CHAPTER SEVENTEEN

In Vitro Reconstitution and Crystallization of Cas9 Endonuclease Bound to a Guide RNA and a DNA Target

Carolin Anders, Ole Niewoehner, Martin Jinek[1]
Department of Biochemistry, University of Zurich, Zurich, Switzerland
[1]Corresponding author: e-mail address: jinek@bioc.uzh.ch

Contents

1. Introduction 516
2. Electrophoretic Mobility Shift Assay 518
3. Fluorescence-Detection Size Exclusion Chromatography Assay 523
4. Crystallization of Cas9–RNA–DNA Complexes 530
5. Concluding Remarks 534
Acknowledgments 535
References 535

Abstract

The programmable RNA-guided DNA cleavage activity of the bacterial CRISPR-associated endonuclease Cas9 is the basis of genome editing applications in numerous model organisms and cell types. In a binary complex with a dual crRNA:tracrRNA guide or single-molecule guide RNA, Cas9 targets double-stranded DNAs harboring sequences complementary to a 20-nucleotide segment in the guide RNA. Recent structural studies of the enzyme have uncovered the molecular mechanism of RNA-guided DNA recognition. Here, we provide protocols for electrophoretic mobility shift and fluorescence-detection size exclusion chromatography assays used to probe DNA binding by Cas9 that allowed us to reconstitute and crystallize the enzyme in a ternary complex with a guide RNA and a *bona fide* target DNA. The procedures can be used for further mechanistic investigations of the Cas9 endonuclease family and are potentially applicable to other multicomponent protein–nucleic acid complexes.

Methods in Enzymology, Volume 558
ISSN 0076-6879
http://dx.doi.org/10.1016/bs.mie.2015.02.008

© 2015 Elsevier Inc.
All rights reserved.

1. INTRODUCTION

Cas9 is an RNA-guided endonuclease capable of cleaving double-stranded DNA. The substrate specificity of Cas9 is determined by a 20-nucleotide sequence within a bound guide RNA that directs the enzyme to complementary sequences in dsDNA through base-pairing interactions (Gasiunas, Barrangou, Horvath, & Siksnys, 2012; Jinek et al., 2012). The enzyme originates from bacterial type II CRISPR–Cas (clustered regularly interspaced palindromic repeats–CRISPR-associated) genome defense systems, in which it associates with CRISPR RNA (crRNA) guides and a trans-activating crRNA (tracrRNA) to target the DNA of invading viruses and other mobile genetic elements (Barrangou et al., 2007; Deltcheva et al., 2011; Garneau et al., 2010). Based on the dual crRNA:tracrRNA guide structure, single-molecule guide RNAs (sgRNAs) have been engineered to program the sequence specificity of Cas9 for DNA cleavage *in vitro* and *in vivo* (Cong et al., 2013; Jinek et al., 2012, 2013; Mali, Yang, et al., 2013). The resulting Cas9–sgRNA two-component system has been harnessed to form the core of emerging genome editing technologies, whereby double-strand breaks induced in a specific genomic locus by the Cas9–sgRNA complex are repaired by the nonhomologous end joining or homologous recombination pathways to engineer a desired modification in the target locus (Cong et al., 2013; Jinek et al., 2013; Mali, Yang, et al., 2013). Owing to its easy programmability and a high degree of specificity, CRISPR-Cas9 has brought about revolutionary advances in genetic engineering in numerous model organisms and cell types (reviewed in Doudna & Charpentier, 2014; Hsu, Lander, & Zhang, 2014; Mali, Esvelt, & Church, 2013). The Cas9–sgRNA system has also capabilities that go beyond genome modifications. Catalytically inactive variants of Cas9 (dCas9) that retain RNA-guided DNA-binding activity have been used for RNA-guided transcriptional control (Gilbert et al., 2014, 2013; Konermann et al., 2015; Mali, Aach, et al., 2013) and for marking genomic loci in imaging applications (Chen et al., 2013). Cas9 has also been adapted for RNA binding and cleavage (O'Connell et al., 2014). CRISPR-Cas9 thus represents a transformative technology with many promising applications not just in basic research but also in biotechnology and medicine.

At the molecular level, Cas9 associates with the guide RNA (either the naturally occurring crRNA:tracrRNA duplex structure or an engineered sgRNA) in a binary protein–RNA complex that uses a 20-nucleotide region

at the 5′ end of the guide RNA to locate a matching sequence in a dsDNA target (Jinek et al., 2012; Karvelis et al., 2013). Watson–Crick base pairing between the guide sequence and the target DNA results in DNA strand separation and the formation of an RNA–DNA heteroduplex in an R-loop structure, positioning the two strands of the target DNA for site-specific cleavage. Cas9 contains two magnesium-dependent nuclease domains: an HNH domain that cleaves the complementary (target) strand of the DNA target, and a RuvC domain responsible for cleavage of the displaced non-complementary (non-target) DNA strand (Chen, Choi, & Bailey, 2014; Gasiunas et al., 2012; Jinek et al., 2012). DNA binding requires the presence of a short protospacer adjacent motif (PAM) in the vicinity of the target region in the DNA (Jinek et al., 2012; Sternberg, Redding, Jinek, Greene, & Doudna, 2014). PAM recognition is an initial step in target DNA binding that enables the Cas9–guide RNA complex to catalyze local strand separation in the target DNA and sequential guide RNA–target DNA heteroduplex formation (Sternberg et al., 2014). As a result, DNA sequences perfectly complementary to the guide RNA sequence but lacking the PAM are not recognized as cleavage substrates by Cas9 (Sternberg et al., 2014). Additionally, near-perfect complementarity between the guide RNA and the target DNA strand within a 8–12 base pair (bp) PAM-proximal "seed" region in the guide–target heteroduplex is a critical determinant of target DNA binding (Jinek et al., 2012). However, efficient DNA cleavage *in vivo* requires more extensive guide–target base pairing (Cencic et al., 2014; Kuscu, Arslan, Singh, Thorpe, & Adli, 2014; Wu et al., 2014). Mismatches within the guide–target heteroduplex are nevertheless tolerated in some positions, and this is the chief source of off-target activities in Cas9-based gene targeting applications (Fu et al., 2013; Hsu et al., 2013; Mali, Aach, et al., 2013; Pattanayak et al., 2013).

Structural studies have shed light on the molecular architecture and mechanism of Cas9. Crystal structures of apo-Cas9 defined the two structural lobes in the molecule, while crystal structures and electron microscopic reconstructions of nucleic acid-bound complexes revealed an extensive RNA-driven conformational rearrangement that primes the Cas9–guide RNA complex for target DNA binding (Jinek et al., 2014; Nishimasu et al., 2014). To obtain structural insights into the molecular mechanism of PAM-dependent target DNA recognition by Cas9, we recently determined the crystal structure of *Streptococcus pyogenes* Cas9 (SpyCas9) in complex with an sgRNA guide and a *bona fide* target DNA containing a canonical 5′-NGG-3′ PAM sequence (Anders, Niewoehner, Duerst, & Jinek, 2014).

The structure revealed that the PAM sequence is recognized by direct readout of the GG dinucleotide in the major groove of the PAM-containing DNA duplex. Furthermore, the structure suggested how PAM recognition by the Cas9–guide RNA complex might be coupled to strand separation in the target dsDNA and concomitant formation of the guide RNA–target DNA heteroduplex.

In this chapter, we present biochemical methods that we used to establish a robust protocol for the *in vitro* reconstitution of the Cas9–sgRNA–DNA complex and outline the procedure and considerations for its crystallization. To probe target DNA binding, we initially used electrophoretic mobility shift assays (EMSAs). We subsequently devised an analytical size exclusion chromatography (SEC) assay employing fluorescently labeled DNA oligonucleotides to identify minimal sgRNA and target DNA structures capable of supporting complex formation. These assays can be readily used to analyze target DNA binding by SpyCas9 as well as orthologous Cas9 proteins from other bacterial species. More generally, the methods may be applicable to other multicomponent protein–nucleic acid assemblies. The fluorescence-detection size exclusion chromatography (FSEC) assay permits high-throughput screening of multiple nucleic acid constructs and is particularly advantageous in situations where the analyzed macromolecular complex contains two or more nucleic acid molecules.

2. ELECTROPHORETIC MOBILITY SHIFT ASSAY

EMSAs are commonly used to characterize protein–nucleic acid interactions. These assays rely on the ability to separate bound and unbound fractions of a labeled nucleic acid probe based on their differential mobility through a nondenaturing polyacrylamide or agarose matrix, which is dependent on the molecular masses of the free nucleic acid and the bound protein, the shape of the complex, and overall electrical charge. Typically, nucleic acid probes used in EMSAs are radiolabeled at their 5′ ends. In this way, minute quantities of the nucleic acid can be used in the assay, and apparent dissociation constants can be determined by quantifying the ratio of bound and unbound molecules and fitting the data to a standard binding isotherm. The theory and the general considerations for practical implementation of EMSAs have been reviewed extensively elsewhere (Kerr, 1995; Mitchell & Lorsch, 2014; Ryder, Recht, & Williamson, 2008).

In our study, we sought to establish assays to probe target DNA binding by the Cas9–sgRNA complex. In an initial approach, we performed EMSAs

using the catalytically inactive double mutant (D10A/H840A) of the *S. pyogenes* Cas9 protein (dCas9) (Fig. 1A and B). dCas9 was expressed and purified as described previously (Anders & Jinek, 2014; Jinek et al., 2012) from expression plasmid pMJ841 (available from Addgene). The sgRNA was prepared by *in vitro* transcription using T7 RNA polymerase and a fully double-stranded oligonucleotide DNA template, and was purified by denaturing polyacrylamide gel electrophoresis and ethanol precipitation (Anders & Jinek, 2014). The sgRNA consisted of a 5′-terminal 20-nucleotide guide segment, followed by a repeat:antirepeat duplex (capped with a GAAA tetraloop) and stem loops SL1 and SL2 in the 3′-terminal region. These features are known to be required for efficient DNA cleavage activity by Cas9 *in vitro* as well as *in vivo* (Briner et al., 2014; Hsu et al., 2013; Jinek et al., 2012). The target DNA was a 28-bp synthetic oligonucleotide duplex containing a sequence perfectly complementary to the guide segment of the sgRNA and a canonical PAM (5′-TGG-3′) in the non-target (i.e., noncomplementary) strand. The non-target strand in the duplex was labeled with the ATTO532 fluorophore covalently attached to its 3′ end. We decided to switch from radiolabeling to fluorescence detection for greater convenience, as custom fluorophore-labeled synthetic DNA oligonucleotides are readily available from commercial sources. The ATTO532 label was chosen for its photostability and superior quantum yield, allowing us to use the labeled duplex at a final concentration of 50 n*M* in our binding reactions. Although the working concentration is greater than the apparent dissociation constant of the DNA-bound Cas9–sgRNA complex, as measured previously using radiolabeled DNA probes (Jinek et al., 2012; Sternberg et al., 2014), and therefore does not permit accurate quantitation of binding affinities, the assay format nevertheless provides a reasonable indication of specific target DNA binding (Anders et al., 2014). In particular, the assay can be used to probe the PAM dependency of binding, since the absence of a cognate PAM interaction increases the apparent K_d by at least three orders of magnitude (Anders et al., 2014; Jinek et al., 2012; Sternberg et al., 2014).

In the EMSA, the target DNA duplex is titrated with increasing amounts of preassembled dCas9–sgRNA complex spanning a concentration range from 10 n*M* to 1 μ*M*; a control reaction lacking the dCas9–sgRNA complex is also set up. The sequences of sgRNA and DNA oligonucleotides used in the assay are shown in Fig. 1B and listed in Table 1. The electrophoresis runs were performed in a cold room at 4 °C to minimize complex dissociation and resulting smearing of the gel bands. The following protocol describes the detailed procedure for the EMSA shown in Fig. 1C.

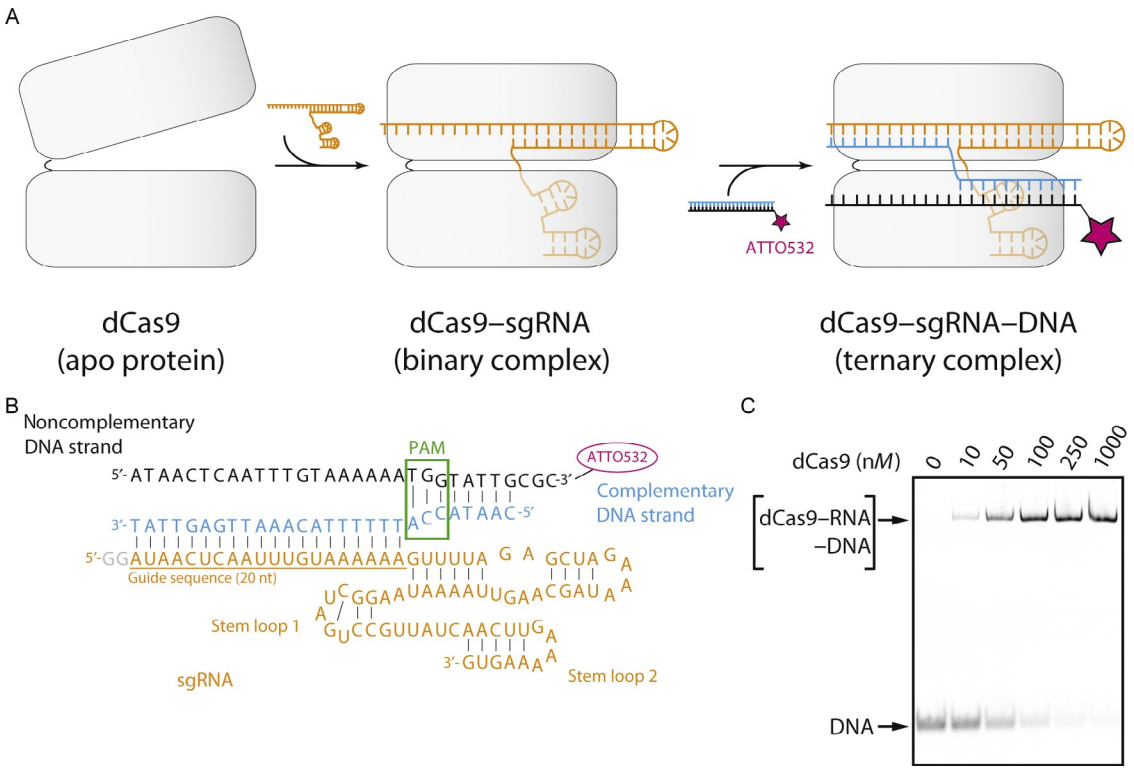


Figure 1 See legend on opposite page.

Table 1 DNA and RNA Sequences Used in EMSA

Description	Sequence[a]
sgRNA	5′-GGATAACTCAATTTGTAAAAAAGTTTTAGA GCTAGAAATAGCAAGTTAAAATAAGGCTAG TCCGTTATCAACTTGAAAAAGTG-3′
ssDNA template sgRNA	5′-CACTTTTTCAAGTTGATAACGGACTAGCCT TATTTTAACTTGCTATTTCTAGCTCTAAAA CTTTTTTACAAATTGAGTTATCCTATAGTG AGTCGTATTA-3′
Reverse primer for PCR amplification of ssDNA template for sgRNA	5′-CACTTTTTCAAGTTGA-3′
Complementary strand	5′-CAATA*CCA*TTTTTTACAAATTGAGTTAT-3′
Noncomplementary strand 1 3′-ATTO532	5′-AAAA**TGG**TATTGCGC-3′-ATTO532
Noncomplementary strand 2 3′-ATTO532	5′-ATAACTCAATTTGTAAAAAA**TGG**TATTGCGC-3′-ATTO532

[a]Nontarget strand PAM sites (5′-NGG-3′) are highlighted in bold. Target-strand complement of the PAM (5′-CCN-3′) is denoted in italic. Target sequence complementary to the guide RNA is underlined.

Figure 1—Cont'd (A) Schematic workflow of the electrophoretic mobility shift assay (EMSA) to probe for ternary complex formation. In a first step, apo-dCas9 and guide RNA are preincubated to form a binary complex. Addition of annealed duplex substrate leads to the assembly of ternary dCas9–guide RNA–target DNA complex. (B) Nucleic acid components used for the EMSA shown in panel (C). sgRNA is shown in orange (light gray in the print version) with the 5′-terminal GG dinucleotide originating from *in vitro* transcription using T7 RNA polymerase marked in gray. The complementary and noncomplementary DNA strands are colored in blue (light gray in the print version) and black, respectively. The noncomplementary strand is labeled at its 3′ end with an ATTO532 fluorophore. The PAM sequence and its complement are marked with a green (gray in the print version) box. (C) Target DNA duplex (50 n*M*) was titrated with increasing concentrations (0, 10, 50, 100, 250, and 1000 n*M*) of *in vitro* reconstituted dCas9–sgRNA complex. Binding reactions were analyzed using a native 8% polyacrylamide gel and visualized by detection of ATTO532 fluorescence using a laser scanner. The lower and upper bands represent the unbound and complex-bound DNA fractions, respectively.

1. Obtain synthetic oligonucleotides for the target and non-target strands of the DNA duplex. The non-target strand carries the ATTO532 label attached to the 3′-hydroxyl group. The labeled oligonucleotides should be PAGE purified prior to use. This can be done by the supplier or alternatively performed in the laboratory using a published protocol (Lopez-Gomollon & Nicolas, 2013). Determine oligonucleotide concentrations by measuring the absorbance at 260 nm. We calculate DNA concentrations using the OligoCalc server (Kibbe, 2007). Dilute both oligonucleotides to make 20 μ*M* stock solutions (20 μl each).
2. Prepare the oligonucleotide DNA duplex probe by annealing target and nontarget strands in a 1.5:1 (target:nontarget) molar ratio to ensure complete hybridization of the labeled strand. In order to do this, mix 15 μl of the complementary strand stock solution (20 μ*M*) with 10 μl of the non-complementary strand. Heat the mixture to 95 °C for 5 min and cool slowly to room temperature. Dilute the mixture with 975 μl binding buffer (20 m*M* HEPES, pH 7.5, 250 m*M* KCl, 2 m*M* $MgCl_2$, 0.01% Triton X-100, 0.1 mg ml^{-1} bovine serum albumin, 10% glycerol) to obtain 200 n*M* stock solution of the labeled DNA duplex.
3. Prepare a native 8% polyacrylamide gel containing 1 × TBE (89 m*M* tris(hydroxymethyl)aminomethane (Tris), 89 m*M* boric acid, 2 m*M* ethylenediaminetetraacetic acid, pH 8.0) supplemented with 2 m*M* $MgCl_2$. We typically pour gels with the dimensions of 300 mm × 180 mm × 1 mm (h × w × d) and use a well comb with 6 mm tooth width and 2 mm spacing.
4. Thaw a fresh aliquot of frozen dCas9 and dilute with storage buffer (20 m*M* HEPES, pH 7.5, 500 m*M* KCl) to prepare a 100 μ*M* stock solution (50 μl). Prepare a stock solution of purified sgRNA by diluting to 400 μ*M* with nuclease-free water. To reconstitute the dCas9–sgRNA complex in a 1:1.5 molar ratio, mix 8.0 μl dCas9 with 3.0 μl sgRNA and incubate for 10 min at room temperature. Dilute the mixture with 189 μl binding buffer to a final concentration of 4 μ*M* and store on ice. Use the 4 μ*M* complex solution to prepare additional dilutions at 1 μ*M* (25 μl) and 200 n*M* (20 μl).
5. For each binding reaction, combine binding buffer and dCas9–sgRNA complex according to Table 2 and incubate for 10 min at room temperature. Add 5 μl oligonucleotide DNA duplex to each reaction and incubate for 10 min at 37 °C (Fig. 1A).
6. Load a 5 μl sample of each binding reaction onto the polyacrylamide gel. Load an additional lane with 5 μl binding buffer supplemented with

Table 2 Electrophoretic Mobility Shift Assay

Final dCas9–sgRNA Concentration	Binding Buffer (µl)	dCas9–sgRNA Mix[a]	Duplex Substrate (µl)[a]	Total Volume (µl)
0 n*M*	15	–	5	20
10 n*M*	14	1 µl 200 n*M*	5	20
50 n*M*	10	5 µl 200 n*M*	5	20
100 n*M*	13	2 µl 1 µ*M*	5	20
250 n*M*	10	5 µl 1 µ*M*	5	20
1 µ*M*	10	5 µl 4 µ*M*	5	20

[a]The target and nontarget DNA strands are annealed prior addition to the reaction. See Section 2, step 2 for the annealing procedure.

0.005% bromophenol blue to be able to monitor the progress of the electrophoresis run. Run the gel using 1 × TBE buffer supplemented with 2 m*M* $MgCl_2$ at 15 W at 4 °C for approximately 75 min, until the bromophenol blue dye front has reached halfway down the gel.

7. Scan the gel using a laser gel scanner (e.g., Typhoon FLA9500, GE Healthcare) with the appropriate excitation and emission wavelength settings for ATTO532 (532 nm laser for excitation and LPG emission filter for detection, respectively).

3. FLUORESCENCE-DETECTION SIZE EXCLUSION CHROMATOGRAPHY ASSAY

Successful crystallization of multicomponent macromolecular assemblies requires extensive coverage of the crystallization parameter space, varying not just the chemical composition of the crystallization buffer but also the precise forms of the components of the crystallized macromolecular complex. In the case of protein–nucleic acid complexes, variation of the sequence and length of the bound nucleic acids is an essential part of the crystallization strategy (Hollis, 2007; Obayashi, Oubridge, Pomeranz Krummel, & Nagai, 2007; Reyes, Garst, & Batey, 2009). This is because specific bases of the nucleic acid ligand often mediate crystal contact formation through base pairing, base stacking, and other tertiary interactions (Anderson, Ptashne, & Harrison, 1984, 1987). In some cases, this has been exploited to engineer crystal contacts for otherwise recalcitrant molecules (Berry, Waghray, Mortimer, Bai, & Doudna, 2011; Ferré-D'Amaré,

Zhou, & Doudna, 1998; Leung et al., 2010). However, this crystallization strategy relies on the ability to efficiently assess complex formation for a broad range of nucleic acid ligands. Although EMSAs are well suited to probe a specific protein–nucleic acid interaction and provide quantitative information about binding, this approach is very time consuming because the analysis of multiple nucleic acid ligands typically requires running many gels. Additionally, although EMSA can differentiate between bound monomers and multimers, the method otherwise yields few clues about the nature of the protein–nucleic acid complex. Finally, it is additionally difficult to determine whether binding is quantitative (i.e., whether all or only a fraction of the protein is active in binding to the ligand) and therefore whether stoichiometric amounts of complex can be reconstituted in large scale in order to prepare sufficient amounts of material for crystallization screening. Analytical SEC is an attractive alternative to EMSA because it provides additional information on the approximate molecular mass of the complex and the degree of its monodispersity (Winzor, 2003). Both protein and nucleic acid components can be monitored due to their UV absorption at 280 and 260 nm, respectively, and complex formation manifests in shifts in the elution volumes of the corresponding peaks and changes in their A_{260}/A_{280} ratios. By comparing the ratio of areas under the peaks in the chromatogram, the relative fractions of protein and/or nucleic acid ligand incorporated into the complex can be determined.

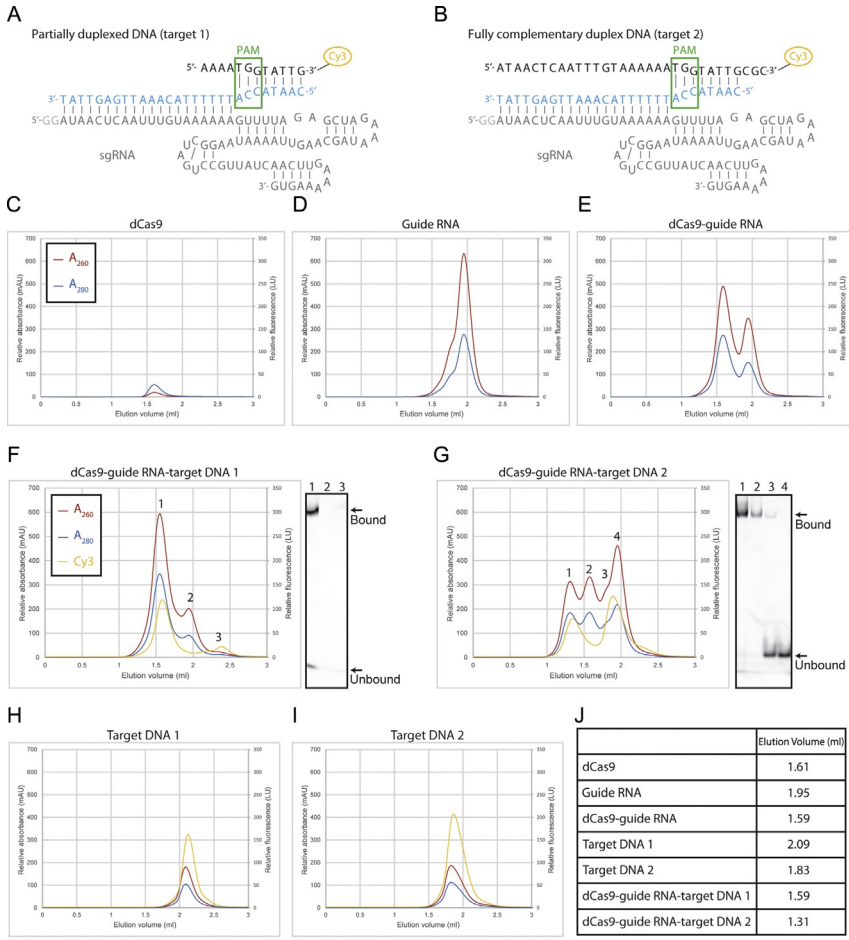
To obtain a crystal structure of Cas9 bound to a guide RNA and a PAM-containing DNA target, our efforts focused exclusively on SpyCas9 due to the availability of the crystal structure of the apo protein and EM reconstructions of its nucleic acid complexes. Aiming to generate dCas9–sgRNA–target DNA complexes suitable for crystallization, we sought to define the minimal features in the sgRNA and the target DNA that would support stable complex formation. To this end, we used analytical SEC to screen a panel of sgRNA structures and target DNAs. Due to the need to be able to monitor the simultaneous binding of two nucleic acid ligands to dCas9 (the sgRNA and the target DNA duplex), we devised a fluorescence-detection strategy in which the target DNA duplex carried a Cy3 fluorophore covalently attached to the 3′ end of the non-target strand. The sgRNA, being the larger of the two nucleic acid components in the ternary complex, was monitored by observing changes in the peak profile in the 260 nm channel. Our FSEC assay utilized an Agilent 1200 series high-performance liquid chromatography (HPLC) instrument equipped with UV–vis and single-channel fluorescence detectors connected in series. We analyzed samples on a

Table 3 Summary of Parameters for FSEC Assay Using a Superdex 200 5/150 Column

Column volume	3 ml
Flow rate	0.1 ml min^{-1}
Pressure limit	15 bar
Excitation wavelength	550 nm
Emission wavelength	570 nm
Maximum injection volume	50 μl
Minimum detectable amount of protein	125 pmol or 20 μg

Superdex 200 5/150 size exclusion column with a 3-ml bed volume and 15 bar pressure limit, eluting at a rate of 0.1 ml min^{-1} (see Table 3 for a summary of parameters). Using this method, as little as 20 μg of dCas9 can be analyzed in a single run within a span of 30 min. The availability of an autosampler allowed us to perform multiple runs in an automated, high-throughput manner. The eluate was fractionated and individual fractions were subsequently analyzed by native polyacrylamide gel electrophoresis.

Previous studies suggested that the Cas9–sgRNA complex recognizes the PAM on the non-target strand of the DNA target (Jinek et al., 2012; Sternberg et al., 2014). Furthermore, the presence of a 5′-truncated non-complementary strand that contained only the 5′-NGG-3′ PAM sequence in the partially duplexed target DNA was sufficient to stimulate cleavage of the complementary strand, suggesting that a truncated non-target strand is sufficient for stable complex formation (Sternberg et al., 2014). We sought to take advantage of these observations to prepare a minimal Cas9–sgRNA–target DNA ternary complex suitable for crystallization. To this end, we tested the binding of dCas9 complexed with a truncated 83-nucleotide sgRNA (containing stem loops SL1 and SL2 in the 3′-terminal region) to two target DNA duplexes that contained a 5′-TGG-3′ PAM sequence. The target DNA strand was identical in the two DNA ligands, whereas the non-target strand was either truncated at its 5′ end to mimic the Cas9 cleavage product (12 nt, Fig. 2A) or fully complementary to the target strand (31 nt, Fig. 2B). Control runs established that apo-dCas9 and unbound sgRNA eluted at 1.61 and 1.95 ml, respectively (Fig. 2C and D). Incubation of dCas9 with a slight excess of sgRNA led to efficient formation of the binary complex, as judged by a change in the A_{260}/A_{280} ratio of the 1.59-ml peak and proportional decrease in the size of the free sgRNA peak



J

	Elution Volume (ml)
dCas9	1.61
Guide RNA	1.95
dCas9-guide RNA	1.59
Target DNA 1	2.09
Target DNA 2	1.83
dCas9-guide RNA-target DNA 1	1.59
dCas9-guide RNA-target DNA 2	1.31

Figure 2 (A and B) Schematic representation of two target DNA variants used in fluorescence-detection size exclusion chromatography (FSEC). The complementary DNA strand (blue (light gray in the print version)) is identical in both DNA targets. The 3′-Cy3-labeled noncomplementary DNA strands (black) differ in length (panel A: target DNA 1, 15 nt; panel B: target DNA 2, 31 nt) to form either a partially or fully double-stranded target substrate. The PAM sequence and its complement are marked with a green (gray in the print version) box. The sgRNA is colored gray. (C–I) FSEC chromatograms for apo-dCas9 (C), sgRNA (D), binary dCas9–sgRNA complex (E), the ternary dCas9–sgRNA–target DNA 1 complex (F), ternary dCas9–sgRNA–target DNA 2 complex (G), target DNA 1 alone (H), and target DNA 2 alone (I). The measured absorbances at 260 and 280 nm are shown with red (black in the print version) and blue (light gray in the print version) lines, respectively. The yellow (light gray in the print version) line depicts the Cy3 fluorescence signal. For the ternary complexes, 20 µl of the labeled peak fractions were analyzed on a native polyacrylamide gel and detected by the Cy3 fluorescence (F and G). (J) Summary of elution volumes from FSEC experiments in (C–I).

(Fig. 2E). The observation that the elution volumes of apo-dCas9 and the dCas9–sgRNA complex were nearly identical was consistent with previous electron microscopic studies indicating that the overall molecular dimensions of the two species are quite similar, despite a major conformational rearrangement in the Cas9 protein upon sgRNA binding (Jinek et al., 2014). Addition of the target DNA duplex containing a truncated noncomplementary strand led to near-stoichiometric formation of the ternary dCas9–sgRNA–target DNA complex, as indicated by a shift of the Cy3 signal from the unbound peak at 2.09 ml to the complex peak at 1.59 ml (Fig. 2F and H). In contrast, addition of the fully complementary target DNA duplex also resulted in ternary complex formation, but only ~50% of the Cy3 label was incorporated in the ternary complex peak (Fig. 2G and I). Taken together, these experiments suggested that both target DNA duplexes were bound by the dCas9–sgRNA complex. However, only in the case of the DNA target containing a truncated noncomplementary strand was ternary complex formation stoichiometric, whereas a substantial fraction of the fully complementary target DNA duplex remained unbound. It is possible that in this case, strand separation by Cas9 was slow, despite the presence of a canonical PAM sequence, which precluded efficient formation of the target DNA-bound complex. Moreover, comparison of the elution volumes of the complexes containing the two target DNA duplexes (1.31 ml for the fully complementary target duplex and 1.59 ml for the partially duplexed target DNA, Fig. 2J) also suggested that the complex containing the partially duplexed target DNA likely adopted a more compact conformation. Together with the observed stoichiometric formation of the complex, these two factors were decisive in identifying the partially duplexed DNA target as the optimal choice for large-scale complex reconstitution and initial crystallization trials.

The following protocol describes the details of the FSEC assay. The sequences of the nucleic acids used in the assay are provided in Table 4. We recommend performing the procedure at 4 °C in a cold room, particularly if multiple complexes are to be analyzed.

1. Prepare 1 l of complex reconstitution buffer (20 m*M* HEPES, pH 7.5, 250 m*M* KCl, 2 m*M* $MgCl_2$). Set up the HPLC instrument and flush the pumps and flowpath with the complex reconstitution buffer.
2. Attach a Superdex 200 5/150 gel filtration column (GE Healthcare) to the HPLC instrument and equilibrate with complex reconstitution buffer at a flow rate of 0.1 ml min^{-1} (see Table 3 for a summary of parameters).

Table 4 DNA and RNA Sequences Used for FSEC Assays

Description	Sequence[a]
sgRNA	5′-GGATAACTCAATTTGTAAAAAAGTTTTAGA GCTAGAAATAGCAAGTTAAAATAAGGCTAG TCCGTTATCAACTTGAAAAAGTG-3′
ssDNA template sgRNA	5′-CACTTTTTCAAGTTGATAACGGACTAGCCT TATTTTAACTTGCTATTTCTAGCTCTAAAA CTTTTTTACAAATTGAGTTATCCTATAGTG AGTCGTATTA-3′
Reverse primer for PCR amplification of ssDNA template for sgRNA	5′-CACTTTTTCAAGTTGA-3′
Complementary strand	5′-CAATA*CCA*TTTTTTACAAATTGAGTTAT-3′
Noncomplementary strand 1 3′-Cy3	5′-AAAA**TGG**TATTG-3′-Cy3
Noncomplementary strand 2 3′-Cy3	5′-ATAACTCAATTTGTAAAAAA**TGG**TATTGCGC-3′-Cy3

[a]Nontarget strand PAM sites (5′-NGG-3′) are highlighted in bold. Target-strand complement of the PAM (5′-CCN-3′) is denoted in italic. Target DNA sequence complementary to the guide RNA is underlined.

3. Thaw a frozen stock solution of dCas9 and prepare 30 µl of a 100 µ*M* (~16 mg ml^{-1}) solution of dCas9 by diluting with storage buffer (20 m*M* HEPES, pH 7.5, 500 m*M* KCl).
4. Prepare 30 µl of a 375 µ*M* (~10 mg ml^{-1}) stock solution of sgRNA by diluting with water. RNA concentrations can be calculated using the OligoCalc server (Kibbe, 2007) from absorbance measured at 260 nm.
5. Mix 94 µl of complex reconstitution buffer with 2.5 µl 100 µ*M* dCas9. Add 1.0 µl 375 µ*M* of sgRNA (1.5-fold molar excess) to form the dCas9–sgRNA binary complex. Incubate for 10 min at room temperature. We recommend adding the sgRNA in 1.5-fold molar excess in order to ensure stoichiometric formation of the binary complex. The final concentrations of dCas9 and sgRNA in the reconstituted sample are 2.5 and 3.75 µ*M*, respectively.

6. In the meantime, prepare double-stranded DNA substrate by annealing complementary and noncomplementary DNA strands at equimolar concentration. To this end, mix 10 µl of each oligonucleotide stock solution (200 µ*M*), heat to 95 °C for 5 min, and cool slowly to room temperature.
7. Add 2.5 µl of 100 µ*M* annealed duplex substrate to the dCas9–sgRNA complex to a final concentration of 2.5 µ*M*. Incubate for 10 min at room temperature. The pipetting scheme of steps 5–7 is summarized in Table 5.
8. Centrifuge for 5 min at 14,000 rpm (16,900 × *g*) at 4 °C to remove any insoluble material from the sample prior to injection onto the Superdex 200 5/150 column.
9. Inject 50 µl of the sample to the equilibrated Superdex 200 5/150 column. Elute with 3 ml complex reconstitution buffer at a flow rate of 0.1 ml min^{-1}. Measure the UV absorbance of the eluate at 260 and 280 nm, and fluorescence by excitation at 550 nm and emission at 570 nm.
10. Fractionate by collecting 20 µl fractions in a 96-well microtiter plate. To analyze fractions by native polyacrylamide gel electrophoresis, add glycerol to each fraction to a final concentration of 5% glycerol and analyze 20 µl of the peak fractions on a native 8% polyacrylamide gel as described for EMSAs in Section 2. The gel can be scanned using a laser gel scanner (e.g., Typhoon FLA9500, GE Healthcare) with the appropriate excitation and emission wavelength settings for Cy3 (532 nm laser excitation and LPG emission filter) (Fig. 2).

Table 5 dCas9–sgRNA–Target DNA Complex Reconstitution for FSEC Assay

	Stock Concentration	Final Concentration	Volume (µl)[a]
Buffer	–	–	94
dCas9	100 µ*M*	2.5 µ*M*	2.5
sgRNA	375 µ*M*	3.75 µ*M*	1.0
Incubate for 10–15 min at room temperature. Then add annealed DNA duplex.			
Duplex substrate[a]	100 µ*M*	2.5 µ*M*	2.5
Total volume			100

[a]The target and nontarget DNA strands are annealed prior addition to the reaction. See Section 3, step 6 for the annealing procedure.

4. CRYSTALLIZATION OF CAS9–RNA–DNA COMPLEXES

Having identified partially duplexed target DNAs as the most promising candidates for initial crystallization trials, we proceeded to large-scale reconstitution of the ternary dCas9–sgRNA–DNA complex to generate milligram quantities of material for crystallization screening. We decided to follow a stepwise reconstitution procedure in which we first purified the binary dCas9–sgRNA complex, bound it to an excess of the target DNA duplex and repurified the resulting ternary complex by SEC. In initial reconstitution experiments, we attempted to reconstitute the dCas–sgRNA complex in 250 m*M* KCl. However, we found that the dCas9–sgRNA complex had limited solubility in low ionic strength solutions, and could not be concentrated to >1 mg ml^{-1} prior to SEC. For this reason, we repeated our FSEC assay using an elution buffer containing 500 m*M* KCl and found that the ternary complex could still be formed (data not shown). We therefore used 500 m*M* KCl in all subsequent purifications.

Our final purification procedure was as follows. dCas9 was recombinantly expressed and purified by nickel affinity and cation exchange chromatographic steps as described previously but the purification stopped short of the final SEC step (Anders & Jinek, 2014). To reconstitute the binary dCas9–sgRNA complex, cation exchange chromatography-purified Cas9 ($\sim$10 mg) was exchanged into SEC buffer (20 m*M* HEPES–KOH, pH 7.5, 500 m*M* KCl) by concentrating the protein to 5–10 mg ml^{-1} in a 50,000 molecular weight cut-off (MWCO) centrifugal filter at 3900 rpm ($\sim 3200 \times g$) and 4 °C, diluting with SEC buffer, and repeating the centrifugal concentration. The concentrated protein was mixed with *in vitro* transcribed sgRNA in a 1:1.5 molar ratio and incubated on ice for 10 min. The sample was centrifuged for 10 min at 14,000 rpm ($16{,}900 \times g$) at 4 °C to remove any insoluble material and injected onto a Superdex 200 16/600 column (equilibrated with SEC buffer) using a 2-ml injection loop. The complex was eluted with 120 ml of SEC buffer at a flow rate of 1 ml min^{-1}. The binary complex typically eluted at a volume of $\sim$69 ml. Peak fractions from the SEC step were pooled and the concentration of the binary complex was calculated from UV absorbance measured at 260 nm using the OlicoCalc server (assuming all of the absorbance was due to just the sgRNA component). $MgCl_2$ was added to a final concentration of 2 m*M* and the binary complex was combined with preannealed target DNA duplex added in 1.5-fold molar excess. The ternary complex sample was concentrated in a

50,000 MWCO centrifugal filter at 3900 rpm and 4 °C to <1.5 ml volume and subsequently centrifuged for 10 min at 14,000 rpm at 4 °C. The sample was again injected onto a Superdex 200 16/600 column equilibrated with SEC buffer 2 (20 m*M* HEPES–KOH, pH 7.5, 500 m*M* KCl, 2 m*M* $MgCl_2$) using a 2-ml injection loop and eluted with 120 ml SEC buffer 2 at a flow rate of 1 ml min^{-1} collecting 1 ml fractions. The ternary complex containing partially duplexed DNA target typically eluted at a volume of ~65 ml and had an A_{260}/A_{280} ratio of ~1.8. Peak fractions from the SEC were pooled and concentrated in a 50,000 MWCO centrifugal filter at 3900 rpm and 4 °C to 10–15 mg ml^{-1} (assuming all of the absorbance at 260 nm was due to just the nucleic acid components). The concentrate was recovered, centrifuged for 10 min at 14,000 rpm at 4 °C, and flash frozen in 50 µl aliquots for storage at −80 °C. For crystallization experiments, the KCl concentration in the sample was reduced to 250 m*M* by diluting 1:1 with 20 m*M* HEPES pH 7.5, 2 m*M* $MgCl_2$ and then adjusting the protein concentration by further dilution with 20 m*M* HEPES pH 7.5, 250 m*M* KCl, 2 m*M* $MgCl_2$.

In our crystallization strategy, we settled on a set of eleven 96-well screens (Table 6) at 20 and 4 °C and explored a range of target DNAs obtained by annealing a small set of target and nontarget strand oligonucleotides. The screens were set up using a Phoenix nanoliter pipetting robot in the 100 + 100 nl format and imaged using a Formulatrix Rock Imager automated system. This screening approach was successful for a dCas9–sgRNA complex containing a partially duplexed target containing a blunt end at the 5′ end of the target strand (Fig. 3A). Initial hits were obtained with a ternary complex sample at a concentration of 1.8 mg ml^{-1} in a range of PEG-based crystallization solutions containing potassium thiocyanate or lithium sulfate (Fig. 3B and C). The outcome of the crystal screening strategy was in agreement with previous studies suggesting that many proteins and multisubunit complexes can be crystallized using a relatively limited set of crystallization screens and that extensive variation of protein constructs and/or nucleic acid ligands is a much more effective way of ensuring successful crystallization (Kimber et al., 2003; Page, Deacon, Lesley, & Stevens, 2005; Thomsen & Berger, 2012). Our future structural studies of Cas9 enzymes will take this into account. To improve crystal size and quality for diffraction experiments, the crystals were subsequently grown using the hanging drop method using Qiagen EasyXtal 15-well trays. In this format, a grid screen vapor diffusion varying the PEG3350 concentration from 12% to 16% (in 1% steps) and KSCN from 0.2 *M* to 0.3 *M* (in 50 m*M* steps) was typically

Table 6 Crystallization Screens Used

Clear Strategy™ Screen I, pH 5.5, pH 6.5, pH 7.5, and pH 8.5 (Molecular Dimensions MD1-14)
Clear Strategy™ Screen II, pH 5.5, pH 6.5, pH 7.5, and pH 8.5 (Molecular Dimensions MD1-15)
Structure Screen I (Molecular Dimensions MD1-01) and Structure Screen II (Molecular Dimensions MD1-02)
Crystallization Basic Kit for Membrane Proteins (Sigma-Fluka Art#73513) and Crystallization Low Ionic Kit for Proteins (Sigma-Fluka #86684)
NeXtal Tubes PACT Suite (Qiagen Art#130718)
NeXtal Tubes Anions Suite (Qiagen Art#130707)
NeXtal Tubes Cations Suite (Qiagen Art#130708)
Crystal Screen™ Kit (Hampton research Art#HR2-110) and Crystal Screen 2™ (Hampton research Art#HR2-112)
PEG4000, 0.2 *M* Li_2SO_4, PEG8000 5–30%, pH gradient:3–10 (custom screen)
PEG4000, 0.2 *M* Li_2SO_4, PEG6000 5–30%, pH gradient:3–10 (custom screen)
PEG400 15–45%//PEG4000 5–30%, pH gradient 4.5–9.4, K-phosphate//Li-Na-sulfate (custom screen)
PEG400 15–45%//PEG4000 5–30%, pH gradient 4.5–9.4, ammonium formate; zinc acetate, potassium iodide (custom screen)

performed with a range of complex concentrations. The resulting thin plate crystals grew best from complex solutions at 1.5 mg ml^{-1} and reached their final size in 1–2 weeks.

For diffraction experiments, the crystals were cryoprotected by a brief transfer into 0.1 *M* Tris–acetate, pH 8.5, 200 m*M* KSCN, 30% PEG3350, 10% ethylene glycol, and flash-cooled in liquid nitrogen. Crystal diffraction was tested at the Swiss Light Source (Paul Scherrer Institut, Villigen, Switzerland). Initially, the crystals diffracted mostly to 3.5 Å, but the diffraction limit could be improved to 2.6 Å by increasing crystal thickness. This was achieved (i) by switching from the catalytically inactive dCas9 protein to the single-mutant (H840A) Cas9 nickase and (ii) by iterative microseeding, whereby seeds generated from an initial batch of complex crystals were used to grow the subsequent batch of crystals for seed generation, until large single crystals of the complex could be obtained reproducibly (Fig. 3D and E). We also found empirically that reversing the traditional

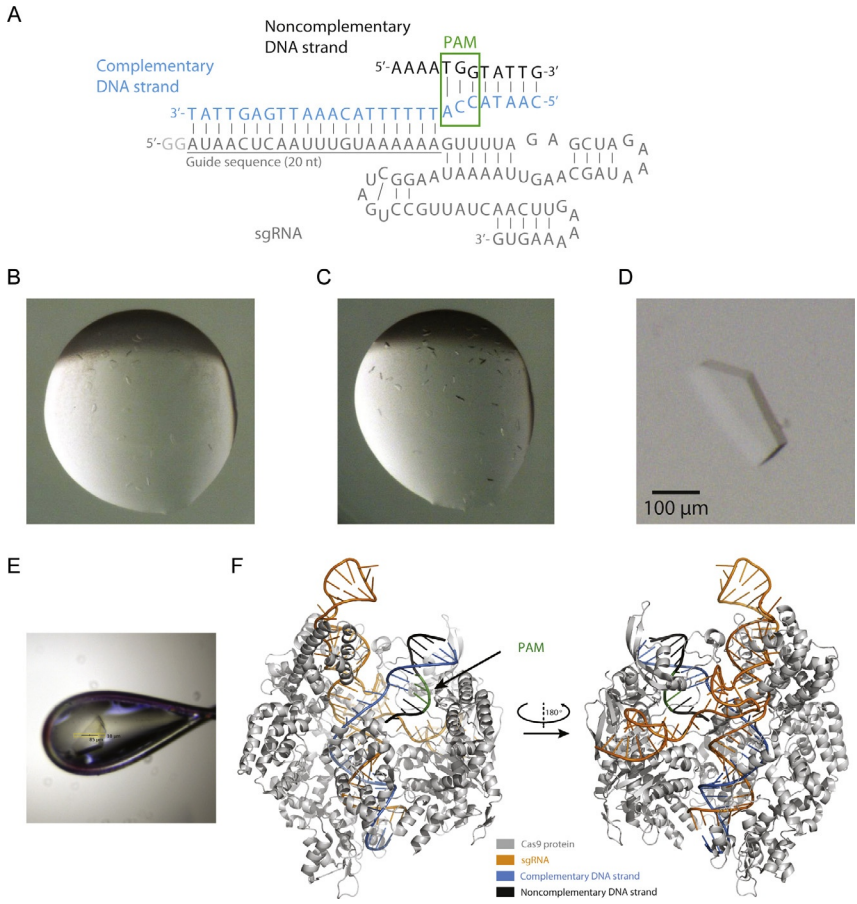


Figure 3 (A) Schematic representation of sgRNA and target DNA sequences in the dCas9–sgRNA–target DNA complex used for crystallization screening. The color coding is as in Fig. 2A. (B and C) Two examples of hits from initial crystallization screenings using 1.8 mg ml^{-1} ternary complex containing dCas9, sgRNA, and target DNA 1 (see panel A). The crystal screenings were set up at 20 °C and incubated either at 4 or 20 °C. Five out of 19 hits grew at 4 °C. Four out of 19 hits were found in KSCN, Tris–acetate, and polyethylene glycols of different chain lengths (PEG1500, PEG3350, and PEG4000). Thirteen out of 19 hits were found in Li_2SO_4, PEG4000, and different buffers spanning the pH range from pH 6.0 to 10.0. Crystals in (B) grew from 0.1 *M* Tris–acetate, pH 8.5, 0.15 *M* KSCN, 18% PEG3350. Crystals in (C) grew from 0.02 *M* CAPS, pH 10.0, 0.2 *M* Li_2SO_4, 15% PEG4000. (D) Optimized crystal after iterative rounds of seeding. The crystallization was performed at 20 °C with 1.5 mg ml^{-1} ternary complex in 20 m*M* HEPES, pH 7.5, 250 mM KCl, 5 m*M* $MgCl_2$. The crystal was grown from 0.1 *M* Tris–acetate, pH 8.5, 0.3 *M* KSCN, 16% PEG3350. (E) Ternary complex crystal harvested in a 200–300 µm nylon loop. (F) Front and rear views of dCas9–sgRNA–target DNA 1 complex shown in cartoon representation.

order of pipetting the protein and reservoir solutions (i.e., pipetting the reservoir solution first and adding protein complex solution second) led to consistently better crystals. Through these incremental improvements, we were able to obtain robust well-diffracting crystals of not just the native ternary complex but also crystals containing selenomethionine-substituted protein for phasing experiments that allowed us to solve the structure (Fig. 3F). We were subsequently able to crystallize dCas9 complexes containing mismatches between the sgRNA and the target strand. Together, the resulting structures have provided critical insights into the molecular mechanism of target DNA recognition by the Cas9–guide RNA complex and highlighted the key role of the PAM in this process (Anders et al., 2014).

5. CONCLUDING REMARKS

Within a short span of time, Cas9 has emerged as an extremely powerful molecular tool for genome editing and gene expression control that combines ease of use with a high degree of precision. Continued investigations of the molecular mechanism of Cas9 are critically dependent on structural studies of its complexes with guide RNAs and DNA substrates. These studies aim to provide detailed insights into the mechanism of target DNA recognition and cleavage and also to establish the structural framework needed for further development of Cas9-based applications. In particular, structure-based engineering of the Cas9–sgRNA platform could lead to improvements in targeting specificity, expand the spectrum of PAM sequences recognized by Cas9, and enhance the efficiency of the CRISPR-Cas9 system in transcriptional and epigenetic control (Konermann et al., 2015).

The assays outlined in this chapter have informed our work toward determining the crystal structure of Cas9 in complex with an sgRNA and a PAM-containing DNA target. Using both EMSA and FSEC approaches, we were able to analyze DNA binding by the Cas9–sgRNA binary complex and identify guide RNA and target DNA features necessary to support ternary complex formation. Using a focused set of target DNAs and crystallization screens, we crystallized the ternary complex and, through systematic improvement of the diffraction quality of the crystals, were able to solve its structure. Lessons learned from this study will influence our approach toward obtaining structures of additional functional states of Cas9 in future studies. Besides providing important tools for structural studies, our assays may prove invaluable for characterizing DNA binding by Cas9 in

application development, particularly with respect to Cas9 orthologs from other bacterial species whose DNA targeting and cleavage activities are as yet uncharacterized. Likewise, we believe that these methods will be readily adaptable to other multicomponent protein–nucleic acid complexes.

ACKNOWLEDGMENTS

We thank A. Duerst for technical assistance. This work was supported by the University of Zurich and the European Research Council (ERC) Starting Grant ANTIVIRNA (337284).

REFERENCES

Anders, C., & Jinek, M. (2014). In vitro enzymology of Cas9. *Methods in Enzymology*, *546*, 1–20.

Anders, C., Niewoehner, O., Duerst, A., & Jinek, M. (2014). Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature*, *513*, 569–573.

Anderson, J., Ptashne, M., & Harrison, S. C. (1984). Cocrystals of the DNA-binding domain of phage 434 repressor and a synthetic phage 434 operator. *Proceedings of the National Academy of Sciences of the United States of America*, *81*, 1307–1311.

Anderson, J. E., Ptashne, M., & Harrison, S. C. (1987). Structure of the repressor-operator complex of bacteriophage 434. *Nature*, *326*, 846–852.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., et al. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, *315*, 1709–1712.

Berry, K. E., Waghray, S., Mortimer, S. A., Bai, Y., & Doudna, J. A. (2011). Crystal structure of the HCV IRES central domain reveals strategy for start-codon positioning. *Structure*, *19*, 1456–1466.

Briner, A. E., Donohoue, P. D., Gomaa, A. A., Selle, K., Slorach, E. M., Nye, C. H., et al. (2014). Guide RNA functional modules direct Cas9 activity and orthogonality. *Molecular Cell*, *56*, 333–339.

Cencic, R., Miura, H., Malina, A., Robert, F., Ethier, S., Schmeing, T. M., et al. (2014). Protospacer adjacent motif (PAM)-distal sequences engage CRISPR Cas9 DNA target cleavage. *PLoS One*, *9*, e109213.

Chen, H., Choi, J., & Bailey, S. (2014). Cut site selection by the two nuclease domains of the Cas9 RNA-guided endonuclease. *Journal of Biological Chemistry*, *289*, 13284–13294.

Chen, B., Gilbert, L. A., Cimini, B. A., Schnitzbauer, J., Zhang, W., Li, G.-W., et al. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, *155*, 1479–1491.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, *339*, 819–823.

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., et al. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, *471*, 602–607.

Doudna, J. A., & Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*, *346*, 1258096.

Ferré-D'Amaré, A. R., Zhou, K., & Doudna, J. A. (1998). A general module for RNA crystallization. *Journal of Molecular Biology*, *279*, 621–631.

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., et al. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnology*, *31*, 822–826.

Garneau, J. E., Dupuis, M.-È., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., et al. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, *468*, 67–71.

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, E2579–E2586.

Gilbert, L. A., Horlbeck, M. A., Adamson, B., Villalta, J. E., Chen, Y., Whitehead, E. H., et al. (2014). Genome-scale CRISPR-mediated control of gene repression and activation. *Cell*, *159*, 647–661.

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, *154*, 442–451.

Hollis, T. (2007). Crystallization of protein-DNA complexes. *Methods in Molecular Biology*, *363*, 225–237.

Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, *157*, 1262–1278.

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology*, *31*, 827–832.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*, 816–821.

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., & Doudna, J. (2013). RNA-programmed genome editing in human cells. *eLife*, *2*, e00471.

Jinek, M., Jiang, F., Taylor, D. W., Sternberg, S. H., Kaya, E., Ma, E., et al. (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science*, *343*, 1247997.

Karvelis, T., Gasiunas, G., Miksys, A., Barrangou, R., Horvath, P., & Siksnys, V. (2013). crRNA and tracrRNA guide Cas9-mediated DNA interference in Streptococcus thermophilus. *RNA Biology*, *10*, 841–851.

Kerr, L. D. (1995). Electrophoretic mobility shift assay. *Methods in Enzymology*, *254*, 619–632.

Kibbe, W. A. (2007). OligoCalc: An online oligonucleotide properties calculator. *Nucleic Acids Research*, *35*, W43–W46.

Kimber, M. S., Vallee, F., Houston, S., Necakov, A., Skarina, T., Evdokimova, E., et al. (2003). Data mining crystallization databases: Knowledge-based approaches to optimize protein crystal screens. *Proteins*, *51*, 562–568.

Konermann, S., Brigham, M. D., Trevino, A. E., Joung, J., Abudayyeh, O. O., Barcena, C., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, *517*, 583–588.

Kuscu, C., Arslan, S., Singh, R., Thorpe, J., & Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nature Biotechnology*, *32*, 677–683.

Leung, A. K. W., Kambach, C., Kondo, Y., Kampmann, M., Jinek, M., & Nagai, K. (2010). Use of RNA tertiary interaction modules for the crystallisation of the spliceosomal snRNP core domain. *Journal of Molecular Biology*, *402*, 154–164.

Lopez-Gomollon, S., & Nicolas, F. E. (2013). Purification of DNA Oligos by denaturing polyacrylamide gel electrophoresis (PAGE). *Methods in Enzymology*, *529*, 65–83.

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., et al. (2013). Cas9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology*, *31*, 833–838.

Mali, P., Esvelt, K. M., & Church, G. M. (2013). Cas9 as a versatile tool for engineering biology. *Nature Methods*, *10*, 957–963.
Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., et al. (2013). RNA-guided human genome engineering via Cas9. *Science*, *339*, 823–826.
Mitchell, S. F., & Lorsch, J. R. (2014). Standard in vitro assays for protein–nucleic acid interactions—Gel shift assays for RNA and DNA binding. *Methods in Enzymology*, *541*, 179–196.
Nishimasu, H., Ran, F. A., Hsu, P. D., Konermann, S., Shehata, S. I., Dohmae, N., et al. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*, *156*, 935–949.
Obayashi, E., Oubridge, C., Pomeranz Krummel, D., & Nagai, K. (2007). Crystallization of RNA-protein complexes. *Methods in Molecular Biology*, *363*, 259–276.
O'Connell, M. R., Oakes, B. L., Sternberg, S. H., East-Seletsky, A., Kaplan, M., & Doudna, J. A. (2014). Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature*, *516*, 263–266.
Page, R., Deacon, A. M., Lesley, S. A., & Stevens, R. C. (2005). Shotgun crystallization strategy for structural genomics II: Crystallization conditions that produce high resolution structures for T. maritima proteins. *Journal of Structural and Functional Genomics*, *6*, 209–217.
Pattanayak, V., Lin, S., Guilinger, J. P., Ma, E., Doudna, J. A., & Liu, D. R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nature Biotechnology*, *31*, 839–843.
Reyes, F. E., Garst, A. D., & Batey, R. T. (2009). Strategies in RNA crystallography. *Methods in Enzymology*, *469*, 119–139.
Ryder, S. P., Recht, M. I., & Williamson, J. R. (2008). Quantitative analysis of protein-RNA interactions by gel mobility shift. *Methods in Molecular Biology*, *488*, 99–115.
Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C., & Doudna, J. A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, *507*, 62–67.
Thomsen, N. D., & Berger, J. M. (2012). Crystallization and X-ray structure determination of an RNA-dependent hexameric helicase. *Methods in Enzymology*, *511*, 171–190.
Winzor, D. J. (2003). Analytical exclusion chromatography. *Journal of Biochemical and Biophysical Methods*, *56*, 15–52.
Wu, X., Scott, D. A., Kriz, A. J., Chiu, A. C., Hsu, P. D., Dadon, D. B., et al. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nature Biotechnology*, *32*, 670–676.