Published online 26 April 2013 *Nucleic Acids Research, 2013, Vol. 41, No. 12* ***6347–6359***
doi:10.1093/nar/gkt315

Double-strand DNA end-binding and sliding of the toroidal CRISPR-associated protein Csn2

Zihni Arslan[1], Reinhild Wurm[1], Oleksandr Brener[1,2], Philipp Ellinger[3], Luitgard Nagel-Steger[1,2], Filipp Oesterhelt[1], Lutz Schmitt[3], Dieter Willbold[1,2], Rolf Wagner[1], Holger Gohlke[4], Sander H. J. Smits[3] and Ümit Pul[1,*]

[1]Institut für Physikalische Biologie, Heinrich-Heine-Universität, Universitätsstr. 1, 40225 Düsseldorf, Germany, [2]ICS-6, Forschungszentrum Jülich, 52425 Jülich, Germany, [3]Institut für Biochemie, Heinrich-Heine-Universität, Universitätsstr. 1, 40225 Düsseldorf, Germany and [4]Institut für Pharmazeutische und Medizinische Chemie, Heinrich-Heine-Universität, Universitätsstr. 1, 40225 Düsseldorf, Germany

Received February 13, 2013; Revised April 4, 2013; Accepted April 5, 2013

ABSTRACT

The adaptive immunity of bacteria against foreign nucleic acids, mediated by CRISPR (clustered regularly interspaced short palindromic repeats), relies on the specific incorporation of short pieces of the invading foreign DNA into a special genomic locus, termed CRISPR array. The stored sequences (spacers) are subsequently used in the form of small RNAs (crRNAs) to interfere with the target nucleic acid. We explored the DNA-binding mechanism of the immunization protein Csn2 from the human pathogen *Streptococcus agalactiae* using different biochemical techniques, atomic force microscopic imaging and molecular dynamics simulations. The results demonstrate that the ring-shaped Csn2 tetramer binds DNA ends through its central hole and slides inward, likely by a screw motion along the helical path of the enclosed DNA. The presented data indicate an accessory function of Csn2 during integration of exogenous DNA by end-joining.

INTRODUCTION

Clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) proteins constitute a prokaryotic adaptive immunity system to prevent viral infections or the invasion by mobile DNA elements (1–3). The immunization is achieved by the specific storage of foreign nucleic acid sequences as 'portraits' of the targets, which are later used in the form of small RNAs for the specific recognition and inactivation of the invading DNA. Ten CRISPR-Cas systems have been described, which differ in the organization of the CRISPR operon, constituted by a set of diverse CRISPR-associated genes (*cas*) and the CRISPR arrays (4). The *cas* genes encode the protein components of the system, which exhibit various biochemical activities (5).

The mechanism of the CRISPR-mediated immunity is divided into three stages (2,6). The first stage describes the immunization of the host cells by the integration of foreign DNA-derived spacer sequences into the CRISPR array (adaptation stage). The second stage comprises the transcription of the CRISPR array to the precursor CRISPR RNA (pre-crRNA), the expression of Cas proteins and the processing of the pre-crRNA to small crRNAs (expression/processing stage). In the third stage, the crRNA-loaded Cas protein complexes screen the foreign DNA for spacer-matching sequences and initiate the nucleolytic hydrolysis of the target DNA (interference stage). According to the recent classification of the different CRISPR-Cas systems, three major types (type I, II and III) have been defined (4), which exhibit mechanistic variability in the maturation of the crRNAs and the inactivation of the target nucleic acids (7–15).

The type II CRISPR-Cas systems, mainly found in pathogenic bacteria (16), consist of the CRISPR array, the *cas* genes encoding for the proteins Cas9, Cas1, Cas2 and Csn2 (in type II-A) or Cas4 (in type II-B) and a gene for the trans-encoded crRNA (tracrRNA). The tracrRNA contains sequence elements complementary to the repeat of the pre-crRNA (11). The RNA duplex formation of the tracrRNA and the pre-crRNA mediates the maturation of the crRNAs by RNaseIII and the Cas9 protein (11,17). Loaded with the tracrRNA:crRNA duplex, the Cas9 protein introduces double-strand DNA breaks at the spacer-matching protospacer region (12). Several studies have shown that the Cas9 protein along with designed crRNA:tracrRNA is active in eukaryotic cells and can be used for multiplex genome editing in bacteria and eukaryotes (18–23). Moreover, engineered Cas9 protein with inactivated nuclease domains has been used for sequence-specific control of transcription in *Escherichia coli* and eukaryotic cells (24).

*To whom correspondence should be addressed. Tel: +49 211 81 14945; Fax: +49 211 81 15167; Email: pul@hhu.de

The activities of the proteins Cas1, Cas2 and Csn2 are dispensable for maturation of the crRNAs and the cleavage of the target DNA, and thus are likely required in the immunization stage (10,11). The nucleases Cas1 and Cas2 are involved in incorporation of new spacer sequences into the CRISPR array (25–27). Little is known about the function of the Csn2 protein; although early studies showed that a knockout of the *csn2* gene prevents the development of phage-insensitive mutants by an uptake of new spacer sequences (7,28), the role of the Csn2 protein during immunization is unknown. The structures and DNA-binding activities of Csn2 proteins from *Enterococcus faecalis*, *S. agalactiae* and *Streptococcus pyogenes* have been analyzed independently (29–31). These Csn2 proteins are ring-shaped homotetramers, and the structure of a protomer consists of a 'head' and a 'leg' domain. Whereas the leg domain consists of a pair of long stretched helices that form the interface for the tetrameric assembly, the function of the head domain is still unclear. More recently, based on the crystal structure of the Cas protein Stu0660 from *Streptococcus thermophilus*, a second Csn2 protein family has been proposed, which shares no sequence similarity to Csn2 proteins; however, its members have a similar tetrameric ring-shaped structure (32). All Csn2 proteins exhibit Ca^{2+}-dependent double-stranded DNA (dsDNA)-binding activities without any notable sequence dependence. The tetrameric structure contains a large positively charged central hole of ~4 nm in diameter, which is stabilized by Ca^{2+}-ions bound to the helices in the leg domain. This hole is large enough to encircle the dsDNA. An alternative DNA binding mode has been suggested by molecular dynamics simulations, which revealed that the Csn2 tetramers can adopt a conformation with a groove outside the ring that could interact with the dsDNA (30).

To elucidate the role of the Csn2 protein in the CRISPR immunization stage, we have analyzed the DNA-binding mechanism of Csn2 from the human pathogen *S. agalactiae* in detail, using different biochemical techniques, atomic force microscopy (AFM) and molecular dynamics (MD) simulations. We provide evidence that Csn2 binds to the ends of linear dsDNA and moves inward through rotation-coupled translocation. The encircling of the dsDNA by Csn2 tetramers was probed by capping the DNA ends of Csn2–DNA complexes with streptavidin, keeping the Csn2 rings arrested on the DNA. The MD simulations give mechanistic insights at the molecular level how a Csn2 ring may rotate along the helical path of the DNA in a Ca^{2+}-dependent manner. The data suggest that Csn2 is a DNA end-loading toroidal protein, whose function is directly related to DNA end metabolism. Csn2 thus resembles proteins involved in the repair of dsDNA breaks and recombination.

MATERIALS AND METHODS

Purification of Csn2

Streptococcus agalactiae Csn2 protein was expressed and purified as previously described (30).

Preparation of relaxed plasmid DNA for competition experiments

Relaxation of supercoiled plasmid DNA was performed with *E. coli* Topoisomerase I (New England Biolabs, NEB). Supercoiled pUC18-1 plasmid DNA (33) was extracted and purified from *E. coli* cells grown to stationary phase in YT medium. Ten micrograms of supercoiled plasmid DNA were incubated in 1 × NEBuffer 4 (50 mM potassium acetate, 20 mM Tris acetate, pH 7.9, 10 mM magnesium acetate, 1 mM DTT) and 1 μg ml^{-1} bovine serum albumin in a total volume of 30 μl. Five units *E. coli* Topoismerase I were added and the reaction mixture was incubated for 1 h at 37°C to allow complete relaxation. After extraction with phenol/chloroform and ethanol precipitation, the DNA was resuspended in deionized water (Millipore). The relaxation was verified on a 1% agarose gel. DNA bands were stained with ethidium bromide and visualized under UV light.

Analyses of protein–DNA interaction by electrophoretic mobility shift assays

The 155 bp DNA fragment was obtained by *Eco*RI-*Dra*I digestion of the plasmid pUC18-1. DNA fragments were purified by agarose gel electrophoresis and end labeled by Klenow (Promega) incorporation of [α-^{32}P]-dATP. Binding reaction mixtures contained 2 nM DNA, 10 mM $CaCl_2$, 20 ng μl^{-1} heparin and indicated concentrations of Csn2 or unlabeled plasmid DNA (supercoiled, relaxed or linearized) as given in the individual experiments. After incubation for 15 min at room temperature, the complexes were separated from free DNA on native 5% (w/v) Tris/glycine polyacrylamide gels. DNA bands were visualized by autoradiography.

Preparation of DNA for AFM

The pCR001 plasmid (5125 bp) used in the AFM analyses was constructed by ligation of *cas1-cas2* genes into the *Nde*I-*Pac*I cleaved pACYCDuet-1 vector (Novagen). The insert with *cas1-cas2* genes were obtained by PCR from genomic DNA of *E. coli* MG1655 using the following primers: forward 5′-GAATGCCATATGACCTGGC TTCCCCTTAAT-3′ and reverse 5′-CCGACCTTAATTA ACATTTCCTTATTATTAAAGATCAGCT-3′.

The plasmid pCR001 was extracted and purified from *E. coli* cells grown to stationary phase in YT medium with the Qiagen Plasmid Midi Kit. To obtain relaxed DNA, 10 μg of the supercoiled plasmid was treated with *E. coli* Topoisomerase I as described above. Linear plasmid DNA used in the AFM analyses was produced by cleavage with the restriction enzyme *Ecl136*II. After relaxation or linearization reactions, the DNA was extracted with phenol/chloroform, purified with PCR purification Kit (Qiagen) and precipitated with ethanol. The relaxed and linearized DNA probes were resuspended in deionized water (Millipore).

Atomic force microscopic (AFM) analyses of Csn2–DNA complexes

Linearized or relaxed pCR001 was incubated with the Csn2 protein in the presence of 2 mM $CaCl_2$ in a total

volume of 5 µl. The DNA and protein concentrations are given in the individual experiments. The samples were incubated 10 min at room temperature to allow complex formation and diluted with 95 µl of adsorption buffer (5 mM $CaCl_2$, 2 mM Tris–HCl, pH 8.0, 10 mM NaCl). Forty-microliter aliquots of the samples were immediately transferred to freshly cleaved mica. After 1 min incubation the mica surface was washed three times with 100 µl deionized water (Millipore). Excess water was dried with compressed N_2. Imaging of Csn2–DNA complexes was performed with Nanowizard II (JPK Instruments AG, Berlin) in intermittent contact in air, using standard silicon cantilevers (OLYMPUS OMCL-AC160).

Biotinylation of dsDNA fragment and Csn2 binding studies

The 256 bp DNA fragment was obtained by *Eco*RI-*Bam*HI digestion of the plasmid pUC18-1. After purification of the DNA fragments by agarose gel electrophoresis, the biotinylation of the ends was achieved by Klenow incorporation of Biotin-11-dUTP. Ten microgram of the DNA fragment was incubated with 5 U Klenow (Promega), each 0.2 mM dCTP, dGTP, dATP and Biotin-11-dUTP (GeneOn) in 1 × Klenow buffer (Promega, 50 mM Tris–HCl, pH 7.2, 10 mM $MgSO_4$ and 100 µM DTT) in a total volume of 50 µl for 1 h at room temperature. After extraction with phenol/chloroform and ethanol precipitation, the DNA was dissolved in 10 µl TE buffer.

The binding studies with the biotinylated DNA were performed by sequential incubation of 168 ng of the biotinylated DNA in the presence of 7.2 mM $CaCl_2$ with 4.7 µg Csn2, 14 mM ethylene glycol tetraacetic acid (EGTA) and/or 2 µg streptavidin (Sigma-Aldrich) in a total volume of 14.4 µl. The volumes of the binding reaction without EGTA or streptavidin were adjusted by addition of deionized water (Millipore). The incubation of the samples occurred in three steps, each for 15 min at room temperature. The complexes were separated from free DNA on a native 2% Tris/Acetate agarose gel. The bands were visualized by ethidium bromide staining.

MD simulations

MD simulations were performed with the AMBER 11 suite of programs (34) together with the force field as described by Cornell *et al.* (35) using modifications suggested by Simmerling *et al.* (36). To generate a starting structure of Csn2 binding to dsDNA, canonical B-DNA of 36 bp length (sequence: 5′-GTTTTAGAGCTGTGCTG TTTCGAATGGTTCCAAAAC-3′) was inserted into the central hole of tetrameric Csn2 such that the DNA is perpendicular to a least-squares fit plane through the protein atoms and that Csn2 is displaced by ∼6 Å from the center of the DNA along the double strand. The tetrameric Csn2 was taken from Ellinger *et al.* (30) (PDB code: 3QHQ), including the Ca^{2+}-ions. In addition, a starting structure of Csn2 binding to dsDNA without Ca^{2+}-ions was generated. The starting structures were placed into octahedral periodic boxes of TIP3P water molecules (37). The distance between the edges of the water box and the closest atom of the solutes was at least 11 Å, resulting in system sizes of ∼186 000 atoms. The systems were minimized by 50 steps of steepest descent minimization followed by 450 steps of conjugate gradient minimization. The particle mesh Ewald method (38) was used to treat long-range electrostatic interactions, and bond lengths involving bonds to hydrogen atoms were constrained using SHAKE (39). The time-step for all MD simulations was 2 fs, with a direct-space, non-bonded cutoff of 8 Å. Applying harmonic restraints with force constants of 5 kcal mol^{-1} $Å^{-2}$ to all solute atoms, canonical ensemble (NVT)-MD was carried out for 50 ps, during which the system was heated from 100 to 300 K. Subsequent isothermal isobaric ensemble (NPT)-MD was used for 150 ps to adjust the solvent density. Finally, the force constants of the harmonic restraints on solute atom positions were gradually reduced to zero during 100 ps of NVT-MD. The following NVT-MD at 300 K with a time constant of 10 ps for heat-bath coupling was used for analysis, with conformations extracted every 20 ps. The simulation lengths amount to 100 ns of which the first 30 ns were discarded. Each simulation was repeated once with varied starting velocities, resulting in four independent simulations in total.

For analyzing the trajectories, conformations were superimposed with respect to the phosphorous atoms of the dsDNA or the C_α atoms of Csn2. The conformational variability of Csn2 binding to dsDNA agrees with that observed in MD simulations of Csn2 alone (30). For determining the translocation along the dsDNA, the distance between the center of the dsDNA and the average coordinates of the C_α atoms of all four helices H5 was computed. For determining the rotation of Csn2 relative to the dsDNA, the angle defined by the point triple (average coordinates of the phosphorous atoms of the five most central nucleotides on one strand; average coordinates of the phosphorous atoms of the five most central nucleotides on the other strand; average coordinates of the C_α atoms of one pair of helices H5 located in one of the legs of the tetrameric structure) was computed. For determining the kink angle of the dsDNA, the angle defined by the point triple (average coordinates of the phosphorous atoms of the five outermost base pairs at one end of the DNA; average coordinates of the phosphorous atoms of the five most central base pairs of the DNA; average coordinates of the phosphorous atoms of the five outermost base pairs at the other end of the DNA) was computed.

RESULTS

Binding of Csn2 requires free DNA ends

The Csn2 tetramers from *S. agalactiae* are stable in the range of pH 7.0–9.0, a wide range of ionic strengths as well as in the presence of Ca^{2+}-chelating EGTA (30). Owing to the high stability of the ring-shaped Csn2 tetramers, it is unlikely that the quaternary structure can be opened wide enough to bind and enclose dsDNA as it is known to occur in several DNA-enclosing proteins (40,41). Therefore, if the Csn2 tetramers encircle the dsDNA the

tetramers should bind DNA by threading onto the DNA ends. In contrast, a binding of the DNA outside the central hole will be independent on the presence of free DNA ends. To distinguish between these two alternative DNA-binding modes, we performed electrophoretic mobility shift assays (EMSAs) and tested whether the binding of Csn2 requires free DNA ends.

First, we incubated a 155 bp ^{32}P-radiolabeled DNA fragment with increasing amounts of Csn2 and separated the protein–DNA complexes on a 5% native Tris/glycine polyacrylamide gel. As can be seen in Figure 1A, Csn2 formed one complex band with an estimated apparent $K_{0.5}$ value between 30 and 60 nM (Figure 1A, lanes 3 and 4). The declining mobility of the complex band with increasing protein concentration indicates the binding of multiple Csn2 molecules to the DNA fragment. To test a putative requirement for a free DNA end for the interaction, we performed competition experiments with unlabeled plasmid DNA, either in closed circular or linearized form. As shown in Figure 1B (lanes 3–6), the complex formation of Csn2 with the linear radiolabeled DNA was challenged with increasing concentrations of the linearized pUC18-1 plasmid. However, the same plasmid in supercoiled or relaxed form almost did not compete for binding of Csn2 (Figure 1B, lanes 7–14; Figure 1C, lanes 19–22). This indicates that free DNA ends are necessary for Csn2-binding and supports the threading of dsDNA into the central channel of the Csn2 tetramers.

To further evaluate the dependence of the Csn2–DNA interaction on free dsDNA ends, we repeated the competition experiments with the plasmid DNA cleaved by restriction enzymes at two sites, with the aim to double the concentration of the DNA ends. The competition efficiency was higher with DNA cleaved with double cutters *Pvu*II or *Ava*II compared with the *Sca*I-linearized DNA at the same final competitor concentration (Figure 1C, lanes 7–14). Moreover, plasmid DNA cleaved with *Hae*III at 12 sites reduced the binding to the radiolabeled DNA already at the lowest competitor concentration tested (Figure 1C, lanes 15–18). Thus, the challenging efficiency depended on the concentration of free DNA ends at equal total competitor DNA concentration. The appearance of three intermediate complexes during the decay of the Csn2–DNA complexes is consistent with a binding of multiple Csn2 tetramers to the 155 bp dsDNA, and suggests a sliding of the bound proteins along the DNA after loading the DNA from its ends.

Cleavage of the pUC18-1 plasmid with the restriction enzymes *Sca*I, *Pvu*II and *Hae*III produces blunt ends, while cleavage with *Ava*II gives 5′-overhangs of three bases. The comparable challenging extent of the *Pvu*II- and *Ava*II-cleaved DNA demonstrated that the DNA end-binding activity of Csn2 is not restricted to blunt ends but short 5′-protruding DNA ends are also bound by Csn2 (Figure 1C, lanes 7–14). The ability of Csn2 to bind dsDNA with 3′- and/or 5′-overhangs was further examined with different synthetic DNA substrates

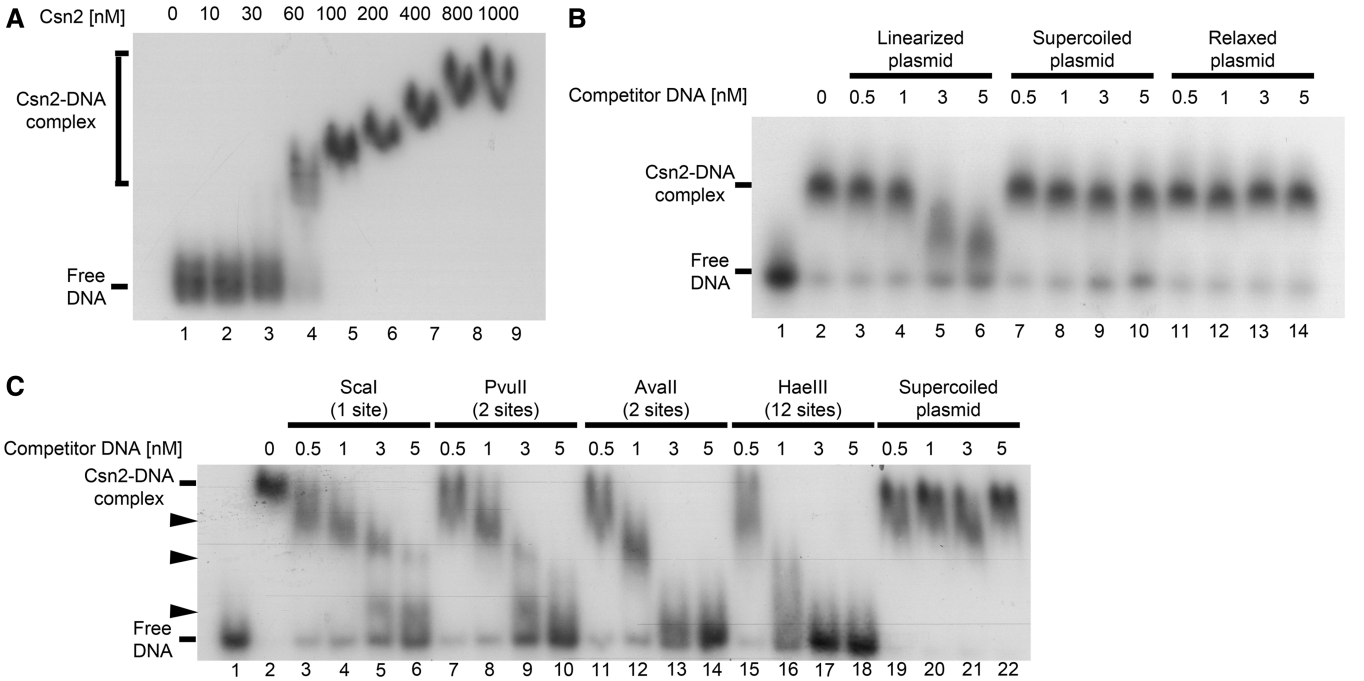

Figure 1. Electrophoretic mobility shift assays of a radiolabeled 155 bp DNA fragment with Csn2 either in the absence (**A**) or in the presence of competitor DNA (**B** and **C**) are presented. In each reaction 2 nM ^{32}P-labeled DNA, 20 ng/µl heparin, and 10 mM $CaCl_2$ were employed. (A) Titration of Csn2 in the range of 0 to 1 µM is shown. (B) Csn2 binding to the radiolabeled DNA fragment was competed with indicated amounts of 2915 bp unlabeled plasmid DNA either in *ScaI*-linearized (lanes 3–6), supercoiled (lanes 7–10) or in relaxed (lanes 11–14) form. The concentration of Csn2 was constant at 60 nM in lanes 2–14. Lanes 1 and 2 show the control reactions, performed either in the absence of Csn2 (lane 1) or in the absence of competitor DNA (lane 2). (C) The same competition experiment as in (B) but with *PvuII*-, *AvaII*-, or *HaeIII*-cleaved competitor plasmid. The numbers of cleavage sites of the different endonucleases are given in the brackets. The black arrowheads indicate intermediate Csn2-DNA complexes, resulting from decomposition of the fully occupied complexes.

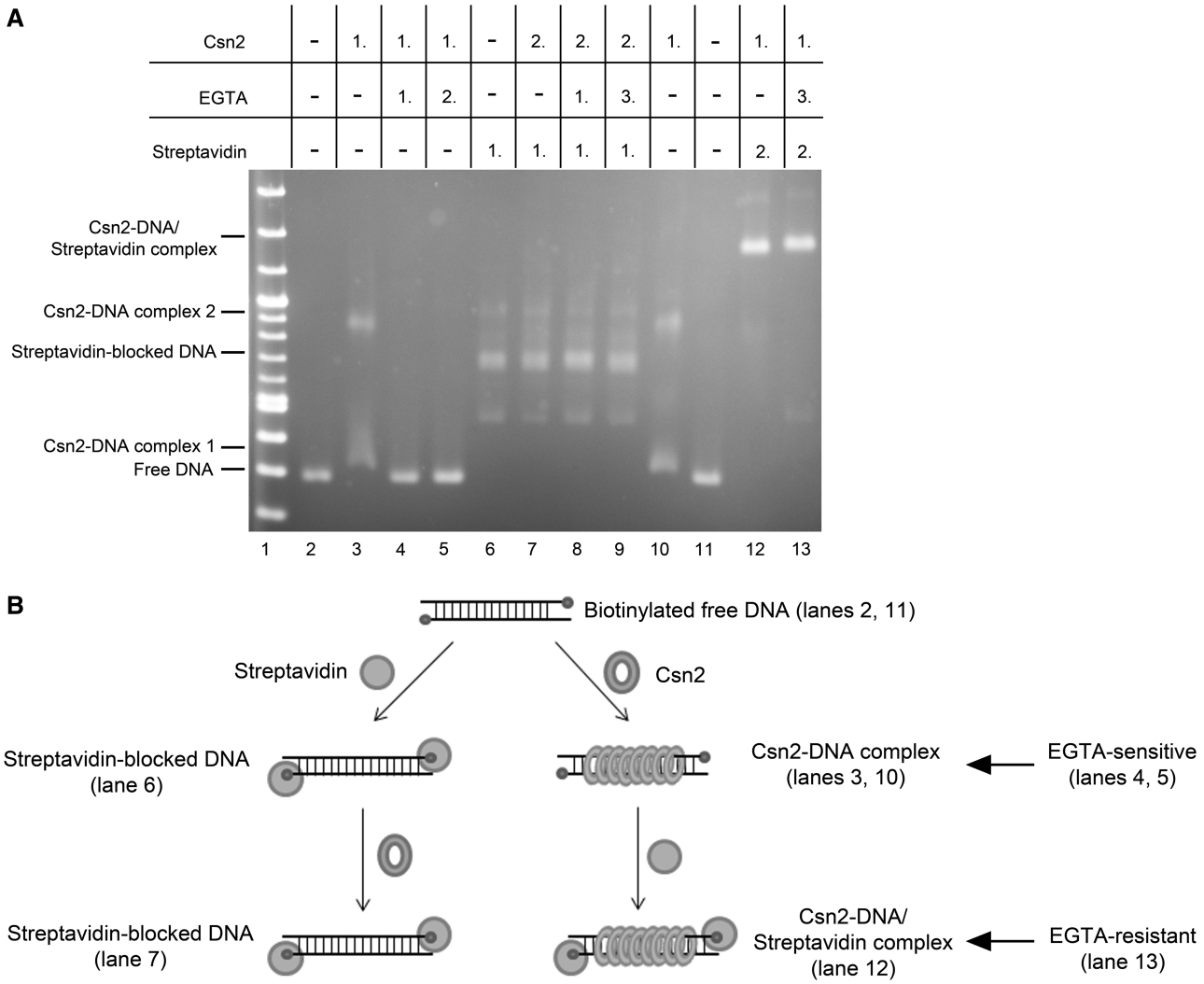

Figure 2. (**A**) Binding analyses of Csn2 in the presence and absence of EGTA and free DNA ends on 2% Tris-acetate agarose gel. In each lane 168 ng linear DNA and 7.2 mM $CaCl_2$ were employed. The numbers above the lanes indicate the order of addition of streptavidin (2 μg), Csn2 (4.7 μg), or EGTA (14 mM) in a total volume of 14.4 μl. Lanes 2–5: Influence of EGTA on Csn2-DNA interaction is shown. Lanes 6–9: 168 ng of the end-biotinylated DNA fragment were incubated first with streptavidin to block the DNA ends. Lanes 10 and 11: Streptavidin was added after binding of Csn2. After separation of the complexes the agarose gel was stained with ethidium bromide. (**B**) Schematic presentation of the binding analysis, shown in (A).

(Supplementary Figure S1). Csn2 did not form complexes with single-stranded DNA (ssDNA) (Supplementary Figure S1A and B, lanes 1–3). However, the hybridization of complementary oligonucleotides to ssDNA restored the Csn2–DNA interaction (Supplementary Figure S1A and B, lanes 4–6). Moreover, Csn2 formed faint complex bands with a DNA substrate that contains a 36-bp dsDNA region flanked by 29 or 34 nt ssDNA tails (Supplementary Figure S1B, lanes 7–9). These results indicate that Csn2 is in principle able to enter dsDNA regions by threading ssDNA tails in its central hole.

Dissociation of Csn2 occurs at the DNA ends: evidence for encircling the DNA by Csn2 rings

The crystal structure of the Csn2 from *S. agalactiae* revealed three Ca^{2+}-ions bound to the leg domain of each protomer (30). These ions have also been found in the crystal structures of Csn2 proteins from *E. faecalis* and *S. pyogenes* (29,31). The loss of one of the Ca^{2+}-ions within one protomer of Csn2 resulted in a reorientation of α-helices in the leg domain, which form the inner rim of the central hole (30) (Supplementary Figure S2). Thus, the Ca^{2+}-ions are considered to be important for the stabilization of the tetramer conformation that is able to bind the dsDNA (29–31).

Because capturing the Ca^{2+}-ions by EGTA did not disassemble the tetramers but abolished their DNA-binding activity (30,31), we suggested that in the presence of EGTA the conformation of the central hole becomes disordered, resulting in weakened interactions of the tetramers with the DNA backbone and leading to the dissociation of the sliding rings from the DNA ends. If this were the case, the obstruction of the DNA ends of

Csn2–DNA complexes should result in arrested Csn2 rings on the DNA, and the complexes should not dissociate in the presence of EGTA. To test this, we attached biotin to both ends of a dsDNA fragment, allowing the blocking of both DNA-termini with streptavidin.

Incubation of Csn2 with biotinylated dsDNA led to the formation of two complex bands, demonstrating that biotin itself did not interfere with Csn2 binding (Figure 2A, lanes 2, 10). As shown previously (29–31), in the presence of EGTA the complex formation was completely inhibited (Figure 2, lane 4). To test whether EGTA causes the dissociation of DNA-bound Csn2 molecules, we incubated the DNA with Csn2 in the presence of Ca^{2+} for 15 min and then added EGTA to the binding reaction followed by incubation for additional 15 min. Indeed, the lack of Csn2–DNA complexes indicated the decomposition of the complexes in the presence of EGTA (Figure 2A, lane 5).

As expected, the tethering of streptavidin to the DNA ends inhibited the binding of Csn2 (Figure 2A, lanes 6 and 7), consistent with the requirement of free DNA ends for Csn2 binding. In contrast, adding streptavidin after the binding of Csn2 resulted in supershifted Csn2-DNA/streptavidin complexes (Figure 2A, lane 12). Moreover, the Csn2 proteins in these complexes were captured on the DNA because adding EGTA did not lead to a decay of the complexes (Figure 2A, lane 13). Thus, both the association as well as the dissociation of Csn2 occurs at the DNA ends, indicating a sliding activity of the Csn2 rings (Figure 2B). In the absence of Ca^{2+}-ions, the tetramers slip down from the DNA ends, likely due to weakened electrostatic interactions between the distorted α-helices of the ring channel and the phosphate backbone of the DNA helix.

AFM of Csn2–DNA complexes

To confirm the DNA end requirement and to probe a binding of multiple Csn2 molecules, which would confirm a sliding activity of the Csn2 rings, we imaged the Csn2–DNA complexes by AFM in intermittent contact mode in air.

The AFM analyses were performed with the 5125 bp plasmid pCR001 as a substrate, either in linearized or relaxed circular form. AFM images of equal amounts of both DNA forms (each 1.3 nM) in the absence of Csn2 are shown Figure 3A and B. Some of the molecules of both DNA forms contained nodes, likely due to a crossover of DNA helices or a distortion/kinking of the DNA during the drying process of the probe on the mica surface (Figure 3B). AFM images of the same DNA mixture in the presence of 176 nM Csn2 showed the formation of protein–DNA complexes (Figure 3C–F). In average the ratio of the heights and widths of bound Csn2 and free DNA (Csn2-DNA/DNA) were 4.2 and 1.3, respectively. In contrast, the ratio of heights and widths of the DNA nodes relative to straight DNA were in average 1.9 and 1.5, respectively. Thus, these DNA nodes, which were also observed in the protein–free sample (Figure 3B), were not considered as protein–DNA complexes.

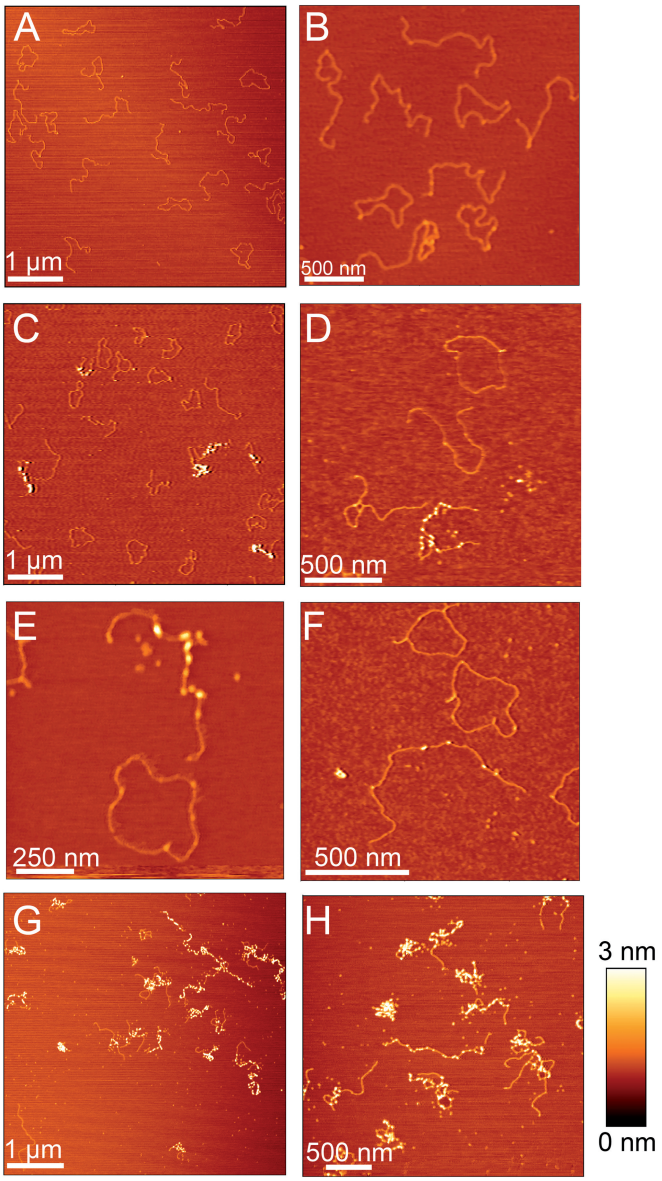

Figure 3. Representative AFM images of 5125 bp plasmid DNA in the absence or presence of Csn2 are shown. (**A**) and (**B**) show images of 1.3 nM relaxed and 1.3 nM linear plasmid DNA in the absence of Csn2; (**C–F**) images of equal amounts of relaxed and linear plasmid DNA (each 1.3 nM) incubated with 176 nM Csn2; (**G**) and (**H**) images of 2.6 nM linear plasmid DNA incubated with 800 nM Csn2. The relative color scale range is 0–3 nm in all images.

Csn2 was exclusively associated with linear DNA (Figure 3C–F), consistent with its DNA end-dependent–binding activity as observed in the mobility shift assays. The binding was not restricted to the DNA ends but internal sites along the entire DNA fragment were also occupied by several Csn2 proteins (Figure 3D–F), indicating a sliding activity of the tetramers after initial binding at the DNA ends. Moreover, the binding of Csn2 seems to be highly cooperative, as some linear DNA molecules were free of any protein, whereas others were bound by several Csn2 proteins in the same sample (Figure 3D–F). Such an apparent cooperative binding

activity is also known for the 'recombination associated protein' RdgC from *E. coli*, a ring-shaped DNA-binding protein with preference for DNA ends (42). As control, incubation of Csn2 with the relaxed plasmid in the absence of linear DNA confirmed that Csn2 does not bind to circular DNA (Supplementary Figure S3).

At higher Csn2 concentration, highly condensed nucleoprotein complexes were observed, likely caused by clustering of multiple Csn2 proteins and DNA-dependent self-association of DNA-bound Csn2 (Figure 3G and H). The formation of these higher-order nucleoprotein complexes should not be caused by an unspecific aggregation of the protein, as previous study has shown that Csn2 exists as stable tetramer in solution (30). To further evaluate the DNA end-dependence and to exclude an unspecific aggregation of Csn2 in the presence of DNA, we analyzed the elution profile of Csn2 in the presence of linear or closed circular plasmid DNA by size exclusion chromatography (Figure 4). Incubation of Csn2 with the linear DNA caused a co-elution of the entire Csn2 tetramers in the void fraction (Figure 4C). In contrast, the same DNA in supercoiled form did not change the elution profile of the tetrameric Csn2 (Figure 4D). The results are consistent with a selective binding of the Csn2 tetramers to the linear dsDNA and exclude an unspecific protein aggregation in the presence of DNA.

MD simulations reveal Ca^{2+}-dependent coupled translocation-rotation motions of Csn2 and kinking of the DNA

To further study Csn2 binding to DNA and the sliding activity of the tetramers, we performed two independent all-atom MD simulations of 100 ns length each of the tetrameric Csn2 with dsDNA of 36 bp length bound within the central hole. The sequence of the dsDNA corresponds to the CRISPR repeat sequence of the type II-A CRISPR system of *S. agalactiae*. Although we do not have an indication for a sequence-specific binding of Csn2, a bioinformatics analysis revealed a correlation between the presence of the *csn2* gene and a particular CRISPR repeat sequence (43).

The simulations reveal sliding motions of Csn2 along the DNA of a magnitude of >10 Å (Figure 5A), which is equivalent to about one-third of the B-DNA pitch. The motion pattern shows both inward and outward movements accompanied by repeated changes in the direction of motion, as expected for a process driven by thermal energy. Yet, during the last 70 ns of the simulations, the Csn2 molecules prefer to be more closely located to the center of the DNA: although distances up to 12 Å away from the center are observed during the simulations (Figure 5A), the combined likelihood from both trajectories of Csn2 being within 6 Å of the center is 66%. Thus, with the present length of the simulations, no preference for Csn2 to move toward one of the DNA ends could be detected.

The sliding motions of Csn2 are weakly ($r^2 = 0.21$) but significantly ($P < 0.001$) correlated to a rotation of the protein around the DNA (Figure 5B, D), as found for both independent MD simulations (data for the second simulation is shown in Supplementary Figure S4). The slope of the correlation line is ~13°/Å, which is only slightly larger than the turn angle per 1 Å rise of ~11° for B-DNA (44). The N-termini of helices H3 are likely responsible for this screw motion as the two helices within one dimer of Csn2, first, are almost perfectly collinearly oriented, second, are tilted by ~30° with respect to the axis of the dsDNA binding within the central hole and, third, possess a Lys (K78) at the N-terminus (Figure 5D). Taken together, this allows the pair of helices H3 of one dimer of Csn2 to fit with their N-termini in between the phosphate backbones of the major groove, while the helix pair of the other dimer is closer to the phosphate backbone of the minor groove. Therefore, on sliding, Csn2 must rotate around the dsDNA for helices H3 need to follow the course of the phosphate backbones, ultimately leading to a screw motion.

To probe the effect of removing Ca^{2+}-ions on Csn2 binding to the DNA, the above MD simulations were repeated without Ca^{2+}. No gross structural distortions of the tetrameric state of Csn2 were observed (Supplementary Figure S5), in agreement with the observed high stability of Csn2 even in the presence of EGTA (30). No release of Csn2 from the DNA was observed either, which is not surprising given the limited size of the water box enclosing the Csn2–DNA complex in the simulation that prevents a slipping down of Csn2. However, in contrast to the simulations with Ca^{2+}, no correlation between a sliding motion of Csn2 and its rotation could be detected in either one of the simulations ($r^2 < 0.07$). Thus, the absence of Ca^{2+} apparently disturbs the screw motion of Csn2, pointing to weakened interactions between Csn2 and DNA.

Both MD simulations furthermore reveal weak ($r^2 = 0.31$) or very weak ($r^2 = 0.06$) but significant ($P < 0.001$) correlations between the translocation of Csn2 from the DNA center and DNA bending in that a large kink angle in the DNA is preferentially observed if Csn2 moves toward a DNA end (Figure 5C and E). The DNA bending is fostered by the insertion of the N-terminus of at least one helix H3 of one of the Csn2 dimers into the major groove, and interactions between the phosphate backbones of the minor groove at two positions with Lys-rich loops between β4 and H2 of the other dimer, respectively (Figure 5E). As no correlation could be detected when plotting the kink angle against the distance of Csn2 from one end of the DNA, no indication was found either that the bending occurs preferentially at one end (and thus at a specific sequence) of the DNA.

DISCUSSION

In this study, we have characterized the DNA end-binding activity of the Cas protein Csn2 and investigated the structure of the Csn2–DNA complexes by atomic force microscopic imaging and MD simulations. Our results assign Csn2 as a new member of the class of multimeric toroidal proteins involved in DNA end-metabolism and indicate a pivotal role of Csn2 as an accessory protein

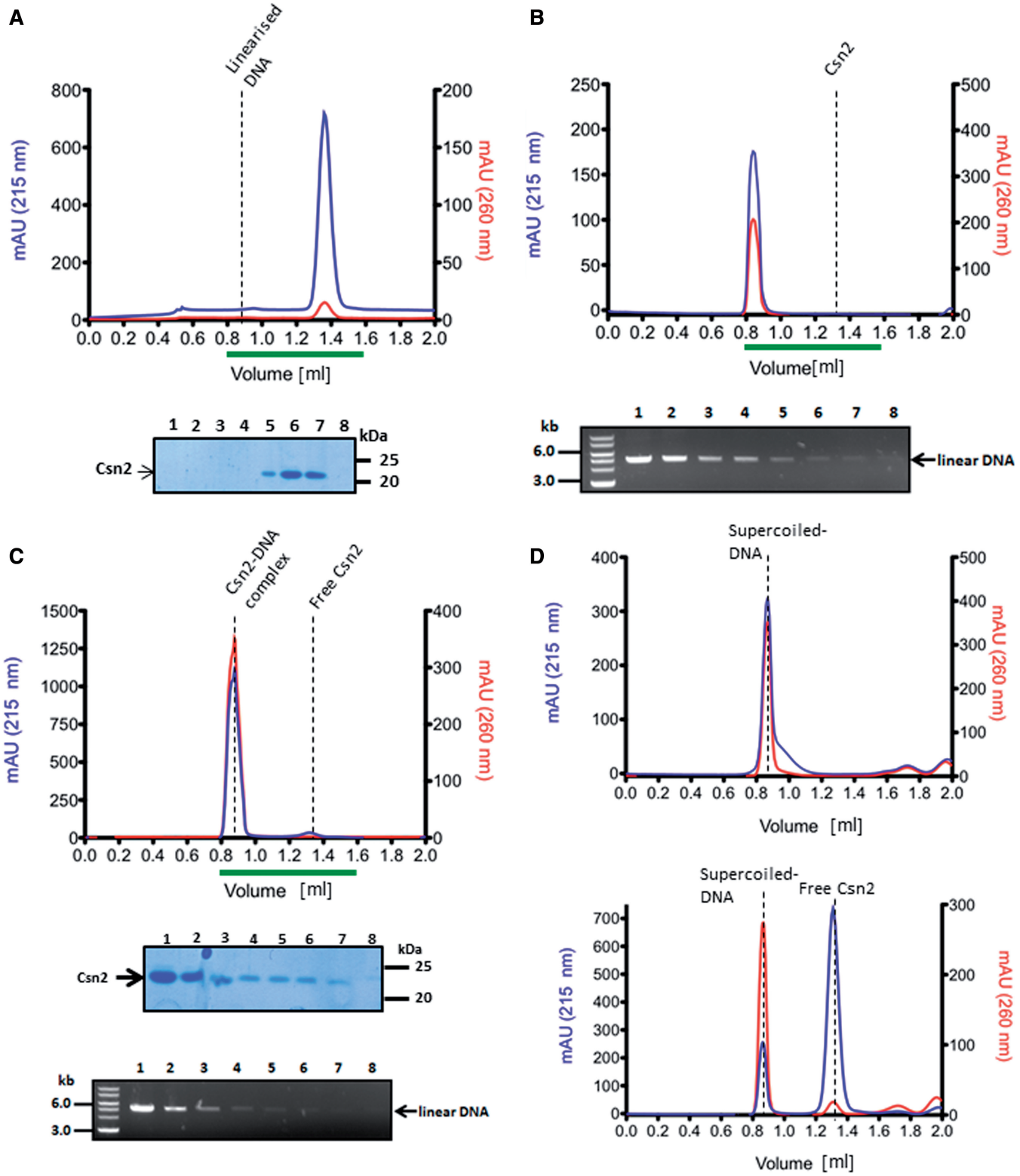

Figure 4. Analytical gel filtration analysis of Csn2 and Csn2-DNA complexes performed with a Superdex 200 PC 3.2/30 column is shown. Elution profiles of 20 µM Csn2 (**A**), 30 nM linearized pCR001 plasmid (**B**) and Csn2-DNA complexes (**C**) are shown. In (**D**) the elution profiles of 30 nM supercoiled pCR001 alone (upper part) or in the presence of 20 µM Csn2 (lower part) are shown. 100 µl fractions were collected starting at an elution volume of 0.8 ml. Aliquots of the fractions 1 to 8, indicated by the green lines below the elution profiles, were analyzed on 10% SDS gels (A, C) and on 1% agarose gels (B, C).

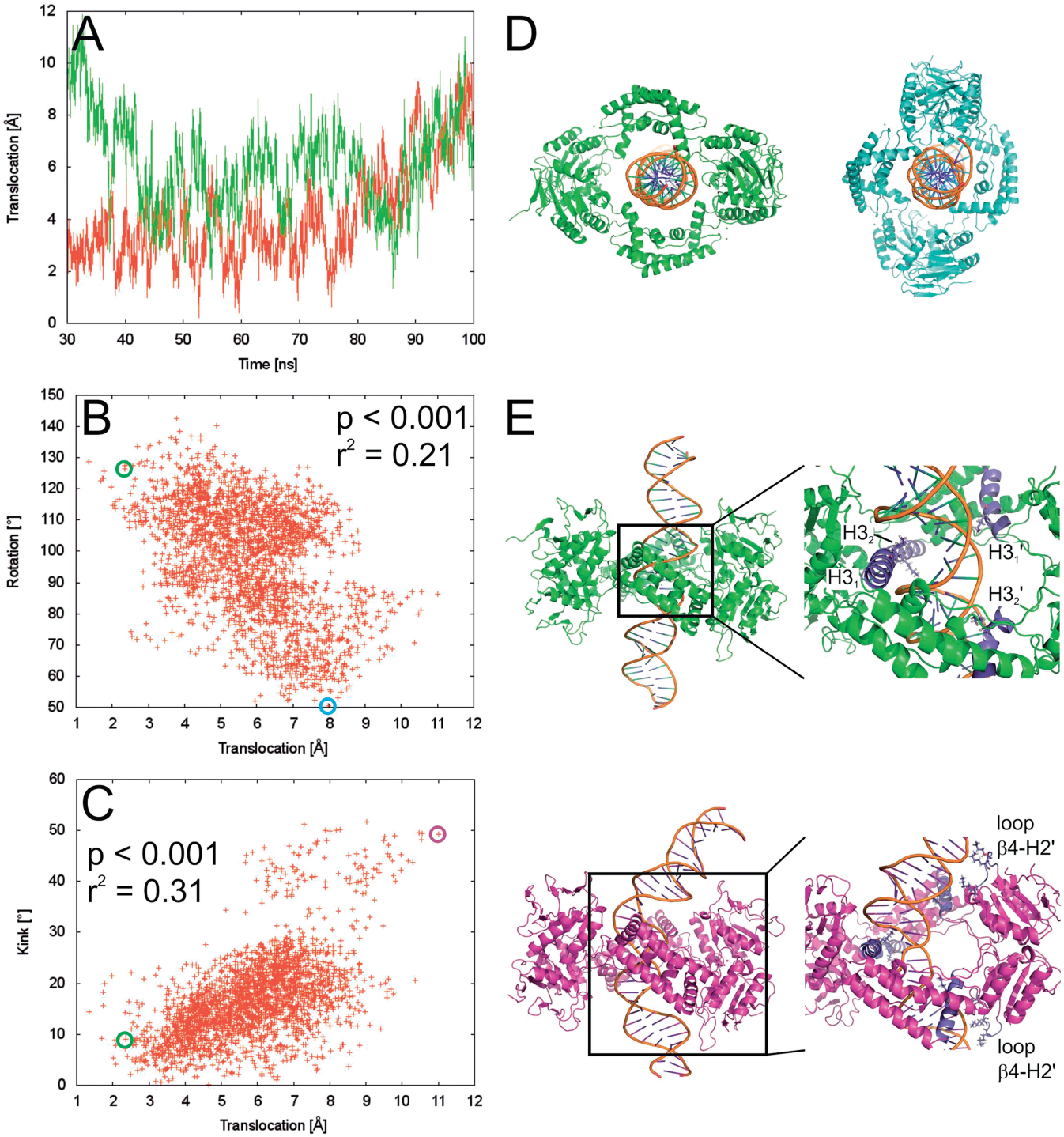

Figure 5. (**A**) Movement of Csn2 along the DNA determined from two independent MD simulations (red and green lines, respectively) over a time course of 70 ns. A translocation value of zero refers to Csn2 being centered on the DNA; translocation values >0 indicate a shift of Csn2 toward the termini of the DNA. (**B**) Coupled motions of translocation and rotation of Csn2 when moving along the DNA over the last 50 ns of the MD simulation. To determine the rotational motion, all conformations of the trajectories were aligned with respect to the phosphorous atoms of the DNA. (**C**) Coupling between DNA bending as determined by a kink angle and the translocation of Csn2 along the DNA over the last 50 ns of the MD simulation. (**D**) View along the DNA axis on Csn2-DNA conformations extracted from the MD trajectory at 50.3 ns (green) and 97.3 ns (cyan). The complexes were aligned with respect to the phosphorous atoms of the DNA and display a rotation of ~75° of Csn2 during a translocation of ~5.7 Å. In panel B, the respective data points are marked by circles. (**E**) Side view on Csn2-DNA complexes extracted from the MD trajectory at 50.3 ns (green) and 98.5 ns (magenta). The complexes display a kinking of the DNA by ~40° during a translocation of Csn2 of ~8.7 Å. In the close-up figures, helices H3 and loops β4-H2 are colored in blue, and Lys residues in these structural elements are depicted as sticks; unprimed labels mark helices and loops that belong to one dimer, primed labels mark objects that belong to the other dimer. In panel C, the respective data points are marked by circles.

during the integration of spacer sequences into the CRISPR array.

Csn2 tetramers bind at dsDNA ends and slide inward along the DNA

The toroidal structure of the DNA-binding protein Csn2 in different studies led to the suggestion that it binds to DNA through its positively charged central hole and encloses the DNA, although alternative binding mechanisms were also considered (29–31). Furthermore, the crystal structure of the Cas protein Stu0660 of the *S. thermophilus* has revealed the presence of a second group of the Csn2 protein family (Stu0660-like Csn2), the members of which share no sequence similarity but have a highly similar tetrameric structure (32). In contrast to the canonical Csn2 proteins, the Csn2-like Stu0660 protein has an extended C-terminal domain and binds to dsDNA without the need of Ca^{2+}-ions. Moreover, due to the lack of a mobility shift with a circular DNA, the authors suggested that Stu0660 'does not or barely' binds to circular DNA but selectively to linear dsDNA (32). Our study is consistent with this suggestion and clearly demonstrates that the Csn2 tetramer binds DNA through its positively charged central hole from the DNA ends and then slides inward along the DNA.

The sliding activity of Csn2 tetramers is not energy-dependent and thus likely not a directed translocation but rather a thermal energy-driven random walk fostered by electrostatic interactions between positively charged amino acids of Csn2 and the negatively charged phosphate backbone of the DNA. MD simulations revealed a rotation of Csn2 when moving along the helical axis of the DNA, similar to a rotation-coupled sliding of the sliding clamp PCNA (45), endonuclease EcoRV (46) or human oxoguanine DNA glycosylase 1 (47). One-dimensional diffusion of DNA-binding proteins is considered to facilitate finding of the specific binding sites (48). Although we have no evidence for a sequence-specific interaction *in vitro*, the reported correlation of the presence of the *csn2* gene and a particular CRISPR repeat sequence (43) could point to a specific role of Csn2 at the repeat sequence. The MD simulations were performed with the CRISPR repeat sequence, and one of the simulations indicated a bending of the DNA preferentially when Csn2 was located at an end. AFM is in principle suitable to study protein-induced DNA bending (49); however, the predicted Csn2-induced kinking is located at the ends of the DNA, and thus difficult to detect with AFM. Therefore, the structural deformation of the DNA ends on Csn2-binding, as proposed by the MD simulations, needs further experimental validation.

The cooperative binding of Csn2, observed in the EMSA and AFM analyses, indicates that multiple Csn2 tetramers are needed at the DNA ends to form stable Csn2–DNA complexes. We suggest that the binding of the first Csn2 tetramer at one DNA end is stabilized by loading of additional tetramers onto the same DNA end, which could impede a slipping down of the initially bound tetramers from the DNA end. This is consistent with the observation that the dissociation of Csn2–DNA complexes occurs through slipping down of the sliding tetramers at the DNA termini (Figure 2). In addition, at high concentrations, Csn2 tends to cluster along the DNA, leading to highly condensed nucleoprotein complexes. This apparent DNA-dependent self-association of the tetramers could point to a physical contact of the tetramers, contributing to the cooperative DNA end-loading of Csn2, as known for the Ku protein (50). The heterodimeric Ku protein forms a ring-like structure, binds preferentially at DNA ends, encircles the DNA (51), can slide along the DNA and promotes looped DNA structures (52). Moreover, it tends to cluster along the DNA as visualized by AFM (53). The DNA-binding properties of Ku are thus similar to the features of Csn2 observed in this study. The function of Ku *in vivo* is to mediate the formation of DNA end-synapsis and recruitment of recombination proteins to allow the repair of double-strand DNA breaks (54).

DNA end-metabolism and spacer integration

A series of multimeric ring-shaped proteins is known, which meet different functions in DNA metabolism, such as replication, transcription, recombination or DNA repair (40,41). In principle, there are three general ways how the DNA can get encircled by a ring-shaped protein: the quaternary structure can be opened wide enough to bind and enclose the DNA [e.g. RdgC (55)], the ring-shaped structure gets directly assembled around the DNA [e.g. PCNA (56)] or the DNA end has to pass through the central hole of the protein [e.g. Ku protein (51)]. In the latter case, the function of the protein is directly linked to DNA ends. Thus, the results presented here strongly suggest that the biological function of the CRISPR adaptation protein Csn2 is related to free DNA ends, which are likely formed during spacer acquisition.

The study of Yosef *et al.* (27) characterized a minimal requirement for spacer integration in type I-E CRISPR systems, and demonstrated the involvement of leader DNA and the nucleases Cas1 and Cas2 in the immunization process. The uptake of new spacers at the leader proximal end is also observed for other CRISPR subtypes, and together with the universality of Cas1 and Cas2, it seems that the main principle mechanism for spacer acquisition is similar in the different CRISPR subtypes. However, some of the CRISPR-Cas subtypes require in addition to Cas1 and Cas2 the proteins Csn2 or Cas4 for the acquisition of new spacer sequences. The Cas4 protein is proposed to be a subunit of the multiprotein complex Cascis (CRISPR-associated complex for the integration of spacers), mediating the CRISPR-mediated immunization (57). A recent study of the Cas4 protein from *Sulfolobus solfataricus* demonstrated that it contains a 5′–3′ DNA exonuclease activity, implying the requirement of DNA end modification/resection during integration of new spacer sequences into the CRISPR array (58). Moreover, the crystal structure of the Cas4 protein from *S. solfataricus* has been deposited in PDB database, showing that the Cas4 protein forms, like Csn2, a ring-shaped structure (PDB ID: 4IC1). Thus, the structural

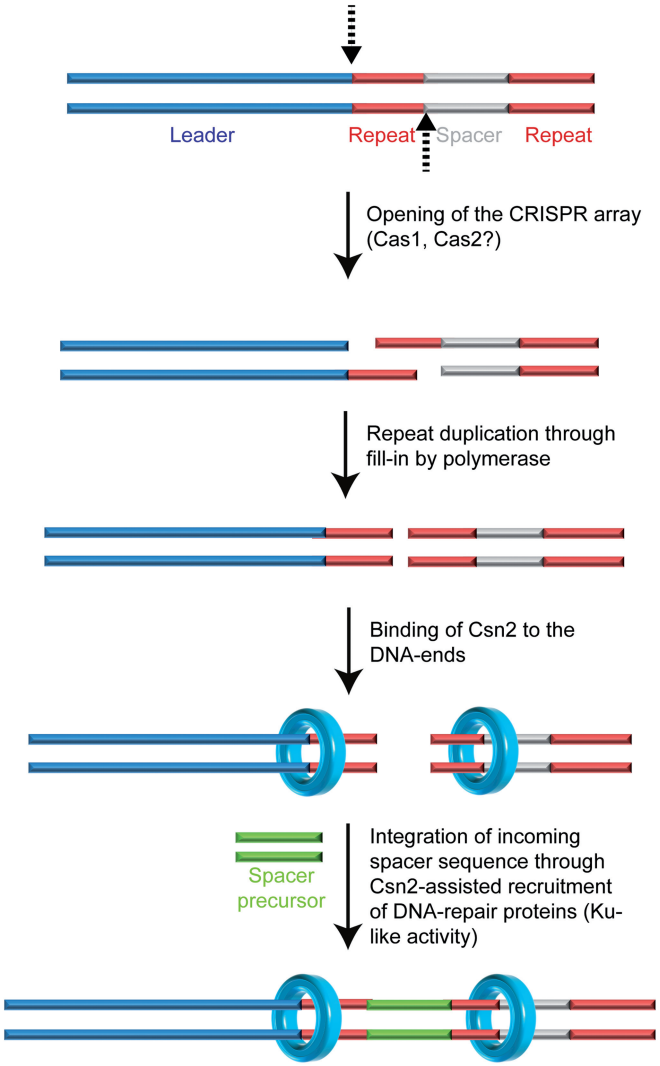

Figure 6. A model for the proposed function of Csn2 DNA end-binding in spacer integration is shown. The leader proximal repeat sequence serves as template for duplication of the repeat sequence (25,26,64), suggesting a cleavage at the leader-repeat and repeat-spacer borders (indicated by the dashed arrows). Complementary strand synthesis and binding of Csn2 to the DNA ends could assist in the integration of new spacer DNA fragments through holding the DNA ends together, while simultaneously recruiting DNA-repair proteins and protecting the double-stranded DNA breaks from exonucleolytic degradation.

similarity of Cas4 to AddB, its RecB-like exonuclease activity (58) and the DNA end-binding of Csn2 presented in this study are in agreement with DNA end-processing steps during the spacer integration. Another indication for the participation of DNA end-processing proteins in the adaptation stage is the genetic association of *nurA* and *herA* genes with type III-A CRISPR operons in Thermoproteales (59), encoding for an exonuclease–helicase complex involved in DNA end-resection (60).

The lack of enzymatic activity of Csn2 suggests that Cas4 and Csn2 are likely not functional homologs. However, the conserved co-localization within the type II systems and the DNA end-related activities of both proteins indicate that they are involved in DNA end-resection processes during integration of spacers. The reported protein–protein interaction of Cas1 with RecB in *E. coli* (61) and the AddB-like structure and activity of the Cas4 protein (58) could support an involvement of DNA repair systems during the uptake of CRISPR spacer sequences. The study of Chayot *et al.* (62) has demonstrated that the insertion of non-related exogenous sequences in *E. coli* depends on the RecBCD complex. The role of RecBCD is to process the DNA ends exonucleolytically on the 3′-end to produce microhomology with the incoming DNA, a mechanism termed alternative end-joining (A-EJ) (62). A similar mechanism could be catalyzed by Cas4 in type II-B or AddAB proteins in type II-A systems lacking Cas4.

The DNA end-binding activity of Csn2 suggests that it has a function as accessory protein on DNA ends. Csn2 could fulfill a task analogous to Ku during acquisition of exogenous DNA by non-homologous end-joining, by binding to the double-strand breaks to hold the ends together and/or by recruiting recombination proteins (63) (Figure 6). Alternatively, the DNA end binding activity of Csn2 could be related to the nuclease activity of Cas9, introducing specific double-strand DNA breaks into the target DNA (12,17). It is possible that after recognition and cleavage of the foreign DNA by Cas9-tracrRNA:crRNA ribonuleoprotein complex, Csn2 proteins are loaded at the DNA ends, e.g. to mark the foreign DNA as a substrate for new spacer sequences and/or to recruit the nucleases for specific generation of new spacer precursors. Indeed, such a coupling of the CRISPR interference with the immunization stage has been described to occur in the type I-E systems, known as 'primed acquisition' of spacer sequences (25). Although our analyses do not directly address the mechanisms of CRISPR spacer integration, the DNA end-binding of the adaptation protein Csn2 and the structure of the Csn2–DNA complexes resemble known recombination proteins acting on DNA end-metabolism. Therefore, we propose that the integration of new spacer DNA may occur by a recombination mechanism similar to non-homologous end-joining.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–5.

ACKNOWLEDGEMENTS

We thank Britta Ries and Andre Abts for assistance during the experiments. S.S. acknowledges the continuing support from the Heinrich Heine University. H.G. is grateful to the 'Zentrum für Informations- und Medientechnologie' (ZIM) at the Heinrich Heine University for computational support. We gratefully thank the members of the DFG Research unit FOR 1680 for helpful discussions.

FUNDING

Deutsche Forschungsgemeinschaft (DFG) [PU 435/1-1]; Strategischer Forschungsfonds at the Heinrich Heine University (to Ü.P.); initiative 'Fit for Excellence' at the Heinrich Heine University (to H.G.). Funding for open access charge: Institutional funds and Strategischer Forschungsfonds of the Heinrich Heine University.

Conflict of interest statement. None declared.

REFERENCES

1. Bhaya,D., Davison,M. and Barrangou,R. (2011) CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu. Rev. Genet.*, **45**, 273–297.
2. Al-Attar,S., Westra,E.R., van der Oost,J. and Brouns,S.J. (2011) Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. *Biol. Chem.*, **392**, 277–289.
3. Wiedenheft,B., Sternberg,S.H. and Doudna,J.A. (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, **482**, 331–338.
4. 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. *et al.* (2011) Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.*, **9**, 467–477.
5. Makarova,K.S., Grishin,N.V., Shabalina,S.A., Wolf,Y.I. and Koonin,E.V. (2006) A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct.*, **1**, 7.
6. van der Oost,J., Jore,M.M., Westra,E.R., Lundgren,M. and Brouns,S.J. (2009) CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem. Sci.*, **34**, 401–407.
7. Barrangou,R., Fremaux,C., Deveau,H., Richards,M., Boyaval,P., Moineau,S., Romero,D.A. and Horvath,P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, **315**, 1709–1712.
8. 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.
9. 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. *et al.* (2012) CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol. Cell*, **46**, 595–605.
10. Sapranauskas,R., Gasiunas,G., Fremaux,C., Barrangou,R., Horvath,P. and Siksnys,V. (2011) The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.*, **39**, 9275–9282.
11. Deltcheva,E., Chylinski,K., Sharma,C.M., Gonzales,K., Chao,Y., Pirzada,Z.A., Eckert,M.R., Vogel,J. and Charpentier,E. (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, **471**, 602–607.
12. Jinek,M., Chylinski,K., Fonfara,I., Hauer,M., Doudna,J.A. and Charpentier,E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816–821.
13. Hale,C.R., Majumdar,S., Elmore,J., Pfister,N., Compton,M., Olson,S., Resch,A.M., Glover,C.V. III, Graveley,B.R., Terns,R.M. *et al.* (2012) Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. *Mol. Cell*, **45**, 292–302.
14. Hale,C.R., Zhao,P., Olson,S., Duff,M.O., Graveley,B.R., Wells,L., Terns,R.M. and Terns,M.P. (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell*, **139**, 945–956.
15. 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. *et al.* (2011) Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat. Struct. Mol. Biol.*, **18**, 529–536.
16. Haft,D.H., Selengut,J., Mongodin,E.F. and Nelson,K.E. (2005) A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.*, **1**, e60.
17. Gasiunas,G., Barrangou,R., Horvath,P. and Siksnys,V. (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl Acad. Sci. USA*, **109**, E2579–E2586.
18. Jinek,M., East,A., Cheng,A., Lin,S., Ma,E. and Doudna,J. (2013) RNA-programmed genome editing in human cells. *Elife*, **2**, e00471.
19. Cho,S.W., Kim,S., Kim,J.M. and Kim,J.S. (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.*, **31**, 230–232.
20. Jiang,W., Bikard,D., Cox,D., Zhang,F. and Marraffini,L.A. (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.*, **31**, 233–239.
21. Hwang,W.Y., Fu,Y., Reyon,D., Maeder,M.L., Tsai,S.Q., Sander,J.D., Peterson,R.T., Yeh,J.R. and Joung,J.K. (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.*, **31**, 227–229.
22. Mali,P., Yang,L., Esvelt,K.M., Aach,J., Guell,M., DiCarlo,J.E., Norville,J.E. and Church,G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823–826.
23. Cong,L., Ran,F.A., Cox,D., Lin,S., Barretto,R., Habib,N., Hsu,P.D., Wu,X., Jiang,W., Marraffini,L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.
24. Qi,L.S., Larson,M.H., Gilbert,L.A., Doudna,J.A., Weissman,J.S., Arkin,A.P. and Lim,W.A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, **152**, 1173–1183.
25. Datsenko,K.A., Pougach,K., Tikhonov,A., Wanner,B.L., Severinov,K. and Semenova,E. (2012) Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat. Commun.*, **3**, 945.
26. Swarts,D.C., Mosterd,C., van Passel,M.W. and Brouns,S.J. (2012) CRISPR interference directs strand specific spacer acquisition. *PLoS One*, **7**, e35888.
27. Yosef,I., Goren,M.G. and Qimron,U. (2012) Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.*, **40**, 5569–5576.
28. Deveau,H., Barrangou,R., Garneau,J.E., Labonte,J., Fremaux,C., Boyaval,P., Romero,D.A., Horvath,P. and Moineau,S. (2008) Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.*, **190**, 1390–1400.
29. Nam,K.H., Kurinov,I. and Ke,A. (2011) Crystal structure of clustered regularly interspaced short palindromic repeats (CRISPR)-associated Csn2 protein revealed Ca2+-dependent double-stranded DNA binding activity. *J. Biol. Chem.*, **286**, 30759–30768.
30. Ellinger,P., Arslan,Z., Wurm,R., Tschapek,B., MacKenzie,C., Pfeffer,K., Panjikar,S., Wagner,R., Schmitt,L., Gohlke,H. *et al.* (2012) The crystal structure of the CRISPR-associated protein Csn2 from *Streptococcus agalactiae*. *J. Struct. Biol.*, **178**, 350–362.
31. Koo,Y., Jung,D.K. and Bae,E. (2012) Crystal structure of *Streptococcus pyogenes* Csn2 reveals calcium-dependent conformational changes in its tertiary and quaternary structure. *PLoS One*, **7**, e33401.
32. Lee,K.H., Lee,S.G., Eun Lee,K., Jeon,H., Robinson,H. and Oh,B.H. (2012) Identification, structural, and biochemical characterization of a group of large Csn2 proteins involved in CRISPR-mediated bacterial immunity. *Proteins*, **80**, 2573–2582.
33. Pul,U., Wurm,R. and Wagner,R. (2007) The role of LRP and H-NS in transcription regulation: involvement of synergism, allostery and macromolecular crowding. *J. Mol. Biol.*, **366**, 900–915.
34. Case,D.A., Cheatham,T.E. III, Darden,T., Gohlke,H., Luo,R., Merz,K.M. Jr, Onufriev,A., Simmerling,C., Wang,B. and Woods,R.J. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem.*, **26**, 1668–1688.

35. Cornell,W.D., Cieplak,C.I., Bayly,I.R., Gould,I.R., Merz,K.M., Ferguson,D.M., Spellmeyer,D.C., Fox,T., Caldwell,J.W. and Kollman,P.A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, **117**, 5179–5197.
36. Simmerling,C., Strockbine,B. and Roitberg,A.E. (2002) All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc.*, **124**, 11258–11259.
37. Jorgensen,W.L., Chandrasekhar,J., Madura,J. and Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926–935.
38. Darden,T., York,D. and Pederson,L. (1993) Particle Mesh Ewald—a Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089–10092.
39. Ryckaert,J.P., Ciccotti,G. and Berendsen,H.J.C. (1977) Numerical-Integration of Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.*, **23**, 327–341.
40. Hingorani,M.M. and O'Donnell,M. (1998) Toroidal proteins: running rings around DNA. *Curr. Biol.*, **8**, R83–R86.
41. Hingorani,M.M. and O'Donnell,M. (2000) A tale of toroids in DNA metabolism. *Nat. Rev. Mol. Cell. Biol.*, **1**, 22–30.
42. Tessmer,I., Moore,T., Lloyd,R.G., Wilson,A., Erie,D.A., Allen,S. and Tendler,S.J. (2005) AFM studies on the role of the protein RdgC in bacterial DNA recombination. *J. Mol. Biol.*, **350**, 254–262.
43. Kunin,V., Sorek,R. and Hugenholtz,P. (2007) Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol.*, **8**, R61.
44. Saenger,W. (1984) *Principles of Nucleic Acid Structure*. Springer-Verlag, New York.
45. Kochaniak,A.B., Habuchi,S., Loparo,J.J., Chang,D.J., Cimprich,K.A., Walter,J.C. and van Oijen,A.M. (2009) Proliferating cell nuclear antigen uses two distinct modes to move along DNA. *J. Biol. Chem.*, **284**, 17700–17710.
46. Dikic,J., Menges,C., Clarke,S., Kokkinidis,M., Pingoud,A., Wende,W. and Desbiolles,P. (2012) The rotation-coupled sliding of EcoRV. *Nucleic Acids Res.*, **40**, 4064–4070.
47. Blainey,P.C., Luo,G., Kou,S.C., Mangel,W.F., Verdine,G.L., Bagchi,B. and Xie,X.S. (2009) Nonspecifically bound proteins spin while diffusing along DNA. *Nat. Struct. Mol. Biol.*, **16**, 1224–1229.
48. von Hippel,P.H. and Berg,O.G. (1989) Facilitated target location in biological systems. *J. Biol. Chem.*, **264**, 675–678.
49. Dame,R.T., van Mameren,J., Luijsterburg,M.S., Mysiak,M.E., Janicijevic,A., Pazdzior,G., van der Vliet,P.C., Wyman,C. and Wuite,G.J. (2005) Analysis of scanning force microscopy images of protein-induced DNA bending using simulations. *Nucleic Acids Res.*, **33**, e68.
50. Ma,Y. and Lieber,M.R. (2001) DNA length-dependent cooperative interactions in the binding of Ku to DNA. *Biochemistry*, **40**, 9638–9646.
51. Walker,J.R., Corpina,R.A. and Goldberg,J. (2001) Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature*, **412**, 607–614.
52. Dynan,W.S. and Yoo,S. (1998) Interaction of Ku protein and DNA-dependent protein kinase catalytic subunit with nucleic acids. *Nucleic Acids Res.*, **26**, 1551–1559.
53. Cary,R.B., Peterson,S.R., Wang,J., Bear,D.G., Bradbury,E.M. and Chen,D.J. (1997) DNA looping by Ku and the DNA-dependent protein kinase. *Proc. Natl Acad. Sci. USA*, **94**, 4267–4272.
54. Feldmann,E., Schmiemann,V., Goedecke,W., Reichenberger,S. and Pfeiffer,P. (2000) DNA double-strand break repair in cell-free extracts from Ku80-deficient cells: implications for Ku serving as an alignment factor in non-homologous DNA end joining. *Nucleic Acids Res.*, **28**, 2585–2596.
55. Briggs,G.S., Yu,J., Mahdi,A.A. and Lloyd,R.G. (2010) The RdgC protein employs a novel mechanism involving a finger domain to bind to circular DNA. *Nucleic Acids Res.*, **38**, 6433–6446.
56. Kong,X.P., Onrust,R., O'Donnell,M. and Kuriyan,J. (1992) Three-dimensional structure of the beta subunit of E. *coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell*, **69**, 425–437.
57. Plagens,A., Tjaden,B., Hagemann,A., Randau,L. and Hensel,R. (2012) Characterization of the CRISPR/Cas subtype I-A system of the hyperthermophilic crenarchaeon T*hermoproteus tenax*. *J. Bacteriol.*, **194**, 2491–2500.
58. Zhang,J., Kasciukovic,T. and White,M.F. (2012) The CRISPR associated protein cas4 Is a 5′ to 3′ DNA exonuclease with an iron-sulfur cluster. *PLoS One*, **7**, e47232.
59. Bernick,D.L., Cox,C.L., Dennis,P.P. and Lowe,T.M. (2012) Comparative genomic and transcriptional analyses of CRISPR systems across the genus Pyrobaculum. *Front. Microbiol.*, **3**, 251.
60. Blackwood,J.K., Rzechorzek,N.J., Abrams,A.S., Maman,J.D., Pellegrini,L. and Robinson,N.P. (2012) Structural and functional insights into DNA-end processing by the archaeal HerA helicase-NurA nuclease complex. *Nucleic Acids Res.*, **40**, 3183–3196.
61. Babu,M., Beloglazova,N., Flick,R., Graham,C., Skarina,T., Nocek,B., Gagarinova,A., Pogoutse,O., Brown,G., Binkowski,A. *et al.* (2011) A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. *Mol. Microbiol.*, **79**, 484–502.
62. Chayot,R., Montagne,B., Mazel,D. and Ricchetti,M. (2010) An end-joining repair mechanism in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **107**, 2141–2146.
63. Doherty,A.J. and Jackson,S.P. (2001) DNA repair: how Ku makes ends meet. *Curr. Biol.*, **11**, R920–R924.
64. Goren,M.G., Yosef,I., Auster,O. and Qimron,U. (2012) Experimental definition of a clustered regularly interspaced short palindromic duplicon in *Escherichia coli*. *J. Mol. Biol.*, **423**, 14–16.