Gly-123 (2.26 Å), O\textsubscript{e} of Glu-128 (2.56 Å) in molecule A, the main chain carbonyl of Ala-132 (2.42 Å) in molecule B (Fig. 5B), as well as two water (2.42 Å) molecules. The two Ca2 sites located at the hinge region near the N-terminal end of the α5-helix are square pyramidally coordinated by the O\textsubscript{6} of Asp-118 (average distance of 2.42 Å) in molecule A and by the O\textsubscript{e} and the main chain carbonyl of Glu-138 (2.40 and 2.52 Å, respectively), O\textsubscript{6} of Asp-142 (2.39 Å) and the O\textsubscript{e} of Glu-150 (2.37 Å) in molecule B (Fig. 5C). The key Ca\textsuperscript{2+} coordinating residues including Asp-118, Asp-122, Glu-128, Glu-138, and Glu-150 are highly conserved (Fig. 3A and supplemental Fig. S3), suggesting that Ca\textsuperscript{2+} binding is a conserved feature among Csn2 proteins. The other two coordinating residues Gly-123 and Ala-132 can be substituted by other residues among the Csn2 family because only the main chain carbonyl groups are involved in coordination.

**Ca\textsuperscript{2+} Influences the Oligomerization and ds-DNA Binding Properties of Csn2**—Because our crystal structure revealed that the Ca1 and Ca2 sites were strategically positioned at the oligo-
merization interface A-B, we speculated that Ca\(^{2+}\) may critically influence the oligomerization state and ds-DNA binding property of the \textit{E. faecalis} Csn2 protein. Size exclusion chromatography was carried out to study the effect of Ca\(^{2+}\) binding on the oligomerization state of Csn2. As shown previously, purified \textit{E. faecalis} Csn2 protein migrated as a mixture of oligomers with an average size of a pentamer or hexamer (Figs. 1A and 5D). The addition of EDTA or EGTA to further remove the co-purifying Ca\(^{2+}\) from the Csn2 protein encouraged the formation of higher molecular weight oligomers, whereas incubating with 20 mM Ca\(^{2+}\) promoted the formation of a Csn2 tetramer. Therefore, structural and biochemical evidence converged in, suggesting that Ca\(^{2+}\) plays an important role in promoting the formation of a Csn2 tetrameric ring.

Next, we studied whether Ca\(^{2+}\) binding may influence the ds-DNA binding property of the Csn2 protein. EMSAs were carried out in the presence of Ca\(^{2+}\), EDTA, EGTA, or Mg\(^{2+}\) using a ds-DNA substrate PCR-amplified from the cloned Csn2-co-purifying DNA (Fig. 1). The Csn2 protein was found to interact strongly with this DNA only in the presence of Ca\(^{2+}\) ion, but not in the presence of EDTA or EGTA (Fig. 5E). Interestingly, Mg\(^{2+}\) was not able to restore the ds-DNA binding by Csn2 either, suggesting again that the Csn2 interacts specifically with the Ca\(^{2+}\) ion. Together with the Ca\(^{2+}\)-dependent oligomerization study above, these results suggest that Ca\(^{2+}\) binding promotes tetramerization and that the tetrameric ring gives rise to the ds-DNA binding property in the Csn2 protein.

**DISCUSSION**

Nmeni CRISPR-Cas subtype provides us a unique opportunity to study CRISPR adaptation because a genetic screen in \textit{S. thermophilus} has identified the key players required for new spacer acquisition (10). The identified proteins included the core \textit{cas} genes \textit{cas1} and \textit{cas2} and a Nmeni subtype-specific gene \textit{csn2}. Although structural models and biochemical data are available for Cas1 and Cas2 proteins, this study provides the first set of such data for the Csn2 protein. Our crystal structure clearly reveals that the conserved family of Csn2 proteins functions at the quaternary structure level; that is, Csn2 assumes its ultimate shape and charge distribution only after tetrameric concatenations at the quaternary structure level; that is, Csn2 assumes its first set of such data for the Csn2 protein. Our crystal structure are available for Cas1 and Cas2 proteins, this study provides the gene

---

**REFERENCES**

1. Sorek, R., Kunin, V., and Hugenholtz, P. (2008) \textit{Nat. Rev. Microbiol} 6, 181–186
2. van der Oost, J., Jore, M. M., Westra, E. R., Lundgren, M., and Brouns, S. J. (2009) \textit{Trends Biochem. Sci.} 34, 401–407
3. Marraffini, L. A., and Sontheimer, E. J. (2010) \textit{Nat. Rev. Genet.} 11, 181–190
4. Karginov, F. V., and Hannon, G. J. (2010) \textit{Mol. Cell} 37, 7–19
5. Deveau, H., Garneau, J. E., and Moineau, S. (2010) \textit{Annu. Rev. Microbiol.} 64, 475–493
6. Horvath, P., and Barrangou, R. (2010) \textit{Science} 327, 167–170
7. Waters, L. S., and Storz, G. (2009) \textit{Cell} 136, 615–628
8. Haft, D. H., Selengut, J., Morgonid, E. F., and Nelson, K. E. (2005) \textit{PLoS Comput. Biol.} 1, e60
9. Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., and Koonin, V. E. (2006) \textit{Biol. Direct} 1, 7
10. Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., and Horvath, P. (2007) \textit{Science} 315, 1709–1712
11. 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) \textit{Science} 321, 960–964
12. Carte, J., Wang, R., Li, H., Terns, R. M., and Terns, M. P. (2008) \textit{Genes Dev.} 22, 3489–3496
13. Hale, C. R., Zhao, P., Olson, S., Duff, M. O., Graveley, B. R., Wells, L.
Ca\textsuperscript{2+}-dependent ds-DNA Binding in Csn2

Terns, R. M., and Terns, M. P. (2009) Cell 139, 945–956

14. Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K., and Doudna, J. A. (2010) Science 329, 1355–1358

15. Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., and Charpentier, E. (2011) Nature 471, 602–607

16. Perez-Rodriguez, R., Haitjema, C., Huang, Q., Nam, K. H., Bernardis, S., Ke, A., and DeLisa, M. P. (2011) Mol. Microbiol. 79, 584–599

17. Grissa, I., Vergnaud, G., and Pourcel, C. (2007) Nucleic Acids Res. 35, W52–57

18. Rousseau, C., Gonnet, M., Le Romancer, M., and Nicolas, J. (2009) Bioinformatics 25, 3317–3318

19. Marraffini, L. A., and Sontheimer, E. J. (2008) Science 322, 1843–1845

20. Lillestøl, R. K., Redder, P., Garrett, R. A., and Brügger, K. (2006) Archaea 2, 59–72

21. van der Ploeg, J. R. (2009) Microbiology 155, 1966–1976

22. Tyson, G. W., and Banfield, J. F. (2008) Environ. Microbiol. 10, 200–207

23. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) J. Comput. Chem. 25, 1605–1612

24. Wiedenheft, B., Zhou, K., Jinek, M., Coyle, S. M., Ma, W., and Doudna, J. A. (2009) Structure 17, 904–912

25. Babu, M., Beloglazova, N., Flick, R., Graham, C., Skarina, T., Nocek, B., Gagarinova, A., Pogoutse, O., Brown, G., Binkowski, A., Phanse, S., Joachimiak, A., Koonin, E. V., Savchenko, A., Emili, A., Greenblatt, J., Edwards, A. M., and Yakunin, A. F. (2011) Mol. Microbiol. 79, 484–502

26. Beloglazova, N., Brown, G., Zimmerman, M. D., Proudfoot, M., Makarova, K. S., Kudritska, M., Kochinyan, S., Wang, S., Chruszcz, M., Minor, W., Koonin, E. V., Edwards, A. M., Savchenko, A., and Yakunin, A. F. (2008) J. Biol. Chem. 283, 20361–20371

27. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326

28. Sheldrick, G. M. (2010) Acta Crystallogr. D Biol. Crystallogr. 66, 479–485

29. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) Acta Crystallogr. D Biol. Crystallogr. 66, 213–221

30. Emsley, P., and Cowtan, K. (2004) Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132

31. Winn, M. D., Murshudov, G. N., and Papi, M. Z. (2003) Methods Enzymol. 374, 300–321

32. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) Acta Crystallogr. D Biol. Crystallogr. 66, 12–21

33. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Bioinformatics 23, 2947–2948

34. Gouet, P., Courcelle, E., Stuart, D. I., and Métoz, F. (1999) Bioinformatics 15, 305–308

35. Holm, L., and Rosenström, P. (2010) Nucleic Acids Res. 38, W545–W549

36. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. D Biol. Crystallogr. 54, 905–921

37. Askhenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010) Nucleic Acids Res. 38, W529–533

38. DeLano, W. L. (2010) The PyMOL Molecular Graphics System, version 1.3r1, Schrödinger, LLC, New York