Biodesign

pISSN 2288-6982 | eISSN 2288-7105

CRYSTALLIZATION P 110-113

Crystallization and preliminary structural study of the *Thermobaculum terrenum* Csh2 in the type IB CRISPR system

Do-Heon Gu[1], Suk-Youl Park[2] and Jeong-Sun Kim[1]*

[1]Department of Chemistry, Chonnam National University, 77 Yongbong-ro, Buk-gu, Gwangju 61186, Republic of Korea, [2]Pohang Accelerator Laboratory, Pohang, Gyeongbuk 37673, Republic of Korea. *Correspondence: jsunkim@chonnam.ac.kr

CRISPR/Cas system is an RNA-mediated prokaryotic adaptive immune system against mobile genetic elements. Among three CRISPR/Cas systems, type I system has two functional units in removing nucleic acids, Cascade and Cas3. Type IB Cascade has been analyzed to have components of Csh1, Csh2, Cas4, and Cas5, whose structural and functional information is not available. To provide a structural background for the type IB Cascade, the preliminary structural study of the Csh2 from *Thermobaculum terrenum* is reported here. The purified recombinant protein was successfully crystallized using the precipitant solution composed of 10% (w/v) Polyethylene Glycol 3350, 2% (v/v) Tacsimate at pH 5.0, and 100 mM Sodium Citrate tribasic dihydrate at pH 5.6. Diffraction data were collected to a resolution of 1.93 Å using synchrotron radiation. The crystal belongs to the primitive triclinic P1 space group, with unit cell dimensions a = 31.04 Å, b = 54.65 Å, c = 96.46 Å, α = 73.6 °, β = 88.4 ° and γ = 89.8 °. With two molecules in the asymmetric unit, the crystal volume per unit protein weight corresponds with 2.29 $Å^3 \cdot Da^{-1}$, resulting in 46% solvent.

INTRODUCTION

Most of living organisms have diverse defense systems to protect themselves from outside incomers. For the mobile genetic elements, many prokaryotes have two representative kinds of foreign gene neutralizing systems, called innate and acquired system (Fineran et al., 2012; Sampson et al., 2013; Barrangou et al., 2014; Koonin et al., 2017). As for the acquired immune system, the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated (Cas) system uses fragments of genetic information of past visitors, which had been inserted in the CRISPR array, as a guide to recognize a re-visiting genetic molecule (Barrangou et al., 2007; Sapranauskas et al., 2011; Gong et al., 2014; Plagens et al., 2015; Amitai et al., 2016).

The current classification of CRISPR/Cas systems has three main types (I, II, and III) and eleven subtypes (IA to IF, IIA to IIC, and IIIA to IIIB), based on phylogenetic and often limited functional studies (Kunin et al 2007; Makarova et al., 2011). Compared with other types, type I system has two functional groups in silencing the foreign DNA, Cascade (CRISPR-associated complex for antiviral defense) and Cas3. Cascade can recognize and bind the foreign DNA. However, it cannot degrade foreign DNA directly, which is aided with Cas3 helicase-nuclease (Sinkunas et al., 2011; Beloglazova et al 2011; Mulepati et al 2011; Huo et al 2014; Gong et al., 2014). Cascade forms a multimeric ribonucleoprotein complex that processes the transcript of the CRISPR arrays and binds to CRISPR RNAs. Recognition of the PAM (protospacer adjacent motif) sequence using the CRISPR RNA induces the formation of the hetero-duplex complex with the target DNA, resulting in formation of the R-loop (RNA-mediated loop) (Brouns et al., 2008; Jore et al., 2011; Nam et al., 2012; Westra et al., 2012). Cas3 helicase-nuclease binds to the single-stranded DNA of the R-loop and degrades the target DNA in an endo- and exo-nucleolytic manner (Jore et al., 2011). Even though many literatures on the subtype IE and IF are repeatedly published (Savitskaya et al., 2013; Westra et al., 2013; Pawluk et al., 2014; Richter et al., 2014; Arslan et al., 2014; Gong et al., 2014; Patterson et al., 2015; Bondy-Denomy et al., 2015; Müller-Esparza et al., 2017), information on other subtypes are not available, which are necessary for understanding the mechanism of the complicated and sophisticate type I CRISPR/Cas systems.

Type IB locus contains the annotated gene array of cas7, cas8b, cas6b, cas5, cas3, cas1, cas2, and cas4 (Emmanuelle et al., 2015; Makarova et al., 2011, 2015). Like the type IE system (Savitskaya et al 2013; Westra et al., 2013; Semenova et al., 2016; Fu et al., 2017; Wright et al 2017), Cas1 and Cas2 take the role of a spacer acquisition in the adaptation stage, while Cas6b processes the transcript of a CRISPR array and generates mature CRISPR RNAs in the expression step. By similarity, four proteins of Cas8b (Csh1), Cas7 (Csh2), Cas4, and Cas5 appear to compose a Cascade complex that degrades the invaded DNA with Cas3 in the silencing step. A thermophilic microorganism *Thermobaculum terrenum* has several CRISPR/Cas systems including the type IB system (Kiss et al., 2010; Makarova et al., 2011; Plagens et al., 2015).

In order to add structural information on the CRISPR/Cas system, the preliminary structural studies on the Csh2 from *T. terrenum* were performed and the obtained diffraction data using the synchrotron X-ray source was analyzed.

RESULTS AND DISCUSSION

The cloned csh2 gene from *T. terrenum* was successfully expressed in *E. coli* and the expressed Csh2 protein was purified through sequential chromatographic steps of Ni affinity and ion exchange. The purified recombinant Csh2 protein was concentrated up to 3.4 $mg \cdot mL^{-1}$ in the buffer consisting of 20 mM Tris-HCl at pH 8.0, 200 mM NaCl, and 0.1 mM PMSF (phenylmethylsulfonyl fluoride). Suitable crystals for diffraction experiments were obtained within 10 days at 294 K (Figure 2) using the precipitant composed of 10% (w/v) Polyethylene Glycol 3350, 2% (v/v) Tacsimate at pH 5.0, and 100 mM Sodium Citrate tribasic dihydrate at pH 5.6. The obtained crystal diffracted to a resolution of 1.93 Å, using a synchrotron radiation (Figure 3). The crystal belongs to the primitive triclinic P 1 space group with unit cell parameters of a = 31.04 Å, b = 54.65 Å, c = 96.46 Å, α = 73.6°, β = 88.4°, and γ = 89.8° (Table 1). With two molecules in the asymmetric unit, the crystal volume per unit protein weight corresponds with 2.29 $Å^3 \cdot Da^{-1}$, corresponding to 46% solvent in the crystal (Matthews, 1968). Since no structural homolog that shows a reasonable sequence identity to this Csh2 protein is not available in Protein Data Bank, structural determination of the *T. terrenum* Csh2 is being tried by the crystal derivatized with a heavy atom.

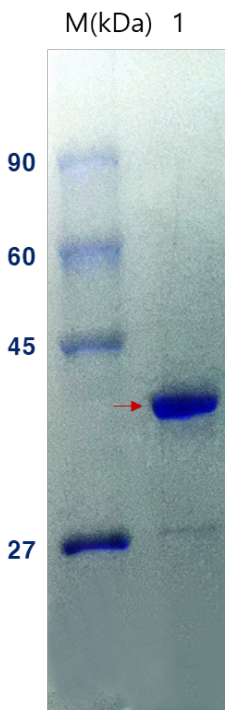

FIGURE 1 | 12 % (w/v) SDS-PAGE analysis of the concentrated Csh2. The purified and concentrated protein was analyzed by 12% (w/v) SDS-PAGE. The position of the *T. terrenum* Csh2 protein is indicated with a red arrow. M, protein size marker; 1, the concentrated Csh2 protein.

METHODS

Construction of a recombinant plasmid, expression, and purification of Csh2

The csh2 gene (Tter2426) in *T. terrenum* has 960 nucleotides which encode a polypeptide of 319 residues. The target gene was amplified from *T. terrenum* YNP1 genomic DNA by polymerase chain reaction (PCR) using the primers that were designed for Ligation-Independent Cloning

TABLE 1 | Data-collection statistics for *T. terrenum* Csh2

Values in parentheses are for the highest resolution shell	
X-ray source	5C (SBII 5C), PAL
Space group	P 1
Unit-cell parameters (Å, °)	a = 31.04, b = 54.65, c = 96.46 α = 73.6, β = 88.4, γ = 89.8
Resolution range (Å)	50 - 1.93
Measured reflections	146705
Unique reflections	43289
Redundancy	3.4 (2.9)
Temperature (K)	95
Matthews coefficient ($Å^3Da^{-1}$)	2.29
Solvent content (%)	46
Completeness (%)	98.8 (95.9)
I / δ (I)	20.5 (2.2)
R_{sym}[1] (%)	8.9 (41.3)

[1] $R_{sym} = \sum_{hkl} \sum_j | I_j - < I > | / \sum_{hkl} \sum_j I_j$, where I is the observed intensity of an individual reflection and < I > is the mean intensity of that reflection.

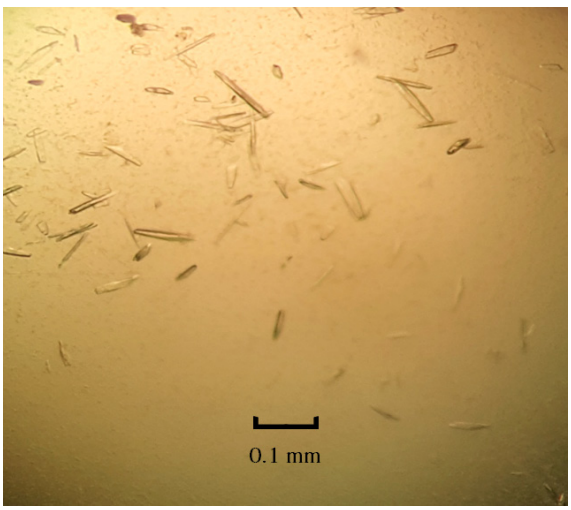

FIGURE 2 | A representative photograph of Csh2 crystals.

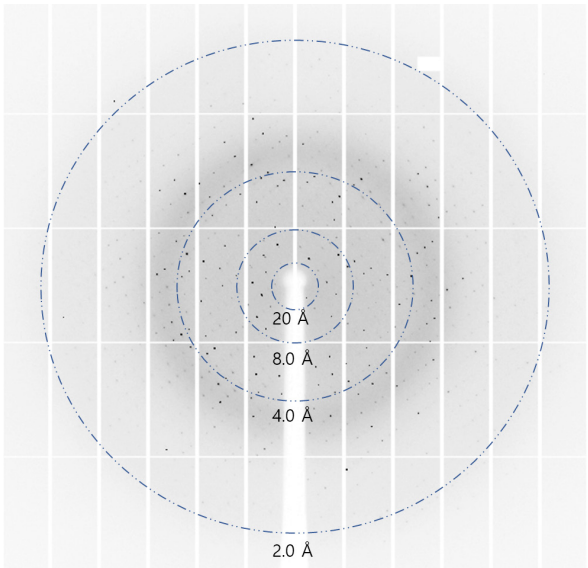

FIGURE 3 | Representative diffraction image of the *T. terrenum* Csh2 crystal. The crystal was exposed for 1 s with per-frame oscillation 0.5°. The overlaid blue circles indicate limits of each resolution.

(LIC) (Aslanidis and de Jong, 1990). The PCR product was treated with the T4 polymerase and ligated with a vector pLIC.B4, a derivative pET21a (Novagen, Madison, WI). The target gene was designed to translate the protein in a fused state with a 6-His-Maltose Binding Protein (MBP) sequence and Tobacco etch virus (TEV) protease cleavage sequence at the N-terminus of the Csh2 protein. The recombinant plasmid was transformed into *E. coli* BL21(DE3) and was grown in Luria-Bertani (LB) media supplemented with 50 mg·mL^{-1} ampicillin at 210 K. When the optical density at 600 nm reached 0.5, the fusion protein was expressed by adding 1.0 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) into the culture media and by incubation for an additional 18 h at 298 K. The culture was harvested by centrifugation at 5,000 g at 277 K. The cell pellet was resuspended in an ice-cold buffer A consisting of 20 mM Tris-HCl pH 8.5, 10 mM β-mercaptoethanol, 500 mM NaCl, and 5% (v/v) Glycerol, and disrupted by ultrasonication. Cell debris was removed by centrifugation at 11,000 g for 30 min. The expressed Csh2 fusion protein was initially purified by His-binding agarose resin (ELPIS, Daejeon, Korea). TEV protease was mixed in order to cleave the His_6-tag and the MBP at the N-terminus and the pool was simultaneously dialyzed to remove the salt. To further purify the protein, anion exchange chromatography using the 5 ml HiTrap™ Q column (GE Healthcare, Uppsala, Sweden) was performed. The final purified protein had a >95% purity based on Coomassie blue-stained SDS-PAGE (Figure 1).

Crystallization

For crystallization, the purified Csh2 protein was concentrated to 3.4 mg·mL^{-1} in the buffer consisting of 20 mM Tris-HCl at pH 8.0, 200 mM NaCl, and 0.1 mM PMSF (phenylmethylsulfonyl fluoride). Protein concentration was determined by considering the extinction coefficient of 1.332 mg·mL^{-1}·cm^{-1} at 280 nm, which was calculated from its amino acid sequence. Crystallization of Csh2 was attempted at 295 K by the sitting-drop vapor-diffusion method. The initial crystals were obtained from a precipitant solution No. 32 of PEG/Ion2 (Hampton research, Riverside, California, USA) containing 12% (w/v) Polyethylene Glycol 3350, 2% (v/v) Tacsimate at pH 5.0, and 100 mM Sodium Citrate tribasic dihydrate at pH 5.6. Crystals suitable for the X-ray diffraction experiment were obtained within 10 days at 294 K using hanging-drop vapor-diffusion method, where 1 μl protein sample and 1 μl precipitant were mixed and equilibrated with 0.2 ml precipitant solution containing 10% (w/v) Polyethylene Glycol 3350, 2% (v/v) Tacsimate at pH 5.0, and 100 mM Sodium Citrate tribasic dihydrate at pH 5.6.

X-ray diffraction data collection

For X-ray diffraction experiment, crystals were briefly immersed into precipitant solution containing 25% (v/v) Ethylene Glycol and immediately placed in a 95 K nitrogen-gas stream. X-ray diffraction data using a native Csh2 crystal were collected on the beamline 5C SBII at Pohang Accelerator Laboratory (PAL, Korea) with per-frame oscillation 0.5° and per-frame exposure of 1 s. 720 images were collected on Pilatus6M detector from a single crystal (Park et al, 2017). The collected data were processed using HKL-2000 (Otwinowski & Minor, 1997).

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by Basic Science Research (NRF Grant Number: 2017R1D1A3B03032278) of the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology of Korea (MEST).

Original Submission: Aug 25, 2017
Revised Version Received: Sep 11, 2017
Accepted: Sep 16, 2017

REFERENCES

Amitai, G., and Sorek, R. (2016). CRISPR-Cas adaptation: insights into the mechanism of action. *Nat Rev Microbiol* **14**, 67-76.

Arslan, Z., Hermanns, V., Wurm, R., Wagner, R., and Pul, Ü. (2014). Detection and characterization of spacer integration intermediates in type IE CRISPR–Cas system. *Nucleic Acids Res* **42**, 7884-7893.

Aslanidis, C., and de Jong, P.J. (1990). Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res* **18**, 6069-6074.

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.

Barrangou, R., and Marraffini, L.A. (2014). CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. *Molecular cell* **54**, 234-244.

Beloglazova, N., Petit, P., Flick, R., Brown, G., Savchenko, A., and Yakunin, A.F. (2011). Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference. *EMBO J* **30**, 4616-4627.

Bondy-Denomy, J., Garcia, B., Strum, S., Du, M., Rollins, M. F., Hidalgo-Reyes, Y., Wiedenheft, B., Maxwell, K.L., and Davidson, A.R. (2015). Multiple mechanisms for CRISPR–Cas inhibition by anti–CRISPR proteins. *Nature* **526**, 136.

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.

Charpentier, E., Richter, H., van der Oost, J., and White, M.F. (2015). Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiol Rev* **39**, 428-441.

Fineran, P.C., and Charpentier, E. (2012). Memory of viral infections by CRISPR-Cas adaptive immune systems: acquisition of new information. *Virology* **434**, 202-209.

Fu, B.X.H., Wainberg, M., Kundaje, A., and Fire, A.Z. (2017). High-Throughput Characterization of Cascade type IE CRISPR Guide Efficacy Reveals Unexpected PAM Diversity and Target Sequence Preferences. *Genetics* **206**, 1727-1738.

Gong, B., Shin, M., Sun, J., Jung, C.H., Bolt, E.L., van der Oost, J., and Kim, J.S. (2014). Molecular insights into DNA interference by CRISPR-associated nuclease-helicase Cas3. *Proc Natl Acad Sci U S A* **111**, 16359-16364.

Horvath, P., and Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**, 167-170.

Huo, Y., Nam, K.H., Ding, F., Lee, H., Wu, L., Xiao, Y., Farchione, MD, Jr., Zhou, S., Rajashankar, K., Kurinov, I., Zhang, R., and Ke. A. (2014). Structures of CRISPR Cas3 offer mechanistic insights into Cascade-activated DNA unwinding and degradation. *Nat Struct Mol Biol* **21**, 771-777.

Jore, M.M., Lundgren, M., van Duijn, E., Bultema, J.B., Westra, E.R., Waghmare, S.P., Wiedenheft, B., Pul, U., Wurm, R., Wagner, R., Beijer, M.R., Barendregt, A., Zhou, K., Snijders, A.P., Dickman, M.J., et al. (2011). Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat Struct Mol Biol* **18**, 529-536.

Kiss, H., Cleland, D., Lapidus, A., Lucas, S., Del Rio, T.G., Nolan, M., Tice, H., Han, C., Goodwin, L., Pitluck, S., Liolios, K., Ivanova, N., Mavromatis, K., Ovchinnikova, G., Pati, A., et al. (2010). Complete genome sequence of 'Thermobaculum terrenum'type strain (YNP1 T). *Stand Genomic Sci* **3**, 153-162.

Koonin, E.V., Makarova, K.S., and Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Curr Opin Microbiol* **37**, 67-78.

Kunin, V., Sorek, R., and Hugenholtz, P. (2007). Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol* **8**, R61.

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J., Wolf, Y.I., Yakunin, A.F., van der Oost, J., and Koonin, E.V. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**, 467-477.

Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J., Charpentier, E., Haft, D.H., Horvath, P., Moineau, S., Mojica, F.J., Terns, R.M., Terns, M.P., et al. (2015). An updated evolutionary classification of CRISPR–Cas systems. *Nat Rev Microbiol* **13**, 722-736.

Matthews, B.W. (1968). Solvent content of protein crystals. *J Mol Biol* **33**, 491-497.

Mulepati, S., and Bailey, S. (2011). Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3). *J Biol Chem* **286**, 31896-31903.

Müller-Esparza, H., and Randau, L. (2017). Comment: Type I CRISPR-Cas targets endogenous genes and regulates virulence to evade mammalian host immunity. *Front Microbiol* **8**, 319.

Nam, K.H., Haitjema, C., Liu, X., Ding, F., Wang, H., DeLisa, M.P., and Ke, A. (2012). Cas5d protein processes pre-crRNA and assembles into a cascade-like interference complex in subtype IC/Dvulg CRISPR-Cas system. *Structure* **20**, 1574-1584.

Otwinowski, Z., and Minor, W. (1997). [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307-326.

Park, S.Y., Ha, S.C., and Kim, Y.G. (2017). The Protein Crystallography Beamlines at the Pohang Light Source II.

Patterson, A.G., Chang, J.T., Taylor, C., and Fineran, P.C. (2015). Regulation of the Type IF CRISPR-Cas system by CRP-cAMP and GalM controls spacer acquisition and interference. *Nucleic Acids Res* **43**, 6038-6048.

Pawluk, A., Bondy-Denomy, J., Cheung, V.H., Maxwell, K.L., and Davidson, A.R. (2014). A new group of phage anti-CRISPR genes inhibits the type IE CRISPR-Cas system of Pseudomonas aeruginosa. *MBio* **5**, e00896-00814.

Plagens, A., Richter, H., Charpentier, E., and Randau, L. (2015). DNA and RNA interference mechanisms by CRISPR-Cas surveillance complexes. *FEMS Microbiol Rev* **39**, 442-463.

Richter, C., Dy, R.L., McKenzie, R.E., Watson, B.N., Taylor, C., Chang, J.T., McNeil, M.B., Staals, R.H., and Fineran, P.C. (2014). Priming in the Type IF CRISPR-Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. *Nucleic Acids Res* **42**, 8516-8526.

Sampson, T.R., Saroj, S.D., Llewellyn, A.C., Tzeng, Y.L., and Weiss, D.S. (2013). A CRISPR-Cas system mediates bacterial innate immune evasion and virulence. *Nature* **497**, 254-257.

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.

Savitskaya, E., Semenova, E., Dedkov, V., Metlitskaya, A., and Severinov, K. (2013). High-throughput analysis of type IE CRISPR/Cas spacer acquisition in E. coli. *RNA Biol* **10**, 716-725.

Semenova, E., Savitskaya, E., Musharova, O., Strotskaya, A., Vorontsova, D., Datsenko, K. A., Logacheva, M.D., and Severinov, K. (2016). Highly efficient primed spacer acquisition from targets destroyed by the Escherichia coli type IE CRISPR-Cas interfering complex. *Proc Natl Acad Sci U S A* **113**, 7626-7631.

Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J* **30**, 1335-1342.

Westra, E.R., Semenova, E., Datsenko, K.A., Jackson, R.N., Wiedenheft, B., Severinov, K., and Brouns, S.J. (2013). Type IE CRISPR-cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. *PLoS Genet* **9**, e1003742.

Westra, E.R., van Erp, P.B., Künne, T., Wong, S.P., Staals, R.H., Seegers, C.L., Bollen, S., Jore, M.M., Semenova, E., Severinov, K., de Vos, W.M., Dame, R.T., de Vries, R., Brouns, S.J., and van der Oost, J. (2012). CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol Cell* **46**, 595-605.

Wright, A.V., Liu, J.J., Knott, G.J., Doxzen, K.W., Nogales, E., and Doudna, J.A. (2017). Structures of the CRISPR genome integration complex. *Science* **357**, 1113-1118.

