Journal of Structural Biology 178 (2012) 350–362

Contents lists available at SciVerse ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

The crystal structure of the CRISPR-associated protein Csn2 from *Streptococcus agalactiae*

Philipp Ellinger [a], Zihni Arslan [b], Reinhild Wurm [b], Britta Tschapek [a], Colin MacKenzie [c], Klaus Pfeffer [c], Santosh Panjikar [e,1], Rolf Wagner [b], Lutz Schmitt [a], Holger Gohlke [d,*], Ümit Pul [b,*], Sander H.J. Smits [a,*]

[a] Institute of Biochemistry, Heinrich Heine University, Universitätsstr. 1, 40225 Düsseldorf, Germany
[b] Institute of Molecular Biology of Bacteria, Heinrich Heine University, Universitätsstr. 1, 40225 Düsseldorf, Germany
[c] Institute of Medical Microbiology and Hospital Hygiene, Heinrich Heine University, Universitätsstr. 1, 40225 Düsseldorf, Germany
[d] Institute of Pharmaceutical and Medicinal Chemistry, Heinrich Heine University, Universitätsstr. 1, 40225 Düsseldorf, Germany
[e] EMBL Hamburg Outstation, c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany

ARTICLE INFO

Article history:
Received 24 February 2012
Received in revised form 3 April 2012
Accepted 7 April 2012
Available online 17 April 2012

Keywords:
CRISPR
Spacer integration
X-ray crystallography
MD-simulation
DNA binding

ABSTRACT

The prokaryotic immune system, CRISPR, confers an adaptive and inheritable defense mechanism against invasion by mobile genetic elements. Guided by small CRISPR RNAs (crRNAs), a diverse family of CRISPR-associated (Cas) proteins mediates the targeting and inactivation of foreign DNA. Here, we demonstrate that Csn2, a Cas protein likely involved in spacer integration, forms a tetramer in solution and structurally possesses a ring-like structure. Furthermore, co-purified Ca^{2+} was found important for the DNA binding property of Csn2, which contains a helicase fold, with highly conserved DxD and RR motifs found throughout Csn2 proteins. We could verify that Csn2 binds ds-DNA. In addition molecular dynamics simulations suggested a Csn2 conformation that can "sit" on the DNA helix and binds DNA in a groove on the outside of the ring.

1. Introduction

Microorganisms have developed several mechanisms to defend invasions by foreign nucleic acids. In contrast to the well-known defense strategies such as the restriction-modification system or inhibition of phage adsorption, the recently discovered CRISPR system (clustered regularly interspaced short palindromic repeats) functions as an inheritable and adaptive immune system of prokaryotes (Al-Attar et al., 2011; Karginov and Hannon, 2010; Marraffini and Sontheimer, 2010a). More than 90% of archaea and nearly 40% of sequenced bacteria are equipped with CRISPR (Grissa et al., 2007), consisting of one or more CRISPR cassettes and a group of CRISPR-associated (Cas) proteins (Haft et al., 2005; Jansen et al., 2002). The CRISPR cassette consists of short, often palindromic, 28–40 bp DNA repeat sequences, separated by non-identical spacer sequences of similar length. In general, the CRISPR cassettes represent the acquired memory of immunity, enabling recognition of the invader nucleic acid. More than 40 Cas protein families are known (Haft et al., 2005). According to the recent "polythetic classification" of the different CRISPR–Cas modules, three different CRISPR types (Type I, II, and III) exist, which differ in the Cas protein composition and defense mechanisms (Makarova et al., 2011).

CRISPR defense can be dissected in three steps, (i) CRISPR adaptation, (ii) CRISPR expression/processing and (iii) CRISPR interference. The first step (immunization or adaptation stage) describes the capturing of new spacers that originate from the cleavage of invading DNA to short pieces; these spacers are subsequently integrated into the CRISPR array of the host genome (Barrangou et al., 2007). The exact mechanism of this integration step is unknown. Two universal Cas proteins Cas1 and Cas2, which are common to the three CRISPR types, are likely involved in this process (Deveau et al., 2010; van der Oost et al., 2009). Both proteins have been identified as nucleases (DNase/RNase activity for Cas1 and RNase activity for Cas2) (Babu et al., 2011b; Beloglazova et al., 2008; Samai et al., 2010; Wiedenheft et al., 2009). Short motifs within the invading DNA, typically a few nucleotides in length and referred to as proto-spacer-adjacent motifs (PAMs), seem to be

Abbreviations: CRISPR, clustered regularly interspaced short palindromic repeats; Cas, CRISPR-associated; PolDom, polymerase domain.

* Corresponding authors. Fax: +49 211 81 13847 (H. Gohlke), fax: +49 211 81 15167 (Ü. Pul), fax: +49 211 81 15310 (S.H.J. Smits).
E-mail addresses: gohlke@hhu.de (H. Gohlke), pul@hhu.de (Ü. Pul), sander.smits@hhu.de (S.H.J. Smits).
[1] Present address: Australian Synchrotron, 800 Blackburn Road, Clayton, VIC3168, Australia.

1047-8477/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jsb.2012.04.006

involved in recognition/selection of new spacers (Mojica et al., 2009). In addition, the PAM sequences are required to prevent a CRISPR autoimmunity reaction (Marraffini and Sontheimer, 2010b).

The second step comprises the transcription of the CRISPR array into the precursor CRISPR-RNA (pre-crRNA) and expression of Cas proteins; the latter mediate processing of pre-crRNA to active crRNAs. In type I CRISPR systems, this processing is carried out by a multiprotein complex termed Cascade (Brouns et al., 2008). The CRISPR Nmeni-subtype, one of eight subtypes, belongs to the type II CRISPR systems and processes the pre-crRNA with the help of a trans-encoded tracrRNA and RNaseIII and the multidomain protein Cas9 (known as Csn1) (Deltcheva et al., 2011).

The third step is the inactivation of invading foreign DNA, or in some cases RNA, which is also performed by the Cascade complex and Cas3 in type I (Brouns et al., 2008) or the Cmr/Csm complex (Hale et al., 2009) in type III. The foreign DNA is targeted with the help of the crRNA and inactivated by digestion.

In the Nmeni CRISPR subfamily, the Cas9 protein is involved in neutralizing invading DNA (Garneau et al., 2010), whereas the endonucleases (or nuclease) Cas1 and Cas2 as well as the subfamily specific Csn2 protein are involved in spacer integration, although their precise roles in this process are still obscure (Sapranauskas et al., 2011). Most functional information of these proteins comes from knockout studies. For example, deletion of Cas9 in *Streptococcus thermophilus* abolishes the resistance against foreign DNA. Knockdown of Csn2 still confers the ability to target and inactivate invading, foreign DNA with the help of existing spacers (Barrangou et al., 2007; Garneau et al., 2010; Sapranauskas et al., 2011). This led to the assumption that Csn2 proteins are involved in spacer integration. Recently, Csn2 from *E. faecalis* was solved at 2.7 Å resolution, revealing a tetrameric ring-like structure. Furthermore it was shown that Csn2 is capable of binding ds-DNA, and it was proposed that the DNA binds through the center of the ring (Nam et al., 2011).

Here, we report the X-ray structure of Csn2 of *S. agalactiae* ATCC13813 at 2.0 Å resolution and reveal that Csn2 is a stable tetramer that can undergo significant conformational changes. Furthermore, we show that Csn2 binds ds-DNA. Csn2 could have two binding modes: (I) "sitting" on the DNA and (II) the DNA passes through the center of the ring.

2. Materials and methods

2.1. Recombinant protein expression vectors

Full length Csn2 was subcloned into pET28b (Novagen) on a *Nco*I/*Xho*I fragment generated by PCR from genomic *Streptococcus agalactiae* ATCC13813 DNA using the following primers: For 5′-TTCCTAGGAGATTCCCATGGTCAAGATTAATTTTCCAAT-3′ and Rev 5′-TATTACTGTGCTTTTACTACTCGAGTACCATATTTTCGCC-3′. The coding sequence was introduced in frame with the start codon and a C-terminal hexahistidine purification tag encoded by the vector. The Csn2 mutants $E_{24}A$, $K_{132}A$, $Y_{29}A$, $Y_{29}W$ and $R_{198}A$, $R_{199}A$ were prepared using the QuikChange XL mutagenesis kit (Stratagene). All sequences were verified by in-house sequencing at the Biological Medical Research Centre (BMFZ) Heinrich Heine University, Düsseldorf.

2.2. Recombinant protein expression

pET28b-Csn2 transformed *Escherichia coli* BL21(DE3) pLysS cells (Novagen) were grown in LB medium at a temperature of 37 °C to an OD_{600} of ~0.6 and then protein expression was induced by adding β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Following induction, cells were grown at 30 °C for 5 h and harvested by centrifugation at 7500*g* for 15 min and 4 °C, flash frozen in liquid nitrogen and stored at −80 °C.

2.3. Purification of Csn2

Cells were resuspended in lysis buffer (50 mM Tris–HCl, pH 8.0, 150 mM NaCl, 2 mM β-mercaptoethanol (β-ME) and lysed using a cell disruptor (Constant Systems). The lysate was clarified by centrifugation at 40,000*g* in a Ti45 rotor (Beckmann) for 1 h at 4 °C. Imidazole was added to the Csn2 supernatant to a final concentration of 20 mM before application to a 5 ml HiTrap Chelating column (GE Healthcare) loaded with nickel, which had been pre-equilibrated with 5 column volumes (CV) of running buffer (50 mM Tris–HCl, pH 8.0, 150 mM NaCl, 2 mM β-ME, 20 mM imidazole). After application of the supernatant to the column, it was washed with 5 CV of running buffer and elution was performed by a linear-gradient with running buffer containing 300 mM imidazole. Pooled fractions were loaded onto a HiLoad Superdex 200 prep grade 16/60 column (GE Healthcare) equilibrated with 1.3 CV of running buffer (10 mM Tris–HCl, pH 8.0, 50 mM NaCl, 1 mM DTT). Fractions containing protein were pooled, concentrated using an *Amicon* Ultra-15 filter (Millipore) with a cut-off of 10 kDa and stored at 4 °C.

2.4. Expression and purification of selenomethionine-labelled Csn2

For selenomethionine substitution, *E. coli* B834(DE3) cells were grown in M9 minimal medium supplemented with 50 μg/ml of L-selenomethionine (Molecular Dimensions). Expression and purification were identical with those for native Csn2.

2.5. Native page of Csn2

The oligomeric state of Csn2 and its mutants was analyzed by native gel electrophoresis using a 4–16% NativePAGE™ Novex Bis-Tris gradient gel (Invitrogen), which was stained with Coomassie brilliant blue.

2.6. Multiple angle light scattering

To determine the oligomeric size of Csn2, a multiple angle light scattering (MALS) setup consisting of miniDAWN Treos/optiLAB rex (Wyatt Technologies) connected to an Äkta Purifier (GE Healthcare) using a Superdex 200 10/300 analytical size exclusion column was used. The flow rate was set to 0.2 ml/min and ultraviolet (UV) detection was monitored at 280 nm. Light scattering was detected at angles of 0°, 90°, and 107° and the obtained values were averaged and evaluated using the program ASTRA (Wyatt Technologies). For light scattering experiments different NaCl concentration were used in SEC running buffer (50 mM, 250 mM and 500 mM).

2.7. Crystallization of Csn2

Crystallization trials were carried out at 4 °C. Crystals of Csn2 (at 25 mg/ml) were grown by mixing protein solution with reservoir solution containing 0.1 M Hepes, pH 6.8–7.2, 9–12% (w/v) PEG6000 and 0.05 M phenol in a 1:1 ratio. Crystals normally grew in 7–10 days. Suitable crystals were cryo-protected using crystallization buffer with 30% (v/v) ethylene glycol and then cryo-cooled in liquid nitrogen. Selenomethionine derivatised crystals were obtained and treated using the same method.