Published online 25 November 2021

Nucleic Acids Research, 2021, Vol. 49, No. 21 ***12433–12444***
https://doi.org/10.1093/nar/gkab1139

Efficient DNA interrogation of SpCas9 governed by its electrostatic interaction with DNA beyond the PAM and protospacer

Qian Zhang[1,2,3,†], Ziting Chen[1,†], Fangzhu Wang[4], Siqi Zhang[1,2,3], Hongyu Chen[4], Xueying Gu[4], Fengcai Wen[1,2,3], Jiachuan Jin[4], Xia Zhang[1], Xingxu Huang[1,5], Bin Shen[4,*] and Bo Sun[1,5,*]

[1]School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China, [2]CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai 200031, China, [3]University of Chinese Academy of Sciences, Beijing 100049, China, [4]State Key Laboratory of Reproductive Medicine, Center for Global Health, Gusu School, Women's Hospital of Nanjing Medical University, Nanjing Maternity and Child Health Care Hospital, Nanjing Medical University, Nanjing 211166, China and [5]Gene Editing Center, ShanghaiTech University, Shanghai 201210, China

Received September 13, 2021; Revised October 26, 2021; Editorial Decision October 26, 2021; Accepted October 27, 2021

ABSTRACT

Streptococcus pyogenes **Cas9 (SpCas9), a programmable RNA-guided DNA endonuclease, has been widely repurposed for biological and medical applications. Critical interactions between SpCas9 and DNA confer the high specificity of the enzyme in genome engineering. Here, we unveil that an essential SpCas9–DNA interaction located beyond the protospacer adjacent motif (PAM) is realized through electrostatic forces between four positively charged lysines among SpCas9 residues 1151–1156 and the negatively charged DNA backbone. Modulating this interaction by substituting lysines with amino acids that have distinct charges revealed a strong dependence of DNA target binding and cleavage activities of SpCas9 on the charge. Moreover, the SpCas9 mutants show markedly distinguishable DNA interaction sites beyond the PAM compared with wild-type SpCas9. Functionally, this interaction governs DNA sampling and participates in protospacer DNA unwinding during DNA interrogation. Overall, a mechanistic and functional understanding of this vital interaction explains how SpCas9 carries out efficient DNA interrogation.**

INTRODUCTION

Clustered regularly interspaced short palindromic repeats (CRISPR)-associated (Cas) systems provide bacteria and archaea with RNA-based adaptive immunity against foreign genetic elements (1). In type II CRISPR-Cas systems, a single Cas9 endonuclease complexed with a dual-guide RNA comprising CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA) is sufficient to target and cleave complementary ∼20-base pair (bp) DNA sequences that have a short protospacer adjacent motif (PAM) located immediately downstream of the sequences (2,3). The system can be further simplified by fusing two noncoding RNA molecules into a single guide RNA (sgRNA) while retaining fully functional Cas9 activities (4). Owing to the simplicity and specificity of programmable DNA recognition and cleavage by these complexes, Cas9/sgRNA systems have proven to be extremely versatile tools for genome engineering applications (5). Additionally, catalytically dead Cas9 (dCas9, hereafter the prefix 'd' indicates the dead version of the protein) proteins that lack endonucleolytic activity but can still bind to DNA targets have been widely used for transcriptional perturbation and genomic imaging in living cells (5,6). To ensure their efficient application in medicine and biology, considerable research efforts have been devoted to increasing their fidelities and minimizing off-target effects (7,8). A thorough understanding of the molecular mechanisms of Cas9 proteins would aid in these improvements.

*To whom correspondence should be addressed. Tel: +86 21 20684536; Email: sunbo@shanghaitech.edu.cn
Correspondence may also be addressed to Bin Shen. Tel: +86 25 86869382; Email: binshen@njmu.edu.cn
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.
Present addresses:
Fangzhu Wang. Reproductive Medicine Center, Gansu Provincial Maternity and Child-Care Hospital, Lanzhou 730050, China.
Qian Zhang. School of Chemistry and Chemical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China.

The functional nuclease activities of Cas9/sgRNA complexes typically require DNA sampling, PAM recognition, protospacer DNA unwinding, R-loop formation and subdomain conformational rearrangement (9–15). Two key steps for the complexes to minimize off-target effects are PAM recognition and R-loop formation. For the most commonly used Cas9 from *Streptococcus pyogenes* (SpCas9), DNA targets containing no PAM sequence lead to the SpCas9/sgRNA complex quickly dissociating from them, whereas PAM recognition gives rise to the unwinding of a few base pairs of protospacer DNA proximal to the PAM and enables the initial DNA interrogation by guide RNA (gRNA) (9,16–18). After the initiation of DNA unwinding, R-loop formation occurs directionally from the PAM-proximal side to the PAM-distal side. In this step, the SpCas9/sgRNA complex further unwinds the remaining protospacer DNA, displaces the nontarget strand (NTS) DNA and examines the complementarity of the target strand (TS) DNA with sgRNA (9). RNA–DNA complementarity serves as a second layer of protection for the SpCas9/sgRNA complex against off-target activity. A perfect 8–12 bp match between sgRNA and the DNA target strand at the immediate 5' side of the PAM ('seed' region) is critical for DNA target binding and cleavage by SpCas9, while mismatches in the PAM-distal region can be tolerated (4,14,19). Structural studies of SpCas9/sgRNA complexed with DNA targets have provided valuable insights into the molecular basis of interactions among ternary complexes (18,20). Accordingly, several SpCas9 derivatives have been developed to improve and broaden PAM recognition and to exert higher stringency in RNA–DNA complementarity before cleavage (7,8,12). Understanding how SpCas9–DNA interactions precisely control specific steps along the pathway toward catalysis would be extremely helpful for further improvements.

Recently, our single-molecule study identified an unexpected SpCas9–DNA interaction that is located approximately 14 bp downstream of the PAM (termed post-PAM interaction) (21). This interaction site is beyond the apparent footprint of SpCas9 on DNA. Moreover, a similar DNA interaction site has also been detected for *Streptococcus aureus* Cas9 (SaCas9), highlighting an evolutionarily conserved feature of Cas9 proteins (22). Importantly, this site has proven to be critical for SpCas9 activity, as loss or occupation of the site on DNA significantly impairs the DNA binding and cleavage activity of SpCas9. Consistently, a cryoEM structure of SpCas9/sgRNA with a 55-bp DNA substrate indicated the existence of this interaction site (23). This structural analysis demonstrated that the segment 1151–1156 in the PI domain directly contacts the NTS backbone downstream of the PAM. Moreover, the interaction site was postulated to be mediated by the four positively charged lysines among these residues and the negatively charged DNA backbone. However, the functional role of the post-PAM interaction in mediating SpCas9 activity remains largely unknown.

Herein, using single-molecule, biochemical and *in vivo* DNA cleavage assays, we elucidated the mechanism and function of the post-PAM interaction in regulating SpCas9 activity. Experiments with four SpCas9 mutants revealed that this interaction indeed arises from the electrostatic forces between the four positively charged lysines among residues 1151–1156 and the DNA backbone. These mutants exhibited distinct enzymatic activities in DNA target binding and cleavage, which have a strong dependence on the residue charges. However, once successfully associated with the DNA target, they all remain tightly bound to it for hours without dissociation. Moreover, we provided evidence that this post-PAM interaction functions in DNA sampling and protospacer DNA unwinding. Based on these results, a mechanistic model is proposed to explain how the post-PAM interaction assists the SpCas9/sgRNA complex in efficient DNA interrogation.

MATERIALS AND METHODS

Expression and purification of SpCas9 proteins

All SpCas9 proteins were expressed and purified as previously described (4). Briefly, a pET-based expression vector containing a sequence encoding SpCas9 and a C-terminal His_6-tag that was used for protein expression. The codons were optimized by Synbio Tech (Suzhou, China). All SpCas9 mutants were produced by site-directed mutagenesis (Fast Mutagenesis System; Transgen Biotech, China). The SpCas9 proteins were expressed in *Escherichia coli* strain BL21 (DE3) (Transgen Biotech) grown in LB at 37°C for 4–6 h. When the optical density at 600 nm (OD_{600}) reached 0.6, protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; Sangon Biotech), and culture continued to grow at 18°C for 12–16 h. The medium was then discarded, and the cells were harvested. The harvested cells were lysed in 20 mM Tris-HCl (pH 8.0), 250 mM NaCl and 1 mM protease inhibitor cocktail, and passed through a homogenizer three times at ∼1000 bars. The lysed dilution was then ultracentrifuged at 11 000 × *g* for 30–60 min. The supernatant (i.e. clarified cell lysate) was separated from the cellular debris and bound in batch to Ni-NTA agarose resin (Transgen Biotech). The resin was washed extensively with 20 mM Tris-HCl (pH 8.0), 250 mM NaCl and 10 mM imidazole. The bound protein was eluted in a single step with 20 mM Tris-HCl (pH 8.0), 250 mM NaCl and 250 mM imidazole. SpCas9 was further purified over a HiTrap SP HP Sepharose column (GE Health Care). All peak fractions were analyzed using SurePAGE, 4–12% (GenScript). Finally, the SpCas9 sample buffer was exchanged (20 mM HEPES-KOH [pH 7.5], 500 mM KCl, 1 mM DTT) via gel-filtration chromatography on a Superdex 200 16/60 column (GE Health Care). The protein was stored at −80°C.

Preparation of DNA templates and gRNAs

The sgRNAs were prepared as follows. The pUC57-sgRNA expression vectors were linearized by BsaI (NEB) and transcribed using a T7 High-Efficiency Transcription Kit (Transgen Biotech). The sgRNAs were then purified using an EasyPure® RNA Purification Kit (Transgen Biotech). The tracrRNA was prepared similarly. Fluorophore-labeled crRNA was purchased from GenScript (Nanjing, China).

DNA oligonucleotides without or with labels were purchased from Sangon Biotech (Shanghai, China). NTS and TS DNA oligonucleotides were mixed in reaction buffer (50 mM Tris-Cl [pH 7.9], 100 mM NaCl and 10 mM $MgCl_2$),

heated to 90°C and slowly cooled to room temperature over 2 h.

The T-shaped DNA template for the single-molecule unzipping assay consisted of three DNA segments (two arms and a trunk) and was prepared as previously described (Supplementary Figure S2B) (21,24). Briefly, arm 1 was PCR-amplified from plasmid pBR322 using a digoxigenin-labeled primer. The resulting DNA fragment was digested with BstXI (NEB) to create an overhang and was subsequently annealed to a short DNA with a complementary overhang formed by adapter 1 and adapter 2. Arm 2 was PCR-amplified from plasmid pBR322 using a biotin-labeled primer. The resulting DNA fragment was digested with BstXI (NEB) to create an overhang and was subsequently annealed to adapter 3 and adapter 4. Adapter 2 from arm 1 and adapter 4 from arm 2 were partially complementary to each other and were annealed to create a short 35-bp trunk with a 3-bp overhang for trunk ligation. The 1795-bp trunk containing the target DNA sequence was amplified from the plasmid pEGFP-N1 and digested with AlwNI (NEB). The final product was produced by ligating the arms with the trunk at a 1:4 ratio. All DNA products were purified with a SanPrep Column PCR Product Purification Kit (Sangon Biotech).

The sequences of the sgRNAs, crRNA, tracrRNA, DNA binding sites, oligonucleotides, adapters and primers are listed in Supplementary Tables S1 and S2.

In vitro DNA binding and cleavage assays

Bulk DNA assays were conducted essentially as previously described (21). For the bulk DNA binding assay, dSpCas9 was first complexed with sgRNA at a 1:2 ratio at room temperature for 10 min in reaction buffer. Complexed dSpCas9 (100 or 300 nM) was incubated with annealed Cy5-DNA (1 nM) for the indicated time at room temperature. The reaction was quenched with glycerol loading buffer (10 mM Tris-HCl [pH 8.0], 10% glycerol). Reaction products were resolved by 12% PAGE and visualized by phosphorimaging (GE Health Care). The experiment for each condition was repeated at least three times.

For the DNA cleavage assay, the complexed SpCas9 (600 nM) was incubated with the DNA substrate (1 nM) for the indicated time at room temperature before quenching with formamide gel loading buffer supplemented with 50 mM EDTA. The reaction products were resolved by 12% denaturing polyacrylamide gel electrophoresis (7 M urea PAGE) and visualized by phosphorimaging (GE Health Care). Each experiment was repeated in triplicate.

Single-molecule DNA unzipping assay

DNA unzipping experiments were performed on an M-trap optical tweezer from LUMICKS (Amsterdam, Netherlands). The sample chamber preparation was described previously (24). Briefly, glass coverslips were cleaned and functionalized with partially biotinylated polyethylene glycol (PEG) (Laysan Bio) (25) and then coated with streptavidin (Thermo). Biotin-tagged DNA was added and incubated to form the DNA tethers. Antidigoxigenin-coated 0.48-μm polystyrene microspheres (Polysciences) were then added into the chamber. For the single-molecule DNA unzipping experiment, dSpCas9 or mutants (30 nM) were first complexed with sgRNA at a 1:1.5 ratio at room temperature for 10 min and then flowed into the chamber immediately before data acquisition. The experiments were conducted in a climate-controlled room at a temperature of 23.3°C; however, owing to local laser trap heating, the temperature increased slightly to 25 ± 1°C (26). The experiments were conducted by mechanically unzipping the dsDNA at a slow velocity of 50 nm/s to probe the potential interactions at the fork.

Single-molecule DNA unzipping data were acquired at 5 kHz and later filtered to 50 Hz. The acquired data signals were converted into force and DNA extension as previously described (21,27). The elasticity parameters of both dsDNA and ssDNA for data conversion were obtained from DNA force-extension curves (28). In the unzipping experiments, one separated base pair generated two nucleotides of ssDNA. Accordingly, the real-time DNA extension in nm was further converted into the number of base pairs unwound based on the elastic parameters of ssDNA under our experimental conditions. To improve the positional precision and accuracy, the force-versus-unzipped-base-pair curves were aligned to the theoretical curve by the cross-correlation of a region before and after the ternary complex disruption (27). To account for minor instrumental drift, trapping-bead size variations and DNA linker variations, the alignment allowed for a small additive shift (<5 bp) and multiplicative linear stretch (<2%) using algorithms similar to those previously described (27).

Single-molecule fluorescence resonance energy transfer (smFRET) assays

Experimental chambers used in the smFRET assays were prepared by the following protocol. The coverslips and slides used for constructing the chambers were cleaned with acetone and methyl alcohol, followed by further cleaning with sulfuric acid and 30% hydrogen peroxide at 95°C for 2 h. The coverslips were then aminosilanized and reacted with N-hydroxysuccinimide (NHS) ester-modified PEG containing 1% biotin-PEG-NHS (Laysan Bio) for specific tethering. A chamber was constructed by applying double-sided tape to stick a slide and a coverslip together. All imaging was performed using total internal reflection fluorescence (TIRF) microscopy in imaging buffer (100 mM NaCl, 50 mM Tris-HCl [pH 7.9], 10 mM $MgCl_2$, 0.8% [m/v] glucose oxidase, 2 units/μl β-D-glucose, 200 units/μl catalase and 4 mM Trolox). An EMCCD camera (Andor) was used to record videos. The smFRET assays were performed using fluorophore-labeled DNA substrates and crRNA at room temperature (Supplementary Tables S1 and S2).

For the DNA sampling assay, the Cy3-labeled DNA substrate (<50 pM) was immobilized on the glass surface via a streptavidin-biotin interaction. The preassembled dCas9/gRNA-Cy5 complex (30 nM) was injected into the chamber. The video was recorded at an exposure time of 50 ms for 2400 frames or 100 ms for 600 frames. Each frame was further processed to extract single-molecule fluorescence intensities. Only fluorescence spots in the acceptor channel were used for analysis to avoid missing or inactivat-

ing the acceptors. The FRET efficiency of a single molecule was approximated as $E = I_A/(I_D + I_A)$, where I_D and I_A are the backgrounds and leakage-corrected emission intensities of the donor and acceptor, respectively. To generate FRET histograms, the DNA samples were incubated with the dSpCas9/gRNA-Cy5 complex for 5 min, after which the first 20 frames of each DNA substrate were collected and averaged after manually filtering photobleaching effects. More than 2000 DNA substrates were sampled for each experimental condition. Multiple-peak Gaussian distributions were used to fit these histograms.

For the protospacer DNA unwinding assay, dSpCas9 or mutants (300 nM) were preassembled with sgRNA (500 nM) at room temperature for 10 min. Then, 1 nM fluorophore-labeled DNA was added to the solution for a 60-min incubation. The samples were diluted and flowed into the experimental chamber for a 10-min incubation. Free proteins and sgRNAs were washed out before data acquisition. To generate FRET histograms, the first 20 frames of each DNA substrate after washout were collected and averaged after manually filtering photobleaching effects. More than 2000 DNA substrates were sampled for each experimental condition. To obtain real-time FRET trajectories, the preassembled dCas9/sgRNA complex (30 nM) was injected into the chamber after immobilization of the DNA substrate (<50 pM) on the glass surface. The video was recorded with an exposure time of 100 ms for 600 frames after the injection.

In vivo DNA cleavage assay

SpCas9 variant plasmids were amplified from wild-type SpCas9 and Gibson assembled using the ClonExpressII One Step Cloning Kit (Vazyme). The sequences of the primers used to generate Cas9 mutants are listed in Supplementary Table S2. To generate a reporter vector, a sgRNA targeting sequence was inserted between the puromycin resistance gene and eGFP in a PiggyBac transposon vector, resulting in a frameshift of eGFP. sgRNA with the U6 promoter sequence was cloned into the transposon vector.

HEK293FT cells were cultured in DMEM supplemented with 10% FBS and 1% penicillin-streptomycin at 37°C with 5% CO_2. The reporter vector and PiggyBac transposase vector were electroplated into HEK293FT cells to establish a stable reporter cell line under puromycin selection. The SpCas9 plasmid was transfected into the stable cell line using the DS150 program on a Lonza 4-D Nucleofector with the SF Cell Line Kit according to the manufacturer's protocol. The eGFP signal was detected by flow cytometry 24, 42 and 60 h postelectroporation. To measure the *in vivo* protein expression level, transfected cells were analyzed by Western blot 60 h post-transfection. The wild-type SpCas9 plasmid was used as a positive control, and the nontransfected cells were used as a negative control.

RESULTS

The post-PAM interaction originates from electrostatic forces

Since the post-PAM interaction plays an essential role in regulating SpCas9 (hereafter referred to as Cas9) activity (21), we first aimed to determine the nature of this interaction site and address whether modulation of this site could correspondingly alter the enzymatic activity of Cas9. According to a structural study of Cas9/sgRNA in a complex with its DNA target, four positively charged lysines (K1151, K1153, K1155 and K1156) among six amino acids (1151–1156) in the PI domain of Cas9 were speculated to electronically interact with the negatively charged phosphate backbone of the DNA downstream of the PAM (Figure 1A) (23). To test this hypothesis, we developed four site-directed Cas9 mutants in which all four lysine (K) residues in the region were point-mutated to either arginine (R) with more positive charges (termed the KR mutant), histidine (H) with less positive charges (termed the KH mutant), alanine (A) without charges (termed the KA mutant) or aspartate (D) with negative charges (termed the KD mutant), as well as their catalytically dead derivatives. If the post-PAM interaction originated from electrostatic forces, the affinities of this DNA interaction site for these mutants could be expected to rank as follows: KR > Cas9 (K) > KH > KA > KD on the assumption that the substitutions lead to no conformational changes in the protein. Given that this interaction site was proven to be essential for Cas9 activity (21), the DNA binding and cleavage efficiencies of these Cas9 mutants were also expected to vary accordingly. To examine the DNA binding capabilities of the four Cas9 mutants, we used a DNA substrate containing a DNA target immediately close to a PAM and a 30-bp dsDNA downstream of the PAM to conduct the electrophoretic mobility shift assay (EMSA). Compared with the dCas9 protein, the dKR mutant showed comparable binding affinity to the DNA, whereas the other three were less efficient (Figure 1B). Notably, the DNA target binding capability of the dCas9 protein exhibited a strong dependence on the residue charges, as expected (Supplementary Figure S1A). The indistinguishable DNA binding activity between dCas9 and dKR was possibly due to the low temporal resolution of our EMSA. These findings indicate that modulation of the post-PAM interaction site can indeed alter the binding activity of Cas9 to the DNA target.

Since the four Cas9 mutants exhibited distinguishable DNA binding activities, we next asked whether the endonucleolytic activities of these mutants varied accordingly. We compared the cleavage products of the same DNA substrate by catalytically active Cas9 proteins in a denaturing gel. The endonucleolytic activity of the KR mutant was slightly enhanced compared with wild-type Cas9, whereas the other three showed reductions in cleavage efficiency, following the same trend as observed in the DNA binding assay of dCas9 mutants (Figure 1C and Supplementary Figure S1B). To further test whether the mutants exhibit similar *in vivo* cleavage activities, we developed an enhanced green fluorescent protein (eGFP) reporter system. In this system, the cleavage activity of the Cas9 proteins in HEK293FT cells can be quantified by the ratio of the eGFP-positive cells based on the fact that the expression of the eGFP gene can only be realized by the frameshift caused by the cleavage of the designed DNA recognition site (Figure 1D). At all the examined time points (24, 42 and 60 h), the cleavage efficiency of the Cas9 proteins showed the same trend between in the *in vivo* DNA cleavage assay as observed in the *in vitro* assays

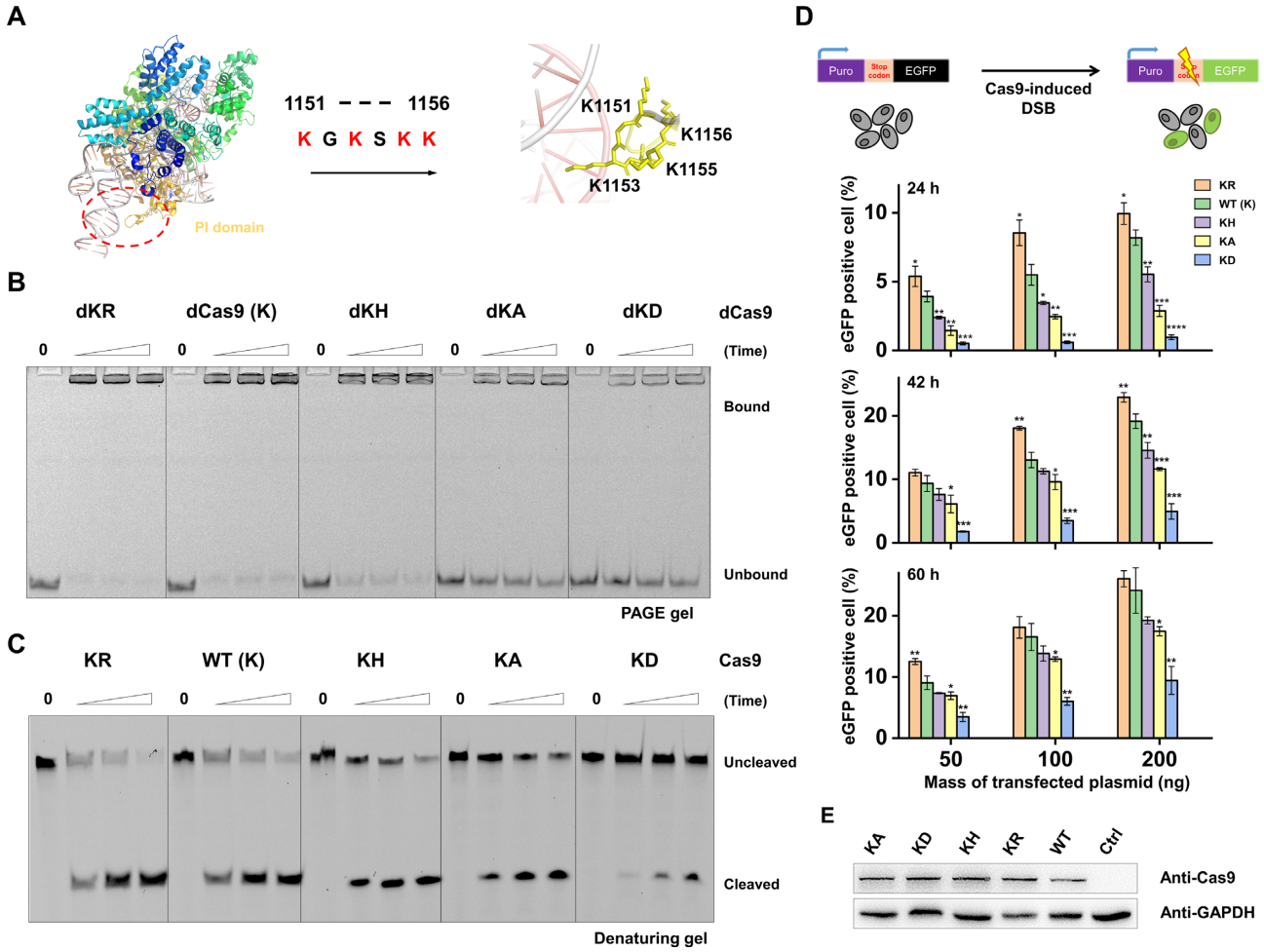

Figure 1. DNA binding and cleavage activities of SpCas9 and mutants. (**A**) Zoomed-in view of the PI domain with the amino acids indicated (PDB accession number: 5Y36). The red ellipse highlights the amino acid region that is associated with the post-PAM DNA. (**B**) A representative PAGE gel showing the DNA substrate containing a 30-bp dsDNA downstream of the PAM bound by dCas9 and four mutants. Reactions were quenched at four time points (0, 1, 10 and 30 min). (**C**) A representative denaturing gel showing the same DNA substrate cleaved by wild-type Cas9 and the four mutants. Reactions were quenched at four time points (0, 1, 5 and 10 min). (**D**) *In vivo* DNA cleavage assay with Cas9 proteins. The cartoon illustrates the experimental design. The percentages of eGFP-positive cells relative to all cells as a function of the mass of the transfected plasmid are shown for each protein at three time points. Error bars represent S.D. of three replicates; *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$; ****: $P < 0.0001$. (**E**) *In vivo* protein expression measured by western blotting.

(Figure 1D). A western blot experiment showed comparable protein expression levels among the mutants, excluding the possibility that the detected difference in this assay arises from protein expression (Figure 1E).

Taken together, our data affirm that the post-PAM interaction originates from the electrostatic forces formed between the four positively charged lysines and the negatively charged DNA phosphate backbone. Cas9 mutants with different residue charges exhibit correspondingly varied DNA binding and cleavage activities.

SpCas9 mutants differ in DNA target association and interaction

Four Cas9 mutants with altered enzymatic activities provided a unique opportunity to decipher the molecular details of the post-PAM interaction in regulating Cas9 activity. To characterize the kinetics and biophysical properties of the DNA target binding of the Cas9 mutants, we performed a previously developed optical tweezer-based DNA unzipping assay (Supplementary Figure S2A) (21,22). This single-molecule approach has proven to be a powerful tool to determine not only the efficiency of an enzyme in searching for and stable binding to its target in a long stretch of DNA but also the position and strength of the protein–DNA interaction site (27,29,30). In this assay, a T-shaped DNA template consisting of two arms and a trunk was constructed (Supplementary Figure S2B). This template was suspended between an optically trapped microsphere and a coverslip surface via two specially labeled arm ends (Supplementary Figure S2A). The 20-bp DNA target and the PAM were located in the middle of the trunk, flanked by ∼1000 bp segments of dsDNA. Moving the coverslip away from the optical trap resulted in the unzipping of the double helix of the DNA trunk and disrupted the Cas9–DNA interaction, if there was any. Catalytically dead Cas9 and mutants

were used in this assay. We performed the experiments by unzipping a single DNA molecule starting from the downstream side of the PAM wherein the post-PAM interaction can first be detected (Supplementary Figure S2A). In all examined traces with the presence of the dCas9/sgRNA complex, the unzipping signatures were either similar to that of naked DNA or showed a single force rise at the target sites (Supplementary Figure S2C and S2D). Control experiments with either sgRNA or dCas9 showed no such rise in force (Supplementary Figure S3). We thus attributed the rise in force exclusively to the disruption of a stably formed dCas9/sgRNA/DNA complex at the expected position.

We first measured the DNA target association rate of the dCas9 proteins as reflected by the fraction of DNA targets tightly bound by dCas9 as a function of time. As the rise in force detected in the DNA unzipping assay represents stable binding of dCas9 to DNA, the DNA target association defined here automatically excludes transient DNA binding of dCas9. Under our experimental conditions, the dKR mutant exhibited efficient DNA target association comparable to that of the dCas9 (K) protein, with saturating DNA target occupancy achieved in <1 min for both proteins (Figure 2A). In contrast, the other three mutants showed significantly reduced DNA target association efficiencies, and notably, it took up to 60 min for the dKA and dKD mutants to completely occupy the DNA targets (Figure 2A). Consequently, the DNA target association rates differed by at least two orders of magnitude (note that the rate constants of the dKR mutant and dCas9 could not be precisely determined due to the limited temporal resolution of our assay) (Figure 2B). We further investigated the DNA target dissociation rates of the mutants by examining the lifetime of dCas9/sgRNA in stably binding to the DNA target. In this assay, after sufficient incubation of dCas9/sgRNA and DNA (up to 2 h) to ensure complete DNA target occupancy by the complex, unbound proteins and sgRNAs were washed from the chamber, and the fraction of DNA bound by the protein was measured at defined time points after washout. More than 80% of the examined traces were found with target sites bound by dCas9 proteins even after 16 h, indicating extremely slow DNA target dissociation for all dCas9 proteins (Figure 2C). This finding is consistent with previous *in vitro* and *in vivo* studies showing an extremely long lifetime of DNA-bound Cas9 (9,31,32).

Next, we carefully analyzed the unzipping signatures of those traces showing distinguished force peaks that could reveal both the locations and the strengths of interactions of the associated complexes with the DNA target. For the sake of data presentation, we designated the protospacer DNA as + 20 (the PAM-distal side) to + 1 (the PAM-proximal side) and the PAM as 0, - 1 and - 2, followed by the downstream DNA starting from - 3 (Figure 2D). The positions of the rises in force were located - 21.7 ± 6.4 bp (mean ± S.D.) for the dKR mutant, -16 ± 4.7 bp for dCas9, -13.0 ± 5.2 bp for the dKH mutant, -7.5 ± 4.4 bp for the dKA mutant and -8.3 ± 3.1 bp for the dKD mutant downstream of the PAM (Figure 2E and Supplementary Figure S4). Unexpectedly, the magnitude of the single peak rose with an average force of ∼25 pN for all four mutants. Furthermore, with the increased positive charge in the mutation region, the detected interaction sites were shifted further away from the PAM, although the strength of the interaction site was maintained (Figure 2E).

Overall, these findings suggest that alteration of the electrostatic forces between Cas9 and DNA modulates the biophysical properties of the post-PAM interaction. Consequently, this modulation affects the DNA target association rate of Cas9 and has a negligible influence on its DNA target dissociation.

The post-PAM interaction governs DNA sampling of SpCas9

We next aimed to explicitly determine the regulatory role of the post-PAM interaction in mediating Cas9 activity. Both EMSAs and DNA unzipping assays demonstrated that the post-PAM interaction modulates the stable binding of Cas9 to the DNA target (Figures 1B and 2A). The very first step for Cas9 to achieve stable DNA target binding is DNA sampling coupled with PAM recognition. We then asked whether the altered DNA target association rates of the Cas9 mutants result from their disparate capabilities of sampling DNA for a PAM. To directly monitor stable and transient interactions of Cas9 with DNA in real time, we carried out a previously developed smFRET assay (33,34), in which biotinylated DNA substrates containing a DNA target adjacent to a PAM were anchored on the coverslip surface in a sample chamber, followed by the injection of the dCas9/gRNA complex into the chamber (Figure 3A). The Cy3 and Cy5 fluorophores were labeled on the modified nucleotide of the TS located 6 bp downstream of the PAM and the 18th nucleotide of the crRNA from the 3' end, respectively (Supplementary Tables S1 and S2). According to the structure of the ternary complex (18), the distance between the fluorophore-labeled sites is 33 Å, and a high FRET would report on a DNA target bound by a correctly oriented dCas9/gRNA complex. We first examined the binding of dCas9/gRNA to a fully matched DNA target. As expected, after the injection of dCas9 or dKR complexed with gRNA-Cy5, real-time FRET trajectories showed a sudden single-step increase in FRET efficiency from 0 to 0.86 in the first 60 s, followed by the maintenance of the high FRET state (Figure 3B). In control experiments using gRNA or dCas9 only, the 0.86 FRET state was not observed, suggesting that this state represents a stably formed dCas9–gRNA–DNA complex (Supplementary Figure S5). However, in stark contrast to dCas9 and dKR, in the first 60 s, the high FRET state was not detected at all with the other three mutants but a stable 0 FRET state instead (Figure 3B). Consistently, after a 5-min incubation of dCas9 with the DNA substrate, FRET histograms showed two populations with dKR and dCas9 centered at 0 and 0.86, and one population with dKH, dKA and dKD centered at 0 (Figure 3B). These findings are consistent with our DNA unzipping results (Figure 2A). We further examined the binding of dCas9/gRNA to a partially matched DNA target. Thirteen mismatches relative to the gRNA were introduced from the PAM-distal side to abolish the stable DNA binding of dCas9 (Supplementary Table S1) (16,35). With this substrate, transient DNA binding events were detected with dKR and dCas9 (Figure 3C). Under our experimental conditions, these two proteins attempted to associate with this partially matched DNA substrate a few

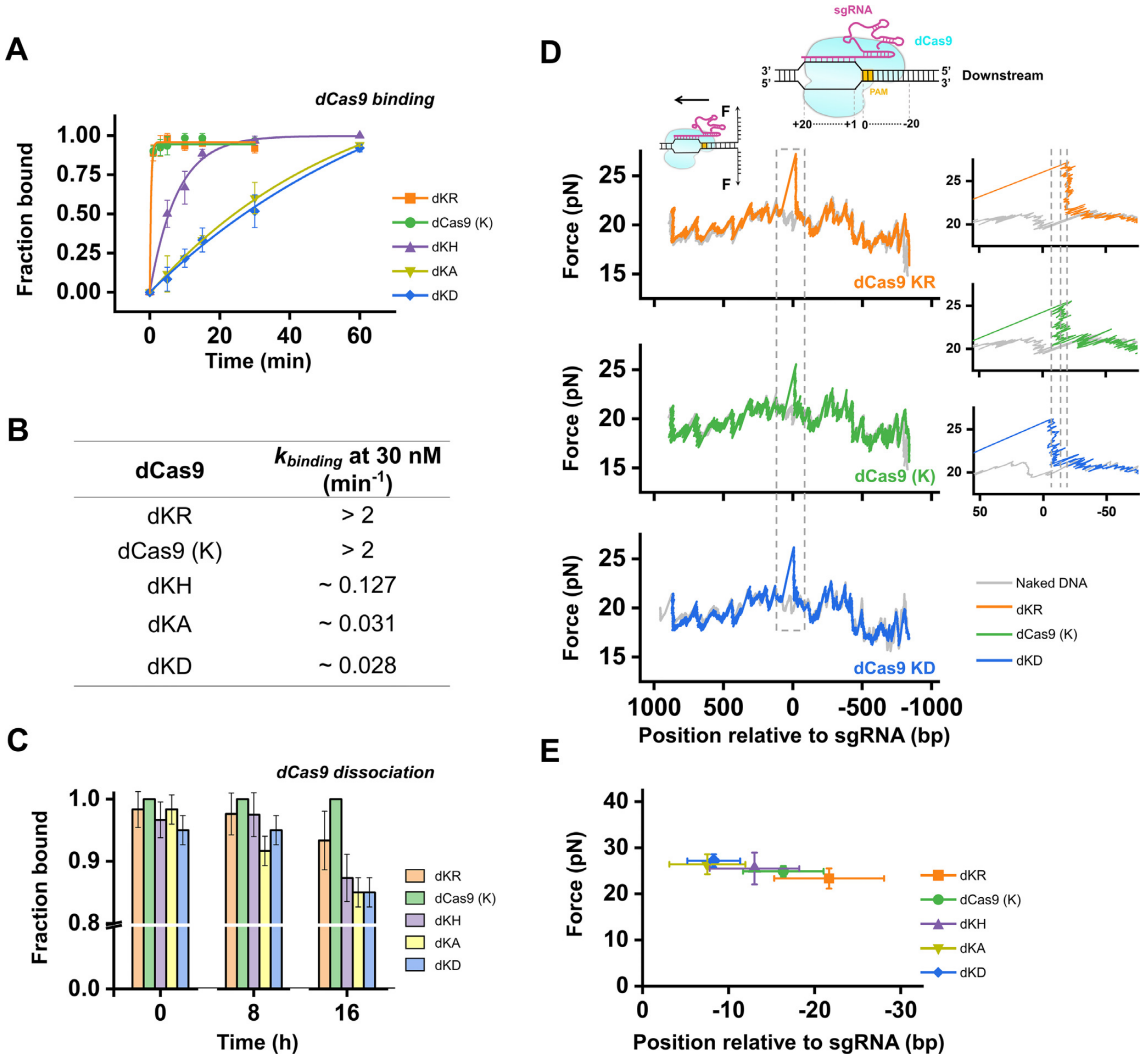

Figure 2. Measurements of DNA association and dissociation rates of Cas9 proteins and locations and strengths of the post-PAM interactions. (**A**) Fractions of DNA bound by dCas9 proteins as a function of time. Data of different dCas9 proteins are shown in different colours and were fitted to $y = 1 - e^{-x*k}$. Each data point is the average of three experiments with >30 traces per experiment. Error bars represent S.D. (**B**) Summary of DNA association rates for five dCas9 proteins. (**C**) DNA dissociation rates of dCas9 and mutants. Each data point is the average of three experiments with 15–30 traces per experiment. (**D**) Representative DNA unzipping traces from the PAM-proximal side with the presence of dKR, dCas9 or dKD showing the force versus the number of unzipped base pairs. Insets: zoomed-in views of the regions with rises in force. (**E**) Locations and strengths of the post-PAM interaction sites ($n_{dKR} = 24$, $n_{dcas9} = 52$, $n_{dKH} = 25$, $n_{dKA} = 33$ and $n_{dKD} = 48$) between dCas9/sgRNA and the DNA target. Error bars represent S.D.

times within 120 s. In contrast, experiments with the other three mutants lacked the detection of such events (Figure 3C). Consistently, statistical analysis revealed a strong dependence of DNA sampling frequency of SpCas9 on the residue charge (Supplementary Figure S6).

Collectively, these data suggest that the Cas9 mutants differ in the first DNA interrogation step and that the post-PAM interaction dictates DNA sampling of the Cas9 protein.

The post-PAM interaction participates in protospacer DNA unwinding

As sufficient protospacer DNA unwinding coupled with nearly full RNA–DNA complementarity is also a key step for Cas9 to realize stable binding to a DNA target (20), we next addressed whether the post-PAM interaction also functions in protospacer DNA unwinding. To this aim, we employed another smFRET assay to directly measure and compare the fractions of protospacer DNA unwound by Cas9 mutants (36). In this assay, biotinylated DNA substrates containing 30-bp downstream dsDNA were labeled with Cy5 and Cy3 on the target (+ 1) and nontarget (+ 9) strands in the protospacer DNA, respectively (Figure 4A and Supplementary Table S1). An initial unwinding of the 20-bp protospacer DNA would result in the separation of these two molecular dyes, thus decreasing the FRET efficiency. With this DNA substrate, a high FRET value ($E \approx 0.77$) was immediately detected, and the addition of either sgRNA or dCas9 led to no change in FRET efficiency (Supplementary Figure S7). To ensure complete DNA target occupancy, we incubated this DNA substrate with dCas9 and

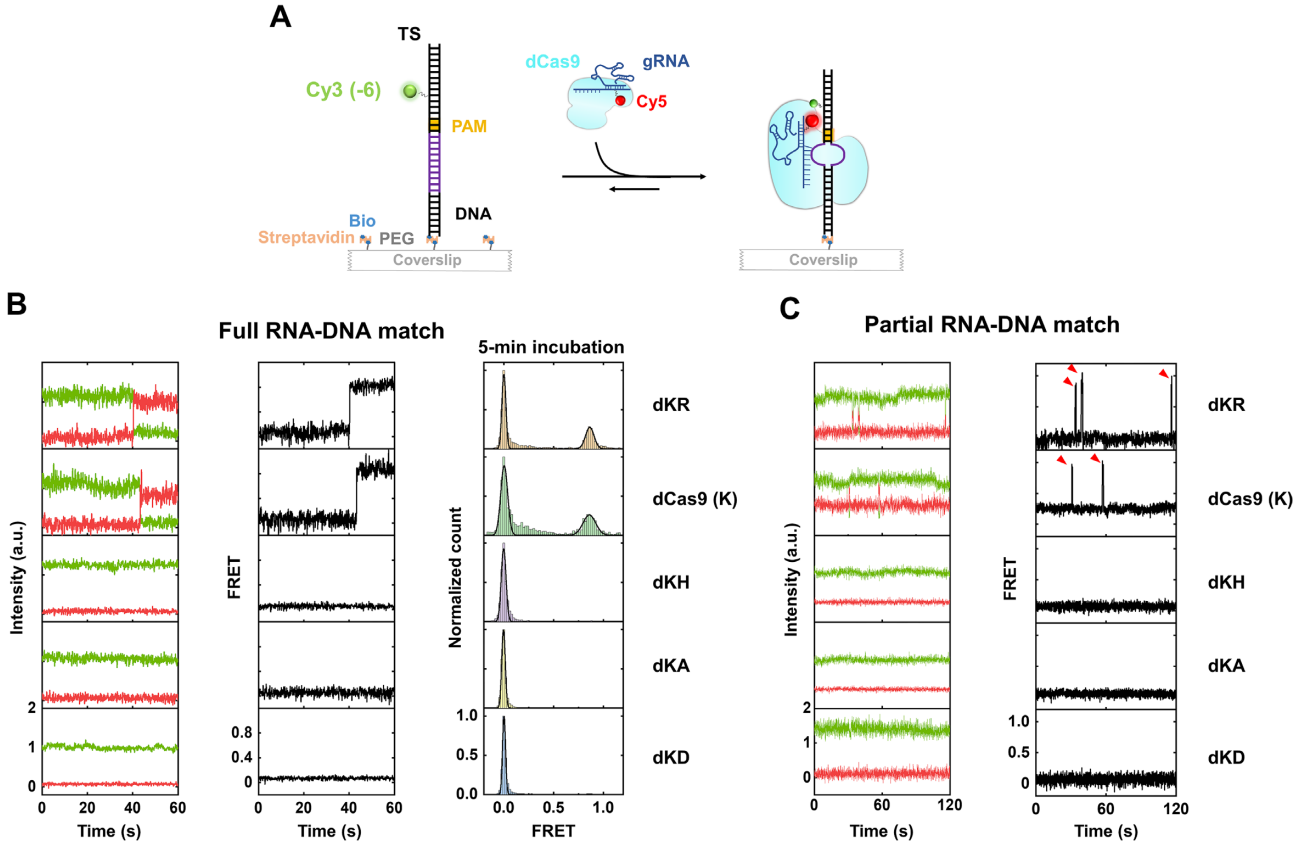

Figure 3. DNA sampling of dCas9 and mutants revealed by smFRET assays. (**A**) A schematic of the smFRET assay for the detection of DNA target sampling of dCas9 and mutants. Correctly oriented DNA target binding by the dCas9/gRNA complex gives rise to high FRET efficiency between Cy3 on the DNA substrate and Cy5 on crRNA. The cognate DNA sequence is shown in purple, and the PAM is shown in yellow. (**B**) Representative smFRET trajectories showing gRNA-Cy5-complexed dCas9 or mutants binding to a fully matched DNA target. FRET histograms of DNA substrates after incubation with dCas9 or mutants for 5 min are shown on the right. (**C**) Representative smFRET trajectories showing gRNA-Cy5-complexed dCas9 or mutants binding to a DNA target containing 13 PAM-distal mismatches. Red arrows indicate the transient DNA binding events.

a fully matched sgRNA for 60 min (Figure 4A). An EMSA confirmed that nearly all DNA substrates were bound by the examined dCas9 or mutants after incubation (Figure 4B). Compared with the naked DNA, all FRET histograms of dCas9- or dCas9 mutant-bound DNA showed a partial shift from the high FRET value ($E \approx 0.77$) to a low FRET value ($E \approx 0.53$), indicating that the protospacer DNA was unwound (Figure 4C). However, the fractions of the low-FRET population representing unwound protospacer DNA were distinguishable among the proteins, and its ranking followed the same trend as the one observed in the DNA binding and cleavage assays (Figure 4D). These data suggest that Cas9 mutants also differ in the capability of unwinding protospacer DNA and that the post-PAM interaction at least partially participates in the separation of the protospacer DNA.

We further examined the dynamics of the DNA unwinding process by introducing the dCas9/sgRNA complex to DNA substrates immobilized on the coverslip surface (Figure 4E). In line with previous findings (36), real-time trajectories revealed that most of the protospacer DNA unwinding events induced by dCas9/sgRNA occurred in a single step which was <1 s ($n = 233$ out of 266) (Figure 4F). Similar single-step DNA unwinding events were also found to dominate with the other four mutants ($n = 213$ out of 239 for dKR, 179 out of 200 for dKH, 180 out of 200 for dKA, and 188 out of 200 for dKD, respectively), albeit occurring in different time frames. The sudden decrease in FRET detected with all Cas9 mutants indicates that the difference among the Cas9 mutants possibly lies in the initiation of DNA unwinding instead of the propagation of unwinding (Figure 4F).

Preseparation of the PAM-proximal DNA compensates for the loss of the post-PAM interaction

To further establish the relationship of the post-PAM interaction with the initiation of protospacer DNA unwinding, we constructed DNA substrates containing no DNA downstream of the PAM to abolish the post-PAM interaction. Meanwhile, the DNA substrates had 1- to 4-bp mismatched 'bubble' segments located in immediate proximity to the PAM; these segments mimicked early DNA intermediates during R-loop formation. Conducting EMSA with these DNA substrates surprisingly revealed that dCas9 DNA target binding could be substantially recovered on the 1-bp bubble substrate (Figure 5A and B). Moreover, further extension of the mismatched segments drastically improved the DNA binding affinity of dCas9, with the fully restored

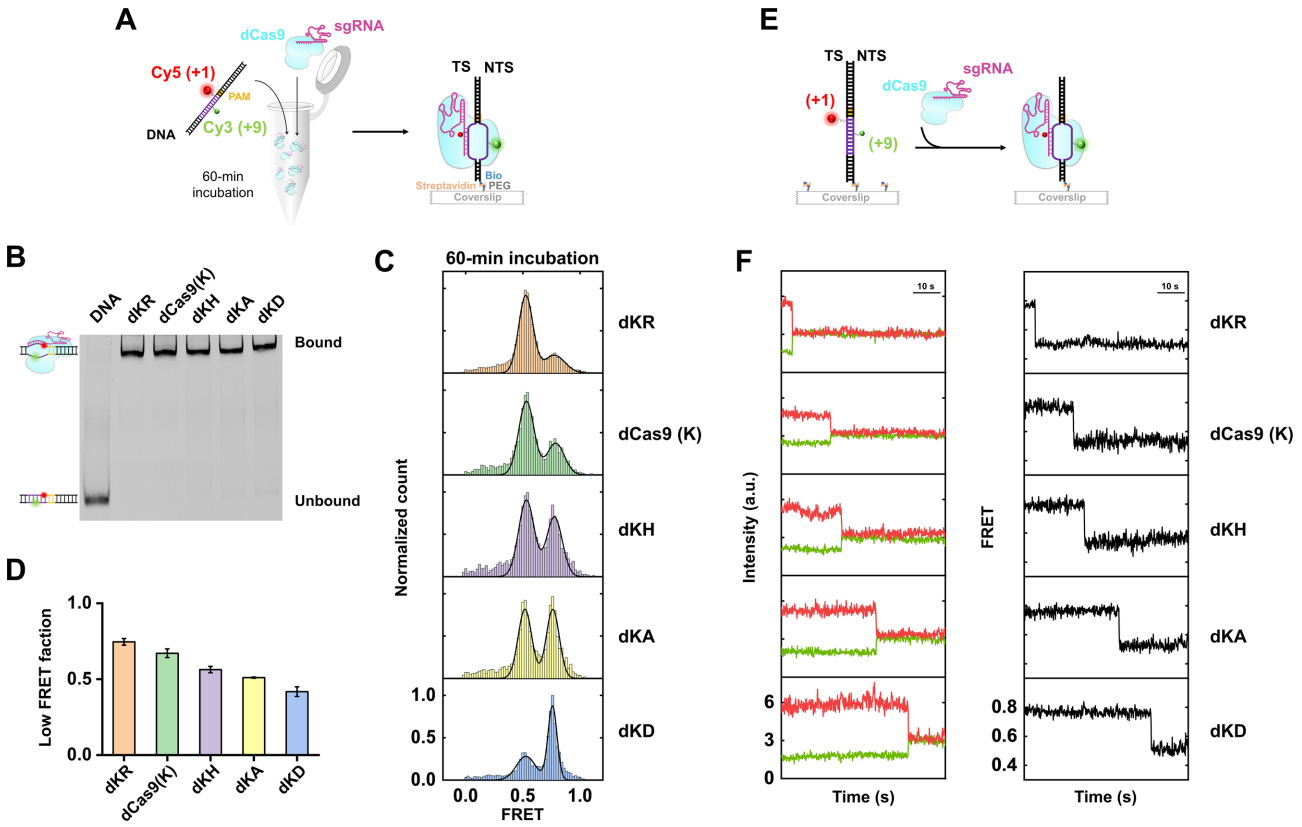

Figure 4. Protospacer DNA unwinding of dCas9 and mutants revealed by smFRET assays. (**A**) A schematic of the smFRET assay for the detection of DNA target unwinding of dCas9 and mutants. After a 60-min incubation of dCas9/sgRNA with a fluorophore-labeled DNA substrate containing a fully matched DNA target, the DNA samples were immobilized on the coverslip surface for FRET measurement. DNA unwinding leads to the separation of the labeled fluorophores on the protospacer DNA, thus decreasing the FRET efficiency. (**B**) A representative PAGE gel showing DNA substrates bound by dCas9 and mutants after a 60-min incubation. (**C**) FRET histograms of dCas9- or mutant-bound DNA substrates after a 60-min incubation. The experimental data and their Gaussian fits are shown. (**D**) Factions of unwound DNA substrates by dCas9 or mutants after a 60-min incubation. Error bars represent S.D. (**E**) A schematic of the smFRET assay for the real-time detection of protospacer DNA unwinding. The dCas9/sgRNA complex is introduced to immobilized DNA substrates. (**F**) Representative smFRET trajectories showing a sudden decrease in FRET efficiency for all examined dCas9 proteins. Note that trajectories were taken in different time frames and the representative ones do not start from the absolute zero second.

activity of dCas9 for the 3-bp and 4-bp 'bubble' substrates (Figure 5A and B).

Next, we examined the cleavage activity of Cas9 toward these DNA substrates by analyzing the cleavage products in a denaturing gel. Wild-type Cas9 showed modest cleavage activity for the 1-, 2- and 3-bp 'bubble' substrates (Figure 5C and D). Moreover, the endonucleolytic activity of Cas9 was even enhanced for the 'bubble' substrates containing four mismatched base pairs close to the PAM (Figure 5C and D). Experiments with another set of DNA substrates containing different sequences confirmed that the detected recovery of Cas9 enzymatic activity on 'bubbled' DNA substrates is unlikely to be sequence-dependent (Supplementary Figure S8). Efficient DNA binding and cleavage activities of the KA and KD mutants with these 'bubbled' DNA substrates were also observed (Supplementary Figure S9). However, Cas9 with completely mismatched sgRNA was incapable of associating with these 'bubbles' DNA substrates, indicating that R-loop initiation and DNA unwinding are both necessary for the stable association (Supplementary Figure S10). These results suggest that regarding Cas9 DNA binding and cleavage activity, preseparation of the PAM-proximal DNA can compensate for the loss of the post-PAM interaction to some extent.

DISCUSSION

This work affirms the electrostatic nature of the post-PAM Cas9–DNA interaction. By mutating the four lysines to amino acids with distinct charges, we provided evidence that the post-PAM interaction is achieved via electrostatic forces between the positively charged amino acids in the 1151–1156 region of the PI domain and the negatively charged DNA backbone (Figure 1). Intriguingly, the DNA-bound KD mutant, which has four negatively charged amino acids in the region, still showed a similar post-PAM interaction that was even closer to the PAM and an interaction strength comparable to that of the other mutants, highlighting the indispensability of the post-PAM site in Cas9 activity (Figure 2E). These shifted interaction sites could be reasoned by new binding sites formed between the mutants and the DNA target. The distinguishable interaction sites detected with the four mutants provided a clue to understanding how mutants without positive charges in the region can still

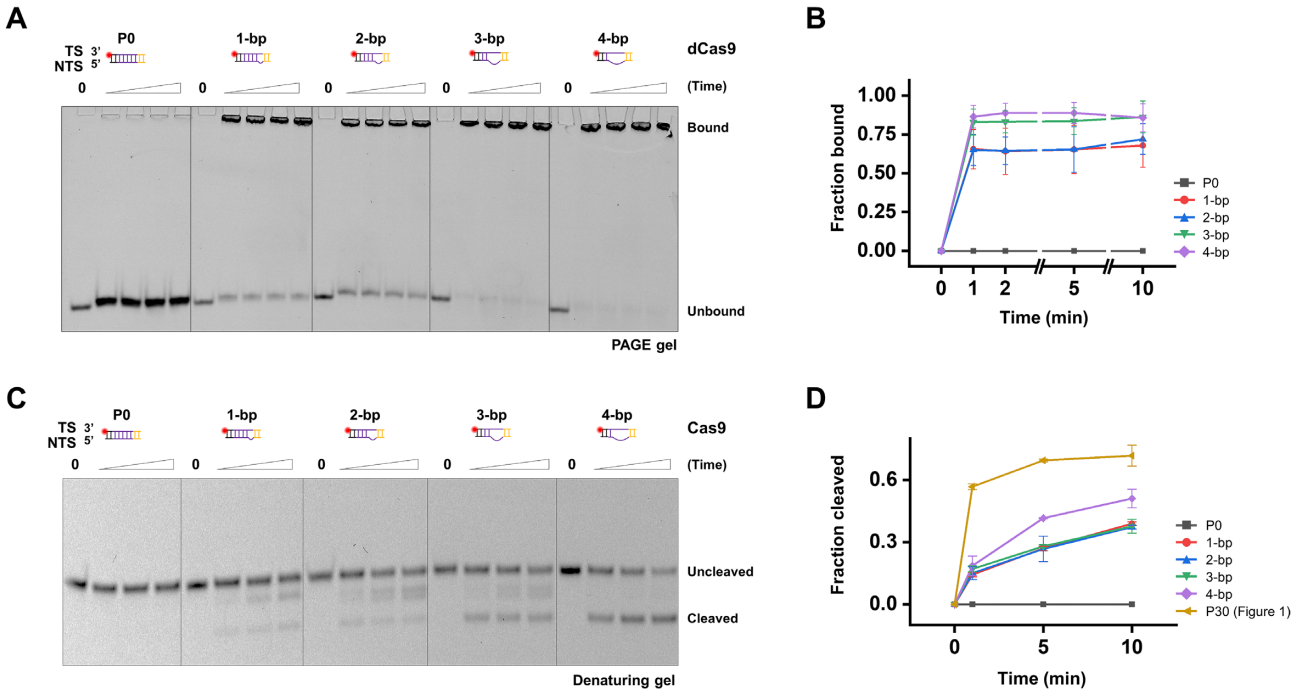

Figure 5. SpCas9 activities towards 'bubbled' DNA templates containing no dsDNA downstream of the PAM. (**A**) A representative gel showing DNA substrates containing 1 to 4 PAM-proximal mismatched base pairs bound by dCas9. These substrates contain no DNA downstream of the PAM or mismatched 'bubble' segments. P0 represents the DNA substrate containing no dsDNA downstream of the PAM. n-bp represents DNA substrates containing n-bp mismatched 'bubble' segments. Reactions were quenched at five time points (0, 1, 2, 5 and 10 min). (**B**) Fractions of DNA bound by dCas9 as a function of time. Error bars represent S.D. of three replicates. (**C**) A representative denaturing gel showing the 'bubble' substrates cleaved by wild-type Cas9. Reactions were quenched at four time points (0, 1, 5 and 10 min). (**D**) Fractions of DNA cleaved by wild-type Cas9 as a function of time. Error bars represent S.D. of three replicates.

bind to the DNA target. There might be a compensation mechanism for the loss of the original post-PAM interaction wherein positively charged residues near the 1151–1156 region in the PI domain may also associate with the DNA at an altered position to execute a similar function of the post-PAM interaction. Alternatively, there might be other interaction sites around the 14-bp region beyond the PAM instead of a single interaction point, which would have been indistinguishable in our DNA unzipping assay. In this case, the loss of one of the major interaction sites would not completely abolish its enzymatic activity, albeit resulting in a significantly reduced DNA binding efficiency (Figure 2A and B).

In addition to unveiling the biophysical properties of the post-PAM interaction, our work also clarified its functional roles in regulating Cas9 activity. Experiments with Cas9 mutants and specifically designed DNA substrates demonstrated that this interaction dictates DNA sampling of Cas9 and contributes to the initiation of protospacer DNA unwinding (Figures 3–5). Our findings explain why loss or attenuation of the post-PAM interaction strongly impairs the DNA target association rate of Cas9, as early DNA interrogation steps are hindered. Based on these findings, a mechanistic model is proposed to explain how the post-PAM interaction mediates Cas9 activity (Figure 6). During the submission of our work, a single-molecule study combined with structural data consistently reported that the lysines within the residues 1151–1156 of SpCas9 mediate its one-dimensional diffusion along DNA for PAM search (37). The astonishingly fast DNA target association rate of wild-type Cas9 is realized by quick DNA sampling and PAM recognition. The post-PAM interaction site provides Cas9 an anchoring point during DNA sampling, thus increasing its efficiency in PAM search. How does this interaction also facilitate the initiation of protospacer DNA unwinding? Increasing evidence suggests that a DNA conformational change is likely involved in the protospacer DNA unwinding and R-loop formation (17,18). The structure of the ternary SpCas9/sgRNA/DNA complex shows that the bound DNA is severely bent at the junction between the PAM and the protospacer, suggesting that local DNA distortion might account for the initiation of DNA unwinding (18). This notion is further substantiated by single-molecule and structural studies with the cascade complex from the Type I CRISPR-Cas system (38–42). Meanwhile, a fluorometric study demonstrated that the initial destabilization of the DNA duplex is achieved primarily through interactions between Cas9/sgRNA and NTS nucleotides in the protospacer close to the PAM (17). This post-PAM interaction may help immobilize the DNA-bound SpCas9/sgRNA complex and, together with the pre-PAM DNA contacts, confer efficient DNA bending and unwinding by Cas9 for DNA interrogation (18,20). The electrostatic nature of the post-PAM interaction ensures that this assistance is sequence-independent.

Tremendous efforts have been devoted to developing Cas9 derivatives to improve their efficiency and minimize off-target effects (7,8,12). Previously, the specificity en-

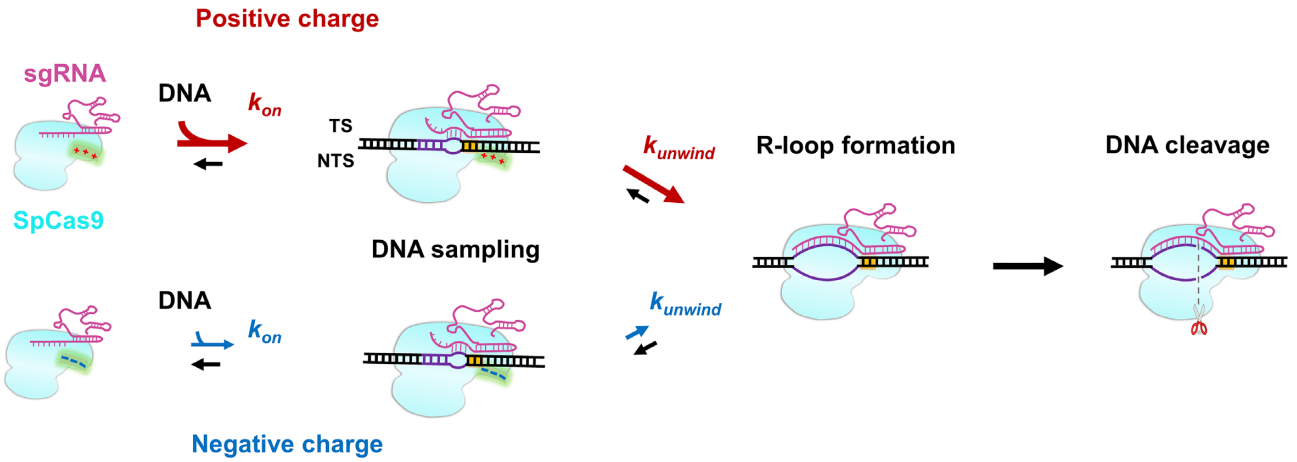

Figure 6. A proposed model for how the post-PAM interaction regulates Cas9 activity. The post-PAM SpCas9–DNA interaction is achieved via electrostatic forces between the positively charged region within 1151–1156 residues (indicated by '+') and the negatively charged DNA backbones downstream of the PAM. This interaction functions mainly in the initiation steps of DNA recognition. With this interaction, a fast PAM search and protospacer DNA unwinding of SpCas9 were realized (indicated by the long red arrows). However, the negatively charged (indicated by '-') SpCas9 mutant weakens its binding affinity to the DNA target and is inefficient in PAM search and protospacer DNA unwinding (indicated by the short blue arrows). After R-loop formation, the post-PAM intention has a negligible influence on the following steps.

hancement of Cas9 was realized mainly by mutating the amino acids interacting with the PAM or the protospacer, such as in Cas9-HF1 and eCas9 (7,8). The newly identified key residues in the PI domain, as well as its electrostatic interaction with the post-PAM DNA identified here, may provide distinct avenues to improve Cas9 efficiency. The more positively charged KR mutant showed a slightly enhanced *in vivo* cleavage capability compared with wild-type Cas9 (Figure 1D), holding promise for the further rational design of highly specific Cas9 variants as genome-modifying tools.

DATA AVAILABILITY

All data are available from the corresponding authors upon reasonable request and/or are included in the manuscript as figure source data or supplementary data.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the Molecular and Cell Biology Core Facility (MCBCF) at the School of Life Science and Technology, ShanghaiTech University for providing technical support.
Author contributions: B. Sun designed the experiments. Q.Z. and Z.C. performed the *in vitro* experiments. S.Z., F. Wen. and X.Z. helped with the single-molecule experiments. F. Wang, H.C. and B. Shen designed and performed the *in vivo* experiments with the help of X.G. and J.J.. B. Sun and B. Shen contributed to data analysis and interpretation. B. Sun, B. Shen and Q.Z. wrote the manuscript with inputs from all authors.

FUNDING

National Key Research and Development Program of China [2017YFA0106700]; National Natural Science Foundation of China [32022048, 31970796, 22104088]; Natural Science Foundation of Shanghai [19ZR1434100]; China Postdoctoral Science Foundation [2021M692053]. Funding for open access charge: National Key Research and Development Program of China [2017YFA0106700].
Conflict of interest statement. None declared.

REFERENCES

1. Marraffini,L.A. and Sontheimer,E.J. (2010) CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.*, **11**, 181–190.
2. Mojica,F.J.M., Diez-Villasenor,C., Garcia-Martinez,J. and Almendros,C. (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, **155**, 733–740.
3. 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.
4. 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.
5. Knott,G.J. and Doudna,J.A. (2018) CRISPR-Cas guides the future of genetic engineering. *Science*, **361**, 866–869.
6. Ma,H.H., Tu,L.C., Naseri,A., Huisman,M., Zhang,S.J., Grunwald,D. and Pederson,T. (2016) Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. *Nat. Biotechnol.*, **34**, 528–530.
7. Kleinstiver,B.P., Pattanayak,V., Prew,M.S., Tsai,S.Q., Nguyen,N.T., Zheng,Z. and Joung,J.K. (2016) High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*, **529**, 490–495.
8. Slaymaker,I.M., Gao,L., Zetsche,B., Scott,D.A., Yan,W.X. and Zhang,F. (2016) Rationally engineered Cas9 nucleases with improved specificity. *Science*, **351**, 84–88.
9. Sternberg,S.H., Redding,S., Jinek,M., Greene,E.C. and Doudna,J.A. (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, **507**, 62–67.
10. Sternberg,S.H., LaFrance,B., Kaplan,M. and Doudna,J.A. (2015) Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature*, **527**, 110–113.
11. Dagdas,Y.S., Chen,J.S., Sternberg,S.H., Doudna,J.A. and Yildiz,A. (2017) A conformational checkpoint between DNA binding and cleavage by CRISPR-Cas9. *Sci. Adv.*, **3**, eaao0027.

12. Chen,J.S., Dagdas,Y.S., Kleinstiver,B.P., Welch,M.M., Sousa,A.A., Harrington,L.B., Sternberg,S.H., Joung,J.K., Yildiz,A. and Doudna,J.A. (2017) Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature*, **550**, 407–410.
13. Yang,M., Peng,S., Sun,R., Lin,J., Wang,N. and Chen,C. (2018) The conformational dynamics of Cas9 governing DNA cleavage are revealed by single-molecule FRET. *Cell Rep.*, **22**, 372–382.
14. Zeng,Y., Cui,Y., Zhang,Y., Zhang,Y., Liang,M., Chen,H., Lan,J., Song,G. and Lou,J. (2018) The initiation, propagation and dynamics of CRISPR-SpyCas9 R-loop complex. *Nucleic Acids Res.*, **46**, 350–361.
15. Osuka,S., Isomura,K., Kajimoto,S., Komori,T., Nishimasu,H., Shima,T., Nureki,O. and Uemura,S. (2018) Real-time observation of flexible domain movements in CRISPR-Cas9. *EMBO J.*, **37**, e96941.
16. Singh,D., Sternberg,S.H., Fei,J.Y., Doudna,J.A. and Ha,T. (2016) Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9. *Nat. Commun.*, **7**, 12778.
17. Mekler,V., Minakhin,L. and Severinov,K. (2017) Mechanism of duplex DNA destabilization by RNA-guided Cas9 nuclease during target interrogation. *Proc. Natl. Acad. Sci. USA*, **114**, 5443–5448.
18. Anders,C., Niewoehner,O., Duerst,A. and Jinek,M. (2014) Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature*, **513**, 569–573.
19. Pattanayak,V., Lin,S., Guilinger,J.P., Ma,E., Doudna,J.A. and Liu,D.R. (2013) High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.*, **31**, 839–843.
20. Jiang,F., Taylor,D.W., Chen,J.S., Kornfeld,J.E., Zhou,K., Thompson,A.J., Nogales,E. and Doudna,J.A. (2016) Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science*, **351**, 867–871.
21. Zhang,Q., Wen,F., Zhang,S., Jin,J., Bi,L., Lu,Y., Li,M., Xi,X.G., Huang,X., Shen,B. *et al.* (2019) The post-PAM interaction of RNA-guided spCas9 with DNA dictates its target binding and dissociation. *Sci. Adv.*, **5**, eaaw9807.
22. Zhang,S., Zhang,Q., Hou,X.M., Guo,L., Wang,F., Bi,L., Zhang,X., Li,H.H., Wen,F., Xi,X.G. *et al.* (2020) Dynamics of staphylococcus aureus Cas9 in DNA target association and dissociation. *EMBO Rep.*, **21**, e50184.
23. Huai,C., Li,G., Yao,R.J., Zhang,Y., Cao,M., Kong,L.L., Jia,C.Q., Yuan,H., Chen,H.Y., Lu,D.R. *et al.* (2017) Structural insights into DNA cleavage activation of CRISPR-Cas9 system. *Nat. Commun.*, **8**, 1375.
24. Sun,B. and Wang,M.D. (2017) Single-molecule optical-trapping techniques to study molecular mechanisms of a replisome. *Methods Enzymol.*, **582**, 55–84.
25. Yardimci,H., Loveland,A.B., van Oijen,A.M. and Walter,J.C. (2012) Single-molecule analysis of DNA replication in Xenopus egg extracts. *Methods*, **57**, 179–186.
26. Peterman,E.J., Gittes,F. and Schmidt,C.F. (2003) Laser-induced heating in optical traps. *Biophys. J.*, **84**, 1308–1316.
27. Hall,M.A., Shundrovsky,A., Bai,L., Fulbright,R.M., Lis,J.T. and Wang,M.D. (2009) High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. *Nat. Struct. Mol. Biol.*, **16**, 124–129.
28. Bustamante,C., Marko,J.F., Siggia,E.D. and Smith,S. (1994) Entropic elasticity of lambda-phage DNA. *Science*, **265**, 1599–1600.
29. Meng,C.A., Fazal,F.M. and Block,S.M. (2017) Real-time observation of polymerase-promoter contact remodeling during transcription initiation. *Nat. Commun.*, **8**, 1178.
30. Koch,S.J., Shundrovsky,A., Jantzen,B.C. and Wang,M.D. (2002) Probing protein-DNA interactions by unzipping a single DNA double helix. *Biophys. J.*, **83**, 1098–1105.
31. Jones,D.L., Leroy,P., Unoson,C., Fange,D., Curic,V., Lawson,M.J. and Elf,J. (2017) Kinetics of dCas9 target search in Escherichia coli. *Science*, **357**, 1420–1424.
32. Ma,H., Tu,L.C., Naseri,A., Huisman,M., Zhang,S., Grunwald,D. and Pederson,T. (2016) CRISPR-Cas9 nuclear dynamics and target recognition in living cells. *J. Cell Biol.*, **214**, 529–537.
33. Singh,D., Wang,Y., Mallon,J., Yang,O., Fei,J., Poddar,A., Ceylan,D., Bailey,S. and Ha,T. (2018) Mechanisms of improved specificity of engineered Cas9s revealed by single-molecule FRET analysis. *Nat. Struct. Mol. Biol.*, **25**, 347–354.
34. Okafor,I.C., Singh,D., Wang,Y., Jung,M., Wang,H., Mallon,J., Bailey,S., Lee,J.K. and Ha,T. (2019) Single molecule analysis of effects of non-canonical guide RNAs and specificity-enhancing mutations on Cas9-induced DNA unwinding. *Nucleic Acids Res.*, **47**, 11880–11888.
35. Bak,S.Y., Jung,Y., Park,J., Sung,K., Jang,H.K., Bae,S. and Kim,S.K. (2021) Quantitative endonuclease Cas9 assessment of engineered Cas9 variants for target specificity enhancement by single-moleculereaction pathway analysis. *Nucleic Acids Res.*, **49**, 11312–11322.
36. Singh,D., Wang,Y., Mallon,J., Yang,O., Fei,J., Poddar,A., Ceylan,D., Bailey,S. and Ha,T. (2018) Mechanisms of improved specificity of engineered Cas9s revealed by single-molecule FRET analysis. *Nat. Struct. Mol. Biol.*, **25**, 347–354.
37. Yang,M., Sun,R., Deng,P., Yang,Y., Wang,W., Liu,J.-J.G. and Chen,C. (2021) Nonspecific interactions between SpCas9 and dsDNA sites located downstream of the PAM mediate facilitated diffusion to accelerate target search. *Chem. Sci.*, **12**, 12776–12784.
38. Szczelkun,M.D., Tikhomirova,M.S., Sinkunas,T., Gasiunas,G., Karvelis,T., Pschera,P., Siksnys,V. and Seidel,R. (2014) Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc. Natl. Acad. Sci. USA*, **111**, 9798–9803.
39. Blosser,T.R., Loeff,L., Westra,E.R., Vlot,M., Kunne,T., Sobota,M., Dekker,C., Brouns,S.J.J. and Joo,C. (2015) Two distinct DNA binding modes guide dual roles of a CRISPR-Cas protein complex. *Mol. Cell*, **58**, 60–70.
40. Rutkauskas,M., Sinkunas,T., Songailiene,I., Tikhomirova,M.S., Siksnys,V. and Seidel,R. (2015) Directional R-Loop formation by the CRISPR-Cas surveillance complex cascade provides efficient off-target site rejection. *Cell Rep.*, **10**, 1534–1543.
41. Hayes,R.P., Xiao,Y., Ding,F., van Erp,P.B., Rajashankar,K., Bailey,S., Wiedenheft,B. and Ke,A. (2016) Structural basis for promiscuous PAM recognition in type I-E Cascade from E. coli. *Nature*, **530**, 499–503.
42. Xiao,Y., Luo,M., Hayes,R.P., Kim,J., Ng,S., Ding,F., Liao,M. and Ke,A. (2017) Structure basis for directional r-loop formation and substrate handover mechanisms in type I CRISPR-Cas system. *Cell*, **170**, 48–60.