

Molecular mechanisms of *Streptococcus pyogenes* Cas9: a single-molecule perspective

Qian Zhang^{1,2,3,4}, Ziting Chen¹, 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 School of Chemistry and Chemical Engineering, Frontiers Science Center for Transformative Molecules and National Center for Translational Medicine, Shanghai Jiao Tong University, Shanghai 200240, China

⁵ Gene Editing Center, ShanghaiTech University, Shanghai 201210, China

Received: 25 June 2021 / Accepted: 11 November 2021

Abstract Cas9 is an RNA-guided endonuclease from the type II CRISPR-Cas system that employs RNA–DNA base pairing to target and cleave foreign DNA in bacteria. Due to its robust and programmable activity, Cas9 has been repurposed as a revolutionary technology for wide-ranging biological and medical applications. A comprehensive understanding of Cas9 mechanisms at the molecular level would aid in its better usage as a genome tool. Over the past few years, single-molecule techniques, such as fluorescence resonance energy transfer, DNA curtains, magnetic tweezers, and optical tweezers, have been extensively applied to characterize the detailed molecular mechanisms of Cas9 proteins. These techniques allow researchers to monitor molecular dynamics and conformational changes, probe essential DNA–protein interactions, detect intermediate states, and distinguish heterogeneity along the reaction pathway, thus providing enriched functional and mechanistic perspectives. This review outlines the single-molecule techniques that have been utilized for the investigation of Cas9 proteins and discusses insights into the mechanisms of the widely used *Streptococcus pyogenes* (Sp) Cas9 revealed through these techniques.

Keywords CRISPR, Cas9, Single-molecule, Molecular mechanism, Nuclease, R-loop

INTRODUCTION

Clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) proteins are widely found in bacterial and archaeal genomes as adaptive immune systems that offer defense against attacks from foreign genetic elements, such as viruses and plasmids (Barrangou *et al.* 2007; Brouns *et al.* 2008; Marraffini and Sontheimer 2008, 2010). The CRISPR–Cas immune response consists of three main stages. Once invasive genetic elements are detected, Cas

proteins first split the invading foreign DNA into small pieces and integrate these fragments into the CRISPR loci region. This process is referred to as adaptation, which gives rise to the formation of genetic memory of invading nucleic acids (Heler *et al.* 2015; Silas *et al.* 2016; Sternberg *et al.* 2016; Wei *et al.* 2015). Subsequently, transcription of the newly formed CRISPR loci region generates precursor CRISPR RNAs (pre-crRNAs), which are further processed into mature crRNAs (crRNA biogenesis) that can be associated with Cas proteins to form effector complexes (Carte *et al.* 2008; Charpentier *et al.* 2015; Deltcheva *et al.* 2011; Liu *et al.* 2017; Staals *et al.* 2013, 2014; Zhang *et al.* 2013b).

✉ Correspondence: sunbo@shanghaitech.edu.cn (B. Sun)

At the last interference stage, the effector complexes locate DNA targets complementary to their crRNAs and carry out the degradation of invading nucleic acids (Jinek *et al.* 2012; Jore *et al.* 2011).

Recent years have witnessed a substantial increase in the diversity of CRISPR–Cas systems. To date, CRISPR–Cas systems have been classified into two classes, six types, and 33 subtypes (Makarova *et al.* 2020). The major difference between Class I and II CRISPR–Cas systems is the composition of the effector modules. The effectors of the Class I systems are composed of multiple Cas proteins, while the Class II CRISPR–Cas systems use a single protein that functions similarly to the entire effector complex of Class I (Hayes *et al.* 2016; Jinek *et al.* 2012; Koonin *et al.* 2017). Based on the presence of the signature Cas proteins, crRNA processing, and target recognition, the two classes are further classified into various types and subtypes. We refer readers to a comprehensive review for more details on the classification of the CRISPR–Cas systems (Makarova *et al.* 2020).

The Cas9 protein is the sole nuclease in the effector complex from Type II of the Class II CRISPR system (Barrangou *et al.* 2007; Garneau *et al.* 2010; Sapranaukas *et al.* 2011). A single Cas9 endonuclease complexed with a dual guide RNA (gRNA) comprising 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 (Deltcheva *et al.* 2011; Garneau *et al.* 2010; Gasiunas *et al.* 2012; Jinek *et al.* 2012; Karvelis *et al.* 2013). The system can be further simplified by fusing the two RNA molecules into a single guide RNA (sgRNA) (Jinek *et al.* 2012). Due to its simplicity and programmability, the Cas9 protein has been widely repurposed as an effective RNA-guided DNA-targeting platform that can easily modify the genome in various species (Knott and Doudna 2018; Zhang 2019). In addition, the nuclease-deficient version of the enzyme (dCas9) has also been widely used in transcriptional regulation and *in vivo* imaging (Bikard *et al.* 2013; Chen *et al.* 2013; Gilbert *et al.* 2013, 2014; Hilton *et al.* 2015; Konermann *et al.* 2013, 2015; Ma *et al.* 2015, 2016b; Maeder *et al.* 2013; Perez-Pinera *et al.* 2013; Qi *et al.* 2013; Tanenbaum *et al.* 2014; Thakore *et al.* 2015). To ensure their efficiency and accuracy, a thorough understanding of the molecular mechanisms of Cas9 proteins would aid in increasing their fidelity and minimizing off-target effects.

The functional nuclease activity of Cas9–sgRNA complexes commonly requires PAM search and recognition, protospacer DNA unwinding, R-loop

formation, and subdomain conformational rearrangement (Anders *et al.* 2014; Jiang *et al.* 2015, 2016b; Sternberg *et al.* 2014, 2015). Although ensemble studies have contributed enormously to the understanding of the molecular mechanisms of Cas9 proteins, these ensemble approaches often reveal the average molecular population and lack the ability to detect intermediate states and distinguish heterogeneity of Cas9 proteins. In the past decade, single-molecule techniques have become the complementary approaches that help understand the detailed molecular dynamics of Cas9 proteins (Cuculis and Schroeder 2017; Globyte *et al.* 2018; Singh and Ha 2018; Whinn *et al.* 2019). These techniques offer the ability to provide enriched information on each specific step of Cas9 proteins along its pathway towards catalysis. Moreover, single-molecule manipulation techniques, such as optical tweezers and magnetic tweezers, offer the capability to apply external mechanical force and torque on DNA substrates (Mullally *et al.* 2020; Newton *et al.* 2019; Szczelkun *et al.* 2014; Zhang *et al.* 2019). Using these techniques, a more comprehensive understanding of the Cas9 mechanism can be generated. In this review, we introduce the major single-molecule approaches used to study Cas9 proteins and highlight examples of insights into the molecular mechanisms of the best-characterized *Streptococcus pyogenes* (Sp) Cas9 obtained using these methods.

SINGLE-MOLECULE TECHNIQUES FOR Cas9 STUDIES

Compared with ensemble studies, single-molecule approaches have shown great advantages in measurements of molecular heterogeneity, distributions in molecular behaviors, and real-time dynamics of single biomolecules (Cordes *et al.* 2015; Cuculis and Schroeder 2017; Leake 2013; Sun 2019; Sun and Wang 2016). A wide variety of fluorescence spectroscopy-based (DNA curtains and fluorescence resonance energy transfer) and force spectroscopy-based (magnetic tweezers, optical tweezers, and atomic force microscopy) single-molecule techniques have been used to investigate different aspects of Cas9 proteins (Cuculis and Schroeder 2017; Globyte *et al.* 2018; Singh and Ha 2018; Whinn *et al.* 2019). In this section, we will briefly describe their principles and capabilities.

DNA curtains

DNA curtains refer to a methodology that aligns hundreds of DNA molecules on the surface of a

microfluidic sample chamber wherein fluorescently labeled DNA and proteins can be monitored simultaneously under total internal reflection fluorescence (TIRF) microscopy (Axelrod 1989; Fazio *et al.* 2008; Greene *et al.* 2010). This approach commonly uses biotin and streptavidin in the phospholipid bilayer on a quartz wafer to anchor multiple DNA molecules (Fig. 1A, left). Due to the barriers in the phospholipid bilayer, one end of DNA molecules is fixed and aligned, and the others, driven by flow, are free-floating (Fazio *et al.* 2008). Alternatively, to avoid the waste of samples and the perturbation from the flow, both ends of a relatively stretched DNA molecule can be anchored (Fig. 1A, right) (Gorman *et al.* 2010). In this scenario, the flow is not needed. DNA curtains are a high-throughput technique that allows real-time visualization of hundreds of fluorescently labeled DNA and proteins. This technique has been successfully applied to investigate PAM search, DNA association and dissociation of SpCas9 (Table 1) (Cuculis *et al.* 2016; Redding *et al.* 2015; Sternberg *et al.* 2014).

Fluorescence resonance energy transfer

Single-molecule fluorescence resonance energy transfer (smFRET) is a powerful technique that can monitor nanometer-scale change in biological macromolecules, such as DNA and protein. By measuring the resonance energy transfer efficiency (E) between a donor and an acceptor dye, this approach reveals the distance change between the labeling sites in real time (Roy *et al.* 2008). smFRET measurements are generally carried out under TIRF microscopy with a light source (usually a laser) irradiating the fluorescently labeled donor molecules on a glass surface. As shown in Fig. 1B, the FRET signals reflect the sensitive change in the distance between the dyes, ranging from 2–8 nm (Roy *et al.* 2008; Selvin 2000). The high spatial and temporal resolution of smFRET allows to detect minute changes and capture transient intermediates between two molecules or within a single molecule. Therefore, smFRET has been involved in the investigation of nearly all catalytic steps of

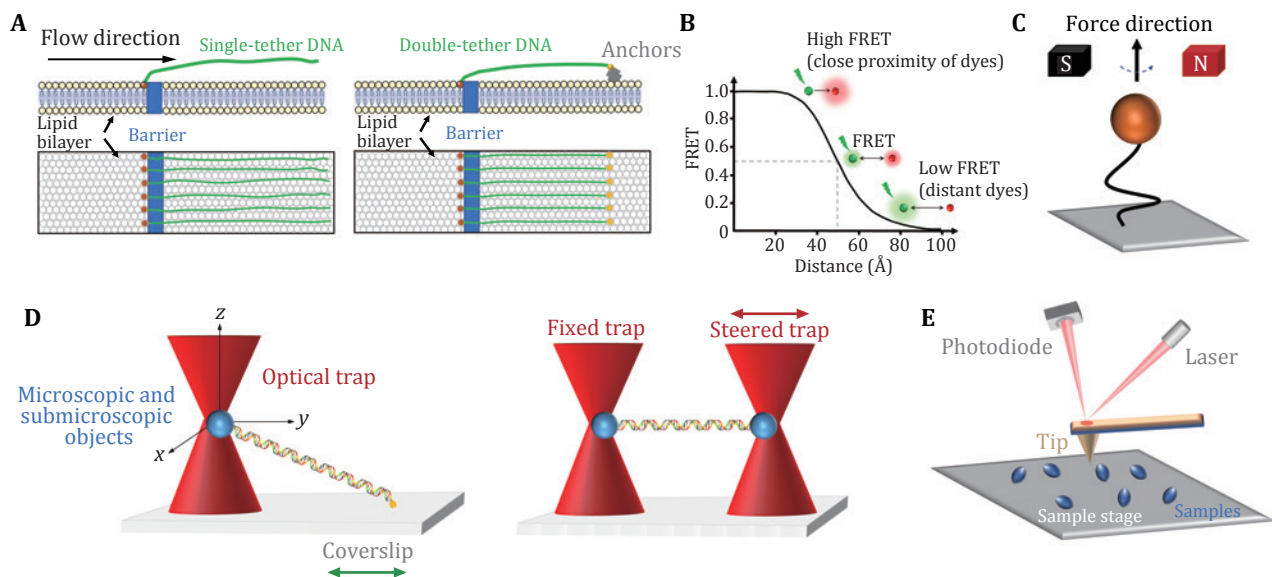


Fig. 1 Single-molecule techniques for the study of Cas9 proteins. **A** DNA curtains. Single-tethered DNA curtains (left). An array of DNA molecules is aligned by a barrier in a lipid bilayer, while laminar flow stretches the molecules away from the barrier. Double-tethered DNA curtains (right). Two ends of a DNA molecule are respectively fixed to the barrier and the anchors (yellow) on the phospholipid bilayer. **B** smFRET. Schematic of the FRET efficiency as a function of the distance (R) between a pair of dyes for $R_0 = 50$ Å. The donor dye transfers energy to the acceptor dye. The transfer efficiency depends on the distance between the two dyes. **C** Magnetic tweezers. One end of a double-stranded DNA molecule is usually attached to the glass surface, and the other end is attached to a magnetic bead. The magnetic field can apply force and torque to the magnetic bead, thereby manipulating the DNA molecules. **D** Optical tweezers. A highly focused laser beam can capture and move microscopic and submicroscopic objects, such as polystyrene beads, thereby manipulating and monitoring the DNA molecule attached to them. Single optical tweezers (left) typically require the ends of a DNA molecule attached to a bead and the coverslip surface. In a dual optical tweezers (right) configuration, two ends of a DNA molecule are attached to two beads manipulated by two traps. **E** AFM. The tip of the needle fluctuates in the direction perpendicular to the surface of the sample under the action of a constant repulsive force so that information on the surface morphology of the sample can be obtained

Table 1 Single-molecule approaches applied to investigate each catalytic step of Cas9

Cas9 activity	Single-molecule methods
PAM search	DNA curtains, FRET
DNA target binding	DNA curtains, FRET, Optical tweezers, AFM
Protospacer unwinding and R-loop formation	Magnetic tweezers, FRET
Conformational rearrangement	FRET, AFM
Dissociation from DNA	DNA curtains, FRET, Optical tweezers

SpCas9 (Table 1) (Bak *et al.* 2021; Chen *et al.* 2017; Dagdas *et al.* 2017; Globyte *et al.* 2019; Lim *et al.* 2016; Okafor *et al.* 2019; Osuka *et al.* 2018; Singh *et al.* 2016, 2018; Sternberg *et al.* 2015; Sung *et al.* 2018; Wang *et al.* 2021; Yang *et al.* 2018, 2021; Zeng *et al.* 2018; Zhang *et al.* 2021).

Magnetic tweezers

Magnetic tweezers are a single-molecule manipulation technology that can apply force and torque to a group of single DNA molecules (Charvin *et al.* 2005; Gupta *et al.* 2009; Jiang *et al.* 2016a; Sarkar and Rybenkov 2016; Strick *et al.* 1998). In this approach, the two ends of a linear DNA molecule are usually attached to a glass surface and a magnetic bead, respectively (Fig. 1C). A gradient magnetic field exerts force and torque on the magnetic beads, thus manipulating the DNA molecule (Jiang *et al.* 2016a; Strick *et al.* 1996, 1998). By adjusting the external magnetic field, the magnetic beads can be pulled or rotated so that the attached DNA molecules can be stretched or twisted (Cheezum *et al.* 2001; Manosas *et al.* 2010; Sbalzarini and Koumoutsakos 2005). Magnetic tweezers have the advantages of naturally maintaining a constant force within 0.1–100 pN on the DNA and introducing DNA supercoiling under low forces. In the torsionally constrained configuration, magnetic tweezers allow for the sensitive detection of minute changes in DNA length caused by the separation of a few base pairs of dsDNA. Therefore, the R-loop formation induced by SpCas9 proteins can be detected (Table 1) (Mullally *et al.* 2020; Szczelkun *et al.* 2014).

Optical tweezers

Optical tweezers apply highly focused laser beams to capture and move microscopic and submicroscopic objects (Fig. 1D) (Ashkin *et al.* 1986; Bustamante *et al.* 2021). The single-beam particle trap is generated by an optical radiation pressure gradient force that can flexibly capture samples from several nm to tens of nm, such as biological macromolecular particles (spheres)

and organelles (Zhang *et al.* 2013a). The ability of optical tweezers to gently manipulate microscale objects suits the study of “fragile” biological macromolecules. Optical traps can manipulate dielectric microspheres that are attached to biomolecules and detect their positions and forces in real time. This technique offers flexible control of the force and extension of the substrate, thus enabling a quick switch between different modes of operation (Finer *et al.* 1994; Smith *et al.* 1996; Wang *et al.* 1998; Zhang *et al.* 2019, 2020). Optical tweezers provide capabilities to analyze the dynamics of molecules on the spatial scale of a nanometer (nm), the time scale of a millisecond (ms), and the force scale of piconewton (pN, 10^{-12} N) (Maragò *et al.* 2010; Polimeno *et al.* 2018; Zhang *et al.* 2013a).

In single-trap optical tweezers, two ends of DNA molecules are typically attached to the microsphere and the surface of the coverslip, respectively (Fig. 1D, left). Dual optical tweezers suspend a DNA molecule via two traps. This experimental configuration isolates the measurements from the sample chamber, thereby reducing instrument noise (Fig. 1D, right). Critical SpCas9–DNA interactions and the force impact on SpCas9 binding were reported via the use of optical tweezers (Table 1) (Newton *et al.* 2019; Zhang *et al.* 2019, 2020, 2021).

Atomic force microscopy

Atomic force microscopy (AFM) studies the surface structure and properties of samples by detecting the extremely weak interatomic interaction between the sample surface and a miniature force-sensitive sensor (Binnig *et al.* 1986). In this approach, one end of a microcantilever that is extremely sensitive to weak forces is fixed, and the other end has a tiny needle tip (Fig. 1E). The needle tip contacts the sample surface lightly. During scanning, the microcantilever with the tip will undulate in the direction perpendicular to the sample surface (Binnig *et al.* 1986). Using the optical detection method, the position changes of the microcantilever corresponding to each scanning point

can be measured, which reflects the sample surface morphology information (Binnig *et al.* 1986). High-speed (HS) AFM allows for the data acquired at high temporal resolution and has successfully revealed real-time conformational changes of SpCas9 (Table 1) (Shibata *et al.* 2017).

MOLECULAR MECHANISMS OF SpCas9 REVEALED BY SINGLE-MOLECULE TECHNIQUES

SpCas9 is a large nuclease composed of 1,368 amino acids (Anders *et al.* 2014; Jiang *et al.* 2016b; Jinek *et al.* 2014; Nishimasu *et al.* 2014). SpCas9 undergoes a series of discrete DNA interrogation steps before cleavage. These steps are governed by critical DNA interactions and are often coupled with protein conformational changes (Anders *et al.* 2014; Jiang and Doudna 2017; Jiang *et al.* 2015, 2016b; Jinek *et al.* 2014; Nishimasu *et al.* 2014). Single-molecule techniques have offered unique experimental approaches to investigate different steps and aspects of the SpCas9 protein in the process of interrogating and cleaving DNA targets (Table 1). Below, we will highlight a few

examples where single-molecule approaches have brought important insights into its molecular mechanisms.

PAM search

Upon complexation with gRNA, the very first step in SpCas9 activity is to locate its DNA targets. Considering the vast amount of DNA in the cell, locating a 20–30-base pair protospacer could be time-consuming. How does the SpCas9–gRNA complex accelerate this process? Addressing this question necessitates the ability to simultaneously monitor both DNA and SpCas9–gRNA at the high spatiotemporal resolution, as DNA target recognition is an intrinsically transient protein–DNA interaction. DNA curtains and smFRET approaches are suitable for this task and have been applied to serve this purpose.

Sternberg *et al.* used a DNA curtain assay to visualize the target searching of a single SpCas9–gRNA complex in real time (Sternberg *et al.* 2014). The search process of the complex was monitored by imaging the YOYO1-stained λ -DNA molecule and quantum dot (QD)-labeled SpCas9 protein in complex with gRNA (Fig. 2A). They

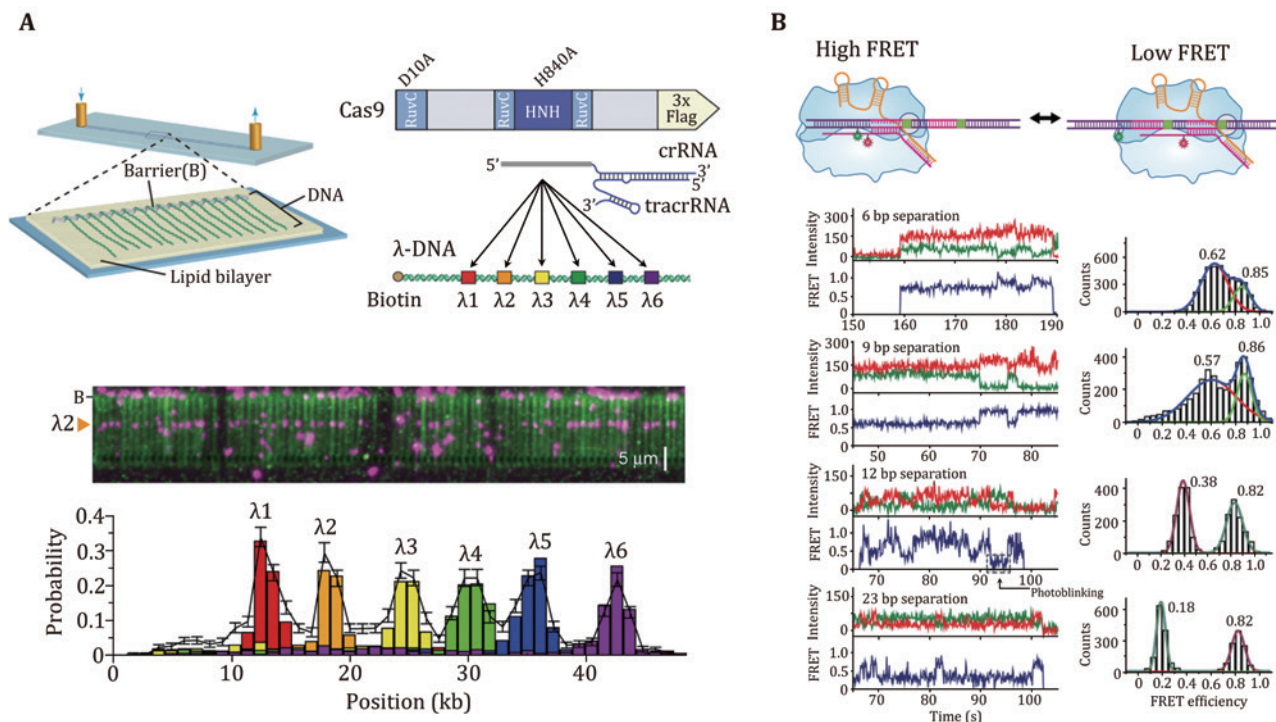


Fig. 2 PAM search of SpCas9 revealed by single-molecule studies. **A** Schematic of the single-tethered DNA curtain for the SpCas9 PAM search assay. SpCas9–gRNA is designed to bind to six DNA target sites. SpCas9 binding sites are detected by DNA stained with YOYO1 (green) and SpCas9 labeled with QDs (magenta). **B** Schematic, smFRET traces, and histograms of SpCas9 binding to the PAMs at different locations, and the distance between PAMs is adjusted. The histograms show two FRET peaks corresponding to either of the target DNA sites. The high FRET peak remains constant across each histogram, while the low FRET peak moved towards the low FRET value as the distance between the targets increases. Adapted from Sternberg *et al.* (2014) and Globyte *et al.* (2019) with permissions

found that SpCas9–gRNA first searches for a dinucleotide PAM (5′-GG-3′) in a three-dimensional (3D) collision manner. The complex only transiently samples off-target sites, which is correlated with the PAM density and ignores sequences that are complementary to the gRNA but lack an adjacent PAM. These data suggest that SpCas9–gRNA begins the PAM search through a random collision with DNA. To accelerate the search, the complex reduces the time spent at non-PAM sites and interrogates only the PAM-flanking DNA for gRNA complementarity. In this assay, they observed no evidence of SpCas9–gRNA associating with targets by one-dimensional (1D) sliding/hopping. However, due to the diffraction limit of light microscopy (~250 nm), the 3D target search model may not be valid for a short length scale of nucleotides. To test that, Globyte *et al.* employed a smFRET experiment to examine the target searching process of a single SpCas9–gRNA complex in a smaller range (Fig. 2B) (Globyte *et al.* 2019). They immobilized a biotinylated SpCas9 complexed with Cy3-labelled gRNA on a PEG-coated quartz surface, followed by the injection of Cy5-labelled DNA into the reaction chamber. Using a series of DNA templates with a few PAMs placed at different distances, the experiment showed that weak PAM binding enables the complex to interrogate adjacent sequences with a range of ~20 bp in a facilitated 1D diffusion manner. These results suggest that the SpCas9–gRNA complex employs a combination of both 3D collision and 1D diffusion to locate its targets.

Stable binding to a DNA target

After PAM recognition, whether SpCas9–gRNA stably binds to or quickly dissociates from a DNA target relies on crRNA–DNA complementarity. To address how mismatches influence target recognition and DNA binding of SpCas9, Singh *et al.* designed smFRET assays to monitor real-time interactions between SpCas9–gRNA and DNA targets (Fig. 3A) (Singh *et al.* 2016). By labeling gRNA with Cy5 and DNA target with Cy3, they found that mismatches proximal to PAM greatly increase the SpCas9 dissociation rate (from $<0.006\text{ s}^{-1}$ to $>2\text{ s}^{-1}$), whereas PAM-distal mismatches still allow for the stable binding of the complex to DNA targets. Specifically, 9–10 PAM-proximal matches are sufficient for ultrastable SpCas9–gRNA binding. Moreover, as the dwell-time analysis shows two characteristic binding times, a two-step mechanism of Cas9–RNA binding involving PAM surveillance and RNA–DNA heteroduplex formation (see the next section) was proposed (Singh *et al.* 2016). In addition

to PAM-distal mismatches, the fluorescence-combined optical tweezers and smFRET assays from the Rueda Laboratory revealed that DNA bubbles, driven by the mechanical forces on the DNA, could also boost stable binding of SpCas9 to off-targets, including mismatches in the PAM-proximal region (Fig. 3B) (Newton *et al.* 2019). Moreover, using magnetic tweezers, a similar enhanced off-target binding effect was also detected with supercoiled DNA substrates (Ivanov *et al.* 2020). These data suggest that bubbles and supercoiling in DNA substrates further increase the promiscuity of stable binding of SpCas9 to off-targets.

The stable binding of SpCas9–gRNA to DNA targets is achieved via direct interactions among the ternary complex. To quantitatively determine these essential interaction sites, Zhang *et al.* used an optical-tweezer-based DNA unzipping technique to probe them along the DNA sequence (Fig. 3C) (Zhang *et al.* 2019). In addition to a strong interaction located within the protospacer, they uncovered an unexpected interaction site located approximately 14 bp downstream of the PAM (post-PAM interaction), which is beyond the PAM and protospacer. Importantly, the loss or occupation of this relatively weak interaction site on DNA significantly attenuates SpCas9 binding. This site was recently verified to mediate DNA sampling and unwinding of SpCas9 (Yang *et al.* 2021; Zhang *et al.* 2021). Consistent with the smFRET data, they also found that a 9-bp PAM-proximal RNA–DNA complementarity was sufficient to support stable SpCas9 binding (Zhang *et al.* 2019). Interestingly, a transient interaction at approximately 15 bp relative to the PAM in the protospacer region among the ternary complex was frequently detected with imperfect RNA–DNA complementarity. This transient interaction was postulated to serve to mediate sensing of RNA–DNA complementarity (see the next section) and/or govern HNH domain mobility for cleavage (see the following section of DNA dissociation after cleavage).

Protospacer DNA unwinding and R-loop formation

To examine crRNA–DNA complementarity, protospacer DNA must be unwound by SpCas9–gRNA in an ATP-independent manner. crRNA–DNA complementarity serves as a second layer of protection for SpCas9–gRNA against off-target binding, and DNA unwinding coupled with R-loop formation has proven to be a primary determinant of SpCas9 activity (Gong *et al.* 2018). In this process, the complex displaces the nontarget strand (NTS) and hybridizes the target strand (TS) of the protospacer with gRNA, allowing for the formation

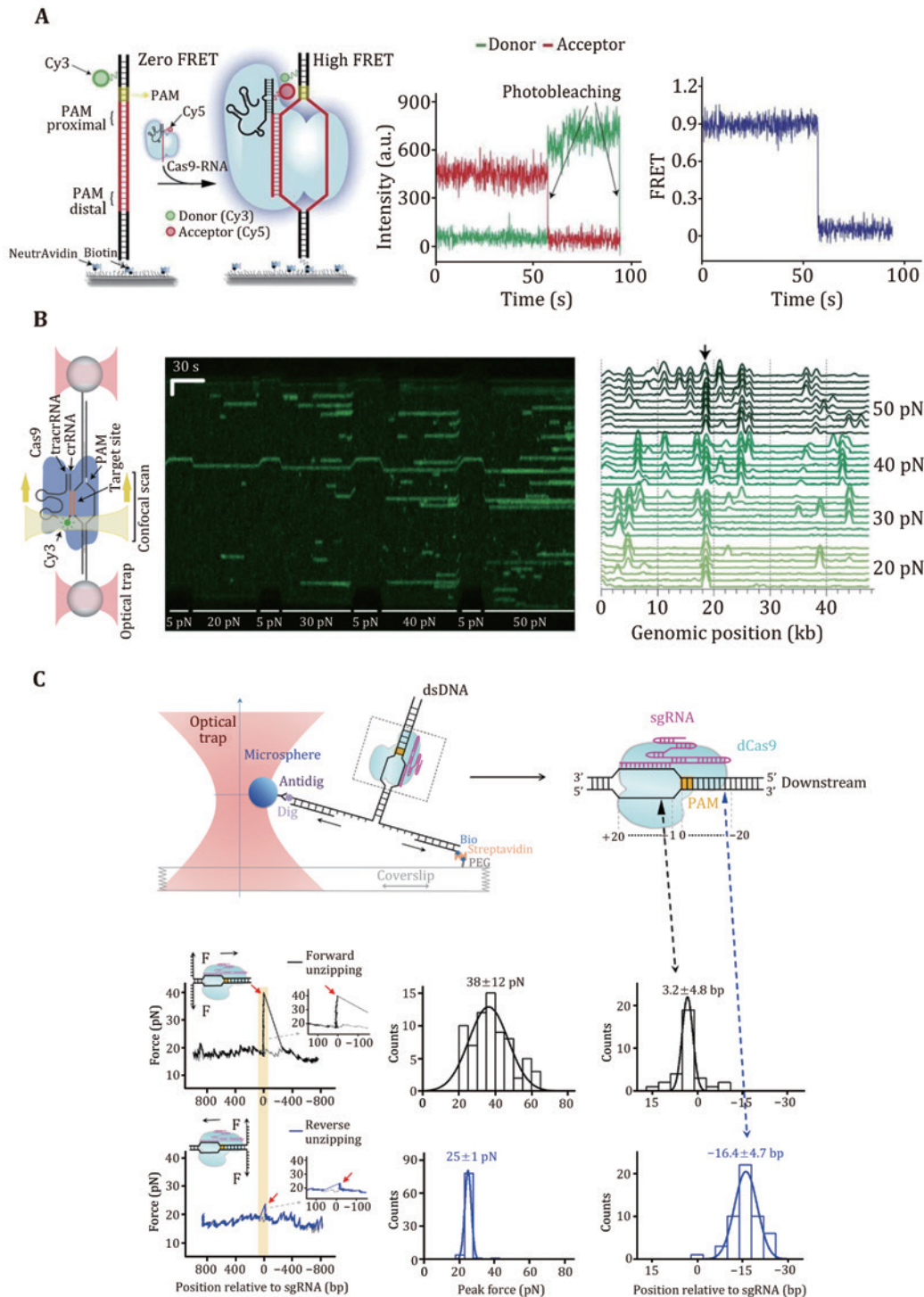


Fig. 3 DNA binding of SpCas9 revealed by single-molecule studies. **A** Schematic and a representative trace of the smFRET assay for the SpCas9 binding target. **B** Kymograph and time-binned intensity histogram of force-stretched lambda DNA in the presence of 5–50 pN force, and off-target binding occurs once the force is higher than 20 pN. **C** Schematic of the single-molecule DNA unzipping experiment and the target DNA coordinate definition. Representative trace, disruption force histogram, and ternary interaction position histogram of forward (black) and reverse unzipping (blue) in the presence of SpCas9 bound to the target DNA. Adapted from Singh *et al.* (2016), Newton *et al.* (2019), and Zhang *et al.* (2019) with permissions

of a three-strand nucleic acid structure known as the R-loop (Jiang and Doudna 2017; Sternberg *et al.* 2014). *In vitro* biochemical assays have demonstrated that SpCas9-gRNA can tolerate PAM-distal mismatches, whereas PAM-proximal mismatches in the first 8–12 nucleotides are more deleterious (Sternberg *et al.* 2014). These findings suggest a unidirectional unwinding mechanism wherein DNA unwinding initiates from the PAM-proximal “seed” DNA sequence and propagates to the PAM-distal region.

Magnetic tweezers were applied to study SpCas9 complexed with noncanonical gRNAs wherein 5' modifications to gRNA were made (Fig. 4A) (Mullally *et al.* 2020). A 5' addition of a 20-nt RNA hairpin to gRNA allowed for stable 9-bp R-loop formation, implying a discrete step in the unwinding of the protospacer DNA. Indeed, a magnetic-tweezer-based rotor bead tracking (RBT) study from the Bryant Laboratory exhibited a transient discrete intermediate in SpCas9-sgRNA-induced R-loop formation, consistent with RNA-DNA hybridization within an initial seed region (Ivanov *et al.*

2020).

Additional evidence for discrete DNA unwinding and R-loop formation came from a series of smFRET studies on the SpCas9 protein (Lim *et al.* 2016; Okafor *et al.* 2019; Singh *et al.* 2018; Zeng *et al.* 2018). In these studies, donor and acceptor dyes were separately placed on the TS and NTS within or around the protospacer region (or on the 5' end of sgRNA), and protospacer DNA unwinding would result in the separation of the two dyes, thus a decrease in the FRET efficiency. A mid-FRET state was occasionally detected between the intact and fully unwound DNA states, suggesting a partially unwound DNA state (Lim *et al.* 2016; Zeng *et al.* 2018). Moreover, PAM-distal mismatches were found to impair DNA unwinding, and fully unwound protospacer DNA required at least 17-bp crRNA-DNA complementarity (Fig. 4B) (Singh *et al.* 2018). Engineered Cas9s, such as eCas9, Cas-HF1, HypaCas9, and SniperCas9, were proven to increase efficiency by partially depopulating the fully unwound state more readily with mismatches (Fig. 4B) (Okafor *et al.* 2019; Singh *et al.* 2018).

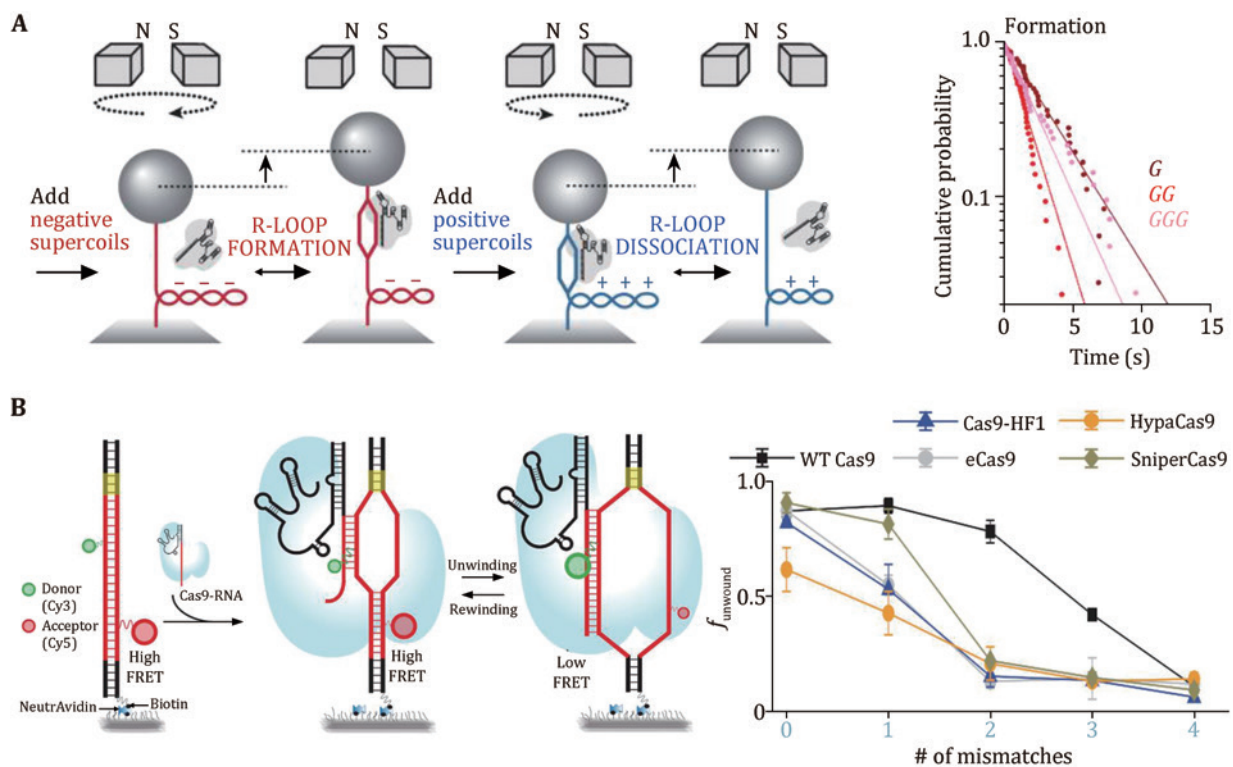


Fig. 4 DNA unwinding and R-loop formation of SpCas9 revealed by single-molecule studies. **A** Schematic of the magnetic tweezers for the DNA unwinding assay. The inverse cumulative probability distribution of the time required for R-loop formation using different gRNAs. G, GG, and GGG represent the number of additional guanines at the 5' end of gRNA. **B** Schematic of the smFRET assay for protospacer DNA unwinding by the SpCas9-gRNA complex. Statistical graph of the protospacer DNA unwinding probability with different numbers of DNA-RNA mismatches for the variety Cas9 mutants. Adapted from Mullally *et al.* (2020) and Okafor *et al.* (2019) with permissions

These single-molecule data support a model in which SpCas9-induced R-loop formation initiates from the PAM-proximal seed sequence and extends to the PAM-distal region. Before realizing the fully opened state, R-loop formation undergoes a partial intermediate state that is cleavage-incompetent. SpCas9-gRNA can remain stably bound to the DNA in this intermediate state (see the previous section “Stable binding to a DNA target”). Fully unwound protospacer DNA coupled with full R-loop formation possibly drives the docking of the HNH domain, thus licensing cleavage-competent SpCas9 (see the following section “DNA dissociation after cleavage”). Modifications of gRNA or the engineering of SpCas9 could rebalance the unwinding-rewinding equilibrium and make it stricter to reach the cleavage-competent state, thus minimizing off-target effects.

Conformational rearrangements in SpCas9 domains

SpCas9 is a multidomain DNA endonuclease. Structures of SpCas9 showed two distinct lobes, the alpha-helical recognition (REC) lobe and the nuclease lobe (NUS), as

well as the more variable C-terminal domain (CTD) (Jinek *et al.* 2014). The NUC lobe contains the conserved HNH and split RuvC nuclease domains that are responsible for cleaving the TS and NTS, respectively. Crystal structures of apo, sgRNA-bound, sgRNA/DNA-bound SpCas9 have revealed the distinct conformational states of the protein, indicating that the protein must undergo conformational changes along its reaction pathway (Jiang and Doudna 2017). Indeed, a comparison of the structures of SpCas9-sgRNA complex and apo-SpCas9 reveals that gRNA can drive a substantial structural rearrangement of SpCas9 to realize a DNA recognition-competent conformation (Jiang *et al.* 2015; Jinek *et al.* 2014). Moreover, upon target binding and R-loop formation, SpCas9 undergoes a further conformation rearrangement that positions the HNH nuclease domain for the TS cleavage (Jiang *et al.* 2016b). A bulk FRET experiment first proved that the HNH domain samples a conformational equilibrium from an inactive state to an activated conformation (Sternberg *et al.* 2015), which was later confirmed by an HF-AFM study (Fig. 5A) (Shibata *et al.* 2017). smFRET studies have further examined the mobility of

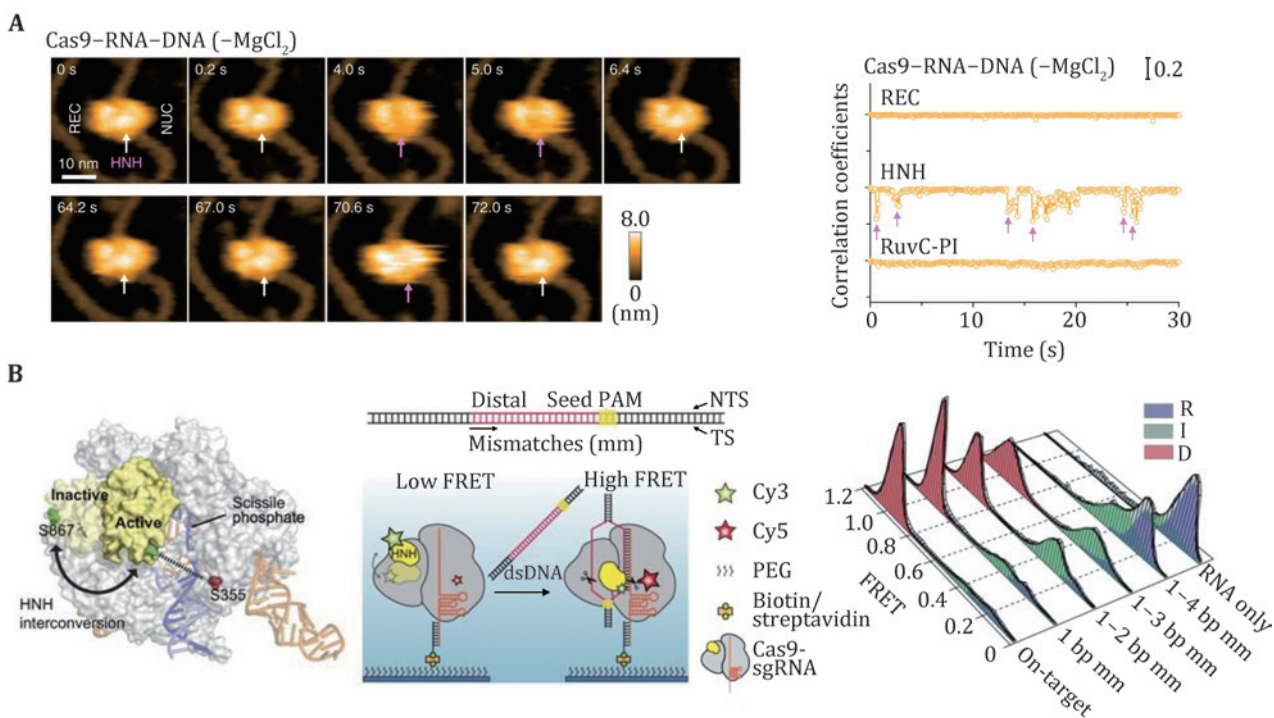


Fig. 5 Conformational rearrangements of SpCas9 revealed by single-molecule studies. **A** Dynamic HS-AFM images of Cas9/RNA/DNA ternary in the absence of magnesium ions. The statistical graph of the correlation coefficient of each domain over time is shown on the right. The white arrows indicate the HNH domain, while the magenta arrows indicate the dynamic change of the HNH domain. The scale bar is 10 nm. **B** Schematic of different SpCas9 conformations (based on PDB 4ZT0 and 5F9R) and the smFRET assay for HNH domain dynamics. The steady-state histograms of SpCas9 with different numbers of DNA-RNA mismatches. Adapted from Shibata *et al.* (2017) and Daggas *et al.* (2017) with permissions

the HNH domain of SpCas9. Using SpCas9 variants labeled with Cy3 and Cy5 dyes, three groups identified an intermediate state of SpCas9 between the open and closed states, which represents the conformational checkpoint between DNA binding and cleavage (Fig. 5B) (Dagdas *et al.* 2017; Osuka *et al.* 2018; Yang *et al.* 2018). High-fidelity SpCas9 variants display slow transition rates to the active conformation, thus enhancing cleavage activity (Singh *et al.* 2018). Furthermore, smFRET studies revealed that a noncatalytic domain, REC3, governs HNH domain mobility. Recently, smFRET studies also demonstrated that the HNH domain of SpCas9 after cleavage was highly flexible (Wang *et al.* 2021).

DNA dissociation after cleavage

One distinguished characteristic of SpCas9 is its stable binding to the on-target site after cleavage. Both *in vitro* and *in vivo* experiments have demonstrated that SpCas9 remains tightly bound to both ends of the cleaved DNA for hours, thus acting as a single-turnover enzyme (Jones *et al.* 2017; Ma *et al.* 2016a; Sternberg *et al.* 2014; Zhang *et al.* 2019). At the single-molecule level, a DNA curtain assay provided evidence that SpCas9 bridges the double-strand break (DSB) ends for a long period unless there is a harsh treatment with 7 mol/L urea (Sternberg *et al.* 2014). An optical tweezer assay further verified that the ternary complex could sustain a rupture force of up to 40 pN (Newton *et al.* 2019).

A fluorescence-based single-molecule digestion assay suggests that after cleavage by a SpCas9 nickase (SpCas9_{dHNH}), the 3' flap generated by the cleaved NTS is possibly exposed and can be digested by exonucleases (Fig. 6A) (Wang *et al.* 2021). Therefore, it is highly likely that different types of CRISPR proteins may employ distinct dissociation mechanisms after cleavage.

The long lifetime of the SpCas9–gRNA–DNA complex limits the efficient usage of each SpCas9 protein and impairs the repair of DSBs (Clarke *et al.* 2018). Increasing the slow off-rate of DNA-cleaved SpCas9 to DNA would be expected to improve its efficiency. Single-molecule experiments have demonstrated that DNA-based motor proteins, *in vitro*, could facilitate the dissociation of DNA-cleaved SpCas9 from DNA. Zhang *et al.* used optical tweezers to examine the consequence of encountering a BLM helicase with a DNA-bound dSpCas9 from both sides (Fig. 6B). They provided a proof of concept that, compared with the upstream side of the PAM, SpCas9 is more readily displaced from the downstream side of the PAM by BLM. These results highlight the importance of the post-PAM interaction in regulating DNA dissociation of SpCas9. Other motor proteins, such as Pif1, RNA polymerase, CMG helicase, and the histone chaperone FACT were also reported to be capable of dislodging DNA-bound SpCas9 (Clarke *et al.* 2018; Schauer *et al.* 2020; Vrtis *et al.* 2021; Wang *et al.* 2020).

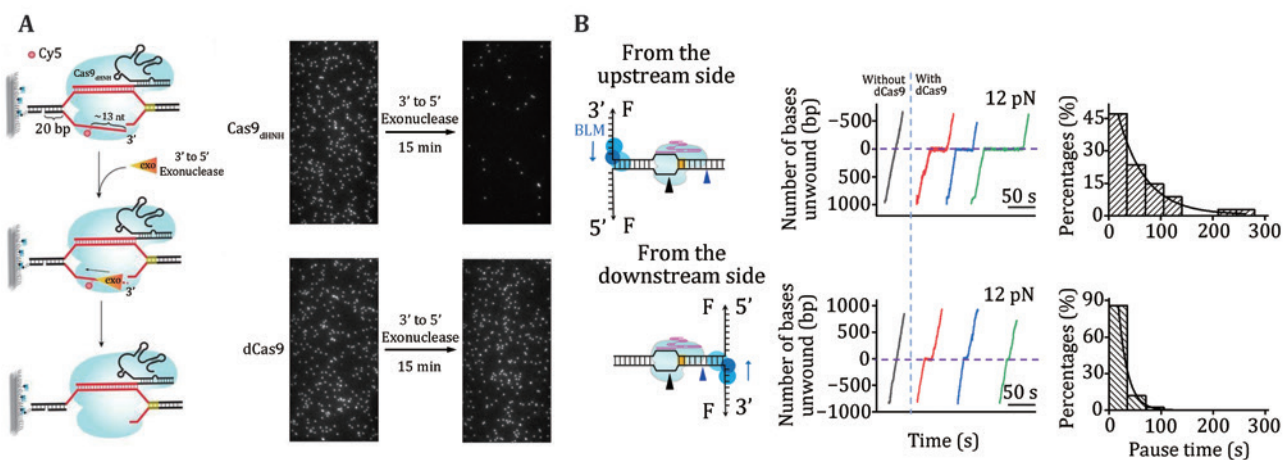


Fig. 6 Single-molecule detection of SpCas9 postcatalytic dissociation. **A** Schematic of the fluorescence observation experiment for the 3' flap NTS digestion. Representative images of the NTS digestion using SpCas9_{HNH} and SpCas9_{dHNH} before and after the Klenow fragments (the 3' to 5' exonuclease). **B** Schematic of the BLM helicase unwinding initiating from either the upstream (top) or downstream (bottom) side of the PAM. Representative traces show the number of unwound base pairs versus time under a constant force of 12 pN with or without dSpCas9. The histograms show the pause time of the BLM helicase at the expected dSpCas9 binding site. A single exponential fitting is used for these distributions. Adapted from Wang *et al.* (2021) and Zhang *et al.* (2019) with permissions

SUMMARY AND PERSPECTIVES

As evident from this review, single-molecule studies provide not only a fundamental understanding of Cas9 mechanisms but also a framework for rational design aiming at improving Cas9 efficiency and minimizing off-target effects. Based on these studies, a detailed dynamic picture of DNA interrogation and cleavage of SpCas9 has been generated (Fig. 7). Upon complexation with sgRNA, SpCas9 first uses a combination of 3D and 1D searching modes to target PAM. The binding of SpCas9 to the PAM initiates PAM-proximal protospacer DNA unwinding and an intermediate R-loop formation to examine crRNA–DNA complementarity. The first 8–10 bp crRNA–DNA matches are sufficient to support the stable binding of SpCas9–sgRNA to the target and will promote further unwinding of the protospacer DNA. Full R-loop formation will not be reached unless 17-nt or more protospacer DNA pairs with the crRNA. Complete annealing of crRNA and target DNA allows the HNH domain to reach a stable, active conformation for TS cleavage. Then, SpCas9 remains stably bound to both DSB ends wherein a cleaved NTS is exposed and

can be accessed by other proteins. DNA-based motor proteins may promote final DNA dissociation from the target for future DNA repair. We believe that single-molecule techniques will continue to contribute to the CRISPR field.

Acknowledgements This work was supported by the National Key Research and Development Program of China (2017YFA0106700), the National Natural Science Foundation of China (32022048 and 22104088), the Natural Science Foundation of Shanghai (19ZR1434100), and the China Postdoctoral Science Foundation (2021M692053). We sincerely apologize to authors whose work could not be included in this manuscript due to the space limitation.

Compliance with Ethical Standards

Conflict of interest Qian Zhang, Ziting Chen and Bo Sun declare that they have no conflict of interest.

Human and animal rights and informed consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any

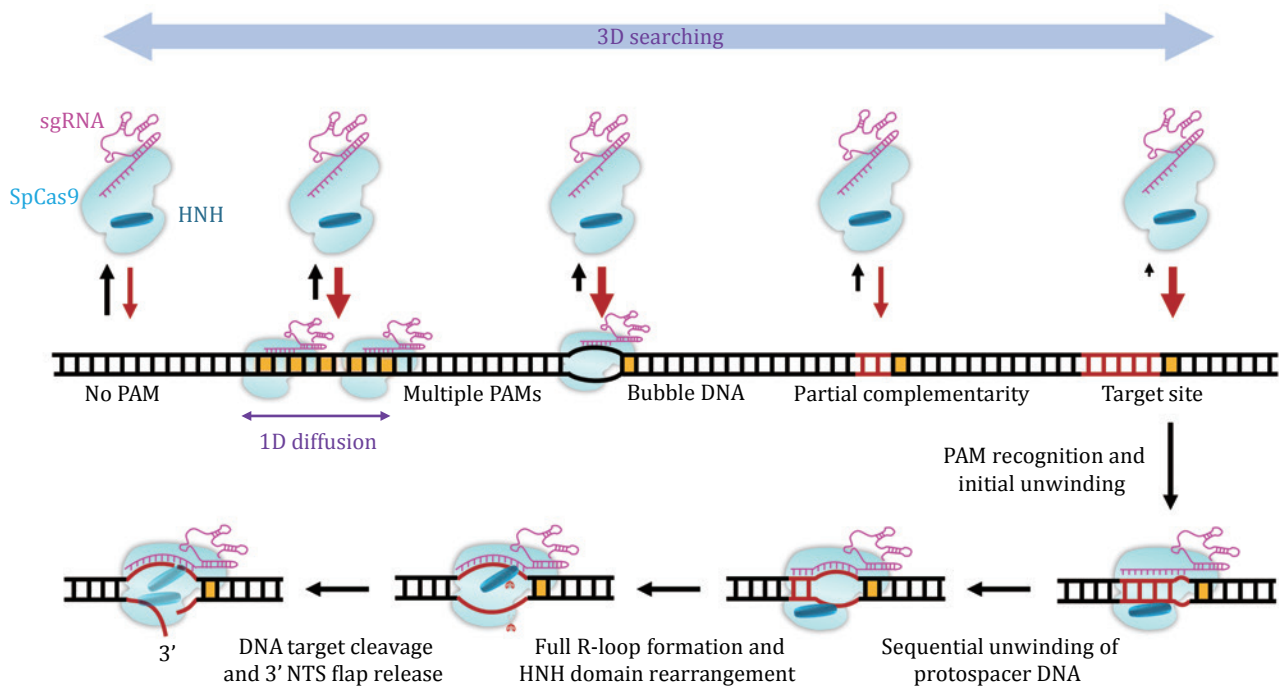


Fig. 7 A dynamic model for the interplay between SpCas9 and DNA. The PAM search is carried out through random 3D collision, and 1D diffusion is performed near the PAM in a close region. DNA bubbles and crRNA–DNA complementarity promote the binding of SpCas9–sgRNA to the DNA. DNA binding by SpCas9–sgRNA induces the unwinding of the PAM-proximal protospacer DNA, giving rise to the formation of an RNA–DNA heteroduplex. The R-loop expansion propagates to the PAM-distal region. Driven by the complete formation of the R-loop, the HNH domain is repositioned to the cleavage site and the DNA is cleaved. SpCas9–sgRNA remains bound to the cleaved site wherein the cleaved 3' flap NTS is first exposed

medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anders C, Niewoehner O, Duerst A, Jinek M (2014) Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* 513(7519): 569–573
- Ashkin A, Dziedzic JM, Bjorkholm JE, Chu S (1986) Observation of a single-beam gradient force optical trap for dielectric particles. *Opt Lett* 11(5): 288–290
- Axelrod D (1989) Total internal reflection fluorescence microscopy. *Methods Cell Biol* 30: 245–270
- Bak SY, Jung Y, Park J, Sung K, Jang H-K, Bae S, Kim SK (2021) Quantitative assessment of engineered Cas9 variants for target specificity enhancement by single-molecule reaction pathway analysis. *Nucleic Acids Res* 49(19): 11312–11322
- Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315(5819): 1709–1712
- Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res* 41(15): 7429–7437
- Binnig G, Quate CF, Gerber C (1986) Atomic force microscope. *Phys Rev Lett* 56(9): 930. <https://doi.org/10.1103/PhysRevLett.56.930>
- Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, Dickman MJ, Makarova KS, Koonin EV, Van Der Oost J (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321(5891): 960–964
- Bustamante CJ, Chemla YR, Liu S, Wang MD (2021) Optical tweezers in single-molecule biophysics. *Nat Rev Methods Primers* 1(1): 25. <https://doi.org/10.1038/s43586-021-00021-6>
- Carte J, Wang R, Li H, Terns RM, Terns MP (2008) Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev* 22(24): 3489–3496
- Charpentier E, Richter H, van der Oost J, White MF (2015) Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiol Rev* 39(3): 428–441
- Charvin G, Strick T, Bensimon D, Croquette V (2005) Tracking topoisomerase activity at the single-molecule level. *Annu Rev Biophys Biomol Struct* 34: 201–219
- Cheezum MK, Walker WF, Guilford WH (2001) Quantitative comparison of algorithms for tracking single fluorescent particles. *Biophys J* 81(4): 2378–2388
- Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li G-W, Park J, Blackburn EH, Weissman JS, Qi LS (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155(7): 1479–1491
- Chen JS, Dagdas YS, Kleinstiver BP, Welch MM, Sousa AA, Harrington LB, Sternberg SH, Joung JK, Yildiz A, Doudna JA (2017) Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* 550(7676): 407–410
- Clarke R, Heler R, MacDougall MS, Yeo NC, Chavez A, Regan M, Hanakahi L, Church GM, Marraffini LA, Merrill BJ (2018) Enhanced bacterial immunity and mammalian genome editing via RNA-polymerase-mediated dislodging of Cas9 from double-strand DNA breaks. *Mol Cell* 71(1): 42–55
- Cordes T, Moerner W, Orrit M, Sekatskii S, Faez S, Borri P, Goswami HP, Clark A, El-Khoury P, Mayr S (2015) Plasmonics, tracking and manipulating, and living cells: general discussion. *Faraday Discuss* 184: 451–473
- Cuculis L, Abil Z, Zhao HM, Schroeder CM (2016) TALE proteins search DNA using a rotationally decoupled mechanism. *Nat Chem Biol* 12(10): 831–837
- Cuculis L, Schroeder CM (2017) A single-molecule view of genome editing proteins: biophysical mechanisms for TALEs and CRISPR/Cas9. *Annu Rev Chem Biomol Eng* 8: 577–597
- Dagdas YS, Chen JS, Sternberg SH, Doudna JA, Yildiz A (2017) A conformational checkpoint between DNA binding and cleavage by CRISPR-Cas9. *Sci Adv* 3(8): ea00027. <https://doi.org/10.1126/sciadv.aao0027>
- Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471(7340): 602–607
- Fazio T, Visnapuu M-L, Wind S, Greene EC (2008) DNA curtains and nanoscale curtain rods: high-throughput tools for single molecule imaging. *Langmuir* 24(18): 10524–10531
- Finer JT, Simmons RM, Spudich JA (1994) Single myosin molecule mechanics: piconewton forces and nanometre steps. *Nature* 368(6467): 113–119
- Garneau JE, Dupuis M-È, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadán AH, Moineau S (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468(7320): 67–71
- Gasiunas G, Barrangou R, Horvath P, Siksnys V (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci USA* 109(39): E2579–E2586
- Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC (2014) Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* 159(3): 647–661
- Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154(2): 442–451
- Globyte V, Kim SH, Joo C (2018) Single-molecule view of small RNA-guided target search and recognition. *Annu Rev Biophys* 47: 569–593
- Globyte V, Lee SH, Bae T, Kim JS, Joo C (2019) CRISPR/Cas9 searches for a protospacer adjacent motif by lateral diffusion. *EMBO J* 38(4): e99466. <https://doi.org/10.15252/embj.201899466>
- Gong S, Yu HH, Johnson KA, Taylor DW (2018) DNA unwinding is the primary determinant of CRISPR-Cas9 activity. *Cell Rep* 22(2): 359–371
- Gorman J, Fazio T, Wang F, Wind S, Greene EC (2010) Nanofabricated racks of aligned and anchored DNA substrates for single-molecule imaging. *Langmuir* 26(2): 1372–1379
- Greene EC, Wind S, Fazio T, Gorman J, Visnapuu M-L (2010) DNA curtains for high-throughput single-molecule optical imaging. *Methods Enzymol* 472: 293–315
- Gupta P, Zlatanova J, Tomschik M (2009) Nucleosome assembly

- depends on the torsion in the DNA molecule: a magnetic tweezers study. *Biophys J* 97(12): 3150–3157
- Hayes RP, Xiao Y, Ding F, van Erp PB, Rajashankar K, Bailey S, Wiedenheft B, Ke A (2016) Structural basis for promiscuous PAM recognition in type I-E Cascade from *E. coli*. *Nature* 530(7591): 499–503
- Heler R, Samai P, Modell JW, Weiner C, Goldberg GW, Bikard D, Marraffini LA (2015) Cas9 specifies functional viral targets during CRISPR–Cas adaptation. *Nature* 519(7542): 199–202
- Hilton IB, D'ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, Gersbach CA (2015) Epigenome editing by a CRISPR–Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat Biotechnol* 33(5): 510–517
- Ivanov IE, Wright AV, Cofsky JC, Aris KDP, Doudna JA, Bryant Z (2020) Cas9 interrogates DNA in discrete steps modulated by mismatches and supercoiling. *Proc Natl Acad Sci USA* 117(11): 5853–5860
- Jiang C, Lionberger TA, Wiener DM, Meyhofer E (2016a) Electromagnetic tweezers with independent force and torque control. *Rev Sci Instrum* 87(8): 084304. <https://doi.org/10.1063/1.4960811>
- Jiang F, Doudna JA (2017) CRISPR–Cas9 structures and mechanisms. *Annu Rev Biophys* 46: 505–529
- Jiang F, Taylor DW, Chen JS, Kornfeld JE, Zhou K, Thompson AJ, Nogales E, Doudna JA (2016b) Structures of a CRISPR–Cas9 R-loop complex primed for DNA cleavage. *Science* 351(6275): 867–871
- Jiang F, Zhou K, Ma L, Gressel S, Doudna JA (2015) A Cas9–guide RNA complex preorganized for target DNA recognition. *Science* 348(6242): 1477–1481
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096): 816–821
- Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, Anders C, Hauer M, Zhou K, Lin S, Kaplan M, Iavarone AT, Charpentier E, Nogales E, Doudna JA (2014) Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* 343(6176): 1247997. <https://doi.org/10.1126/science.1247997>
- Jones DL, Leroy P, Unoson C, Fange D, Curic V, Lawson MJ, Elf J (2017) Kinetics of dCas9 target search in *Escherichia coli*. *Science* 357(6358): 1420–1424
- Jore MM, Lundgren M, Van Duijn E, Bultema JB, Westra ER, Waghmare SP, Wiedenheft B, Pul Ü, Wurm R, Wagner R (2011) Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat Struct Mol Biol* 18(5): 529–536
- 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 Biol* 10(5): 841–851
- Knott GJ, Doudna JA (2018) CRISPR–Cas guides the future of genetic engineering. *Science* 361(6405): 866–869
- Konermann S, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L, Platt RJ, Scott DA, Church GM, Zhang F (2013) Optical control of mammalian endogenous transcription and epigenetic states. *Nature* 500(7463): 472–476
- Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H (2015) Genome-scale transcriptional activation by an engineered CRISPR–Cas9 complex. *Nature* 517(7536): 583–588
- Koonin EV, Makarova KS, Zhang F (2017) Diversity, classification and evolution of CRISPR–Cas systems. *Curr Opin Microbiol* 37: 67–78
- Leake MC (2013) The physics of life: one molecule at a time. *Philos Trans R Soc Lond B Biol Sci* 368(1611): 20120248. <https://doi.org/10.1098/rstb.2012.0248>
- Lim Y, Bak SY, Sung K, Jeong E, Lee SH, Kim JS, Bae S, Kim SK (2016) Structural roles of guide RNAs in the nuclease activity of Cas9 endonuclease. *Nat Commun* 7(1): 13350. <https://doi.org/10.1038/ncomms13350>
- Liu L, Li X, Wang J, Wang M, Chen P, Yin M, Li J, Sheng G, Wang Y (2017) Two distant catalytic sites are responsible for C2c2 RNase activities. *Cell* 168(1–2): 121–134
- Ma H, Naseri A, Reyes-Gutierrez P, Wolfe SA, Zhang S, Pederson T (2015) Multicolor CRISPR labeling of chromosomal loci in human cells. *Proc Natl Acad Sci USA* 112(10): 3002–3007
- Ma H, Tu LC, Naseri A, Huisman M, Zhang S, Grunwald D, Pederson T (2016a) CRISPR–Cas9 nuclear dynamics and target recognition in living cells. *J Cell Biol* 214(5): 529–537
- Ma HH, Tu LC, Naseri A, Huisman M, Zhang SJ, Grunwald D, Pederson T (2016b) Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. *Nat Biotechnol* 34(5): 528–530
- Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK (2013) CRISPR RNA-guided activation of endogenous human genes. *Nat Methods* 10(10): 977–979
- Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, Brouns SJ, Charpentier E, Cheng D, Haft DH, Horvath P (2020) Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nat Rev Microbiol* 18(2): 67–83
- Manosas M, Meglio A, Spiering MM, Ding F, Benkovic SJ, Barre F-X, Saleh OA, Allemand JF, Bensimon D, Croquette V (2010) Magnetic tweezers for the study of DNA tracking motors. *Methods Enzymol* 475: 297–320
- Maragò O, Gucciardi P, Jones P (2010) Photonic force microscopy: from femtonewton force sensing to ultra-sensitive spectroscopy. In: Bhushan B (eds). *Scanning Probe Microscopy in Nanoscience and Nanotechnology*. Springer. pp : 23–56
- Marraffini LA, Sontheimer EJ (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322(5909): 1843–1845
- Marraffini LA, Sontheimer EJ (2010) CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* 11(3): 181–190
- Mullally G, Van Aelst K, Naqvi MM, Diffin FM, Karvelis T, Gasiunas G, Siksnys V, Szczelkun MD (2020) 5' modifications to CRISPR–Cas9 gRNA can change the dynamics and size of R-loops and inhibit DNA cleavage. *Nucleic Acids Res* 48(12): 6811–6823
- Newton MD, Taylor BJ, Driessen RPC, Roos L, Cvetesic N, Allyjaun S, Lenhard B, Cuomo ME, Rueda DS (2019) DNA stretching induces Cas9 off-target activity. *Nat Struct Mol Biol* 26(3): 185–192
- Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O (2014) Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* 156(5): 935–949
- Okafor IC, Singh D, Wang Y, Jung M, Wang H, Mallon J, Bailey S, Lee JK, 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(22): 11880–11888
- Osuka S, Isomura K, Kajimoto S, Komori T, Nishimasu H, Shima T, Nureki O, Uemura S (2018) Real-time observation of flexible domain movements in CRISPR–Cas9. *EMBO J* 37(10): e96941. <https://doi.org/10.15252/embj.201796941>
- Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, Polstein LR, Thakore PI, Glass KA, Ousterout DG, Leong KW (2013) RNA-guided gene activation by CRISPR–Cas9-based

- transcription factors. *Nat Methods* 10(10): 973–976
- Polidino P, Magazzu A, Iati MA, Patti F, Saija R, Boschi CDE, Donato MG, Gucciardi PG, Jones PH, Volpe G (2018) Optical tweezers and their applications. *J Quant Spectr Radiat Transfer* 218: 131–150
- Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152(5): 1173–1183
- Redding S, Sternberg SH, Marshall M, Gibb B, Bhat P, Guegler CK, Wiedenheft B, Doudna JA, Greene EC (2015) Surveillance and processing of foreign DNA by the *Escherichia coli* CRISPR-Cas system. *Cell* 163(4): 854–865
- Roy R, Hohng S, Ha T (2008) A practical guide to single-molecule FRET. *Nat Methods* 5(6): 507–516
- Sapranaukas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V (2011) The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res* 39(21): 9275–9282
- Sarkar R, Rybenkov VV (2016) A guide to magnetic tweezers and their applications. *Front Phys* 4: 48. <https://doi.org/10.3389/fphy.2016.00048>
- Sbalzarini IF, Koumoutsakos P (2005) Feature point tracking and trajectory analysis for video imaging in cell biology. *J Struct Biol* 151(2): 182–195
- Schauer GD, Spenkelink LM, Lewis JS, Yurieva O, Mueller SH, van Oijen AM, O'Donnell ME (2020) Replisome bypass of a protein-based R-loop block by Pif1. *Proc Natl Acad Sci USA* 117(48): 30354–30361
- Selvin PR (2000) The renaissance of fluorescence resonance energy transfer. *Nat Struct Biol* 7(9): 730–734
- Shibata M, Nishimasu H, Kodera N, Hirano S, Ando T, Uchihashi T, Nureki O (2017) Real-space and real-time dynamics of CRISPR-Cas9 visualized by high-speed atomic force microscopy. *Nat Commun* 8(1): 1430. <https://doi.org/10.1038/s41467-017-01466-8>
- Silas S, Mohr G, Sidote DJ, Markham LM, Sanchez-Amat A, Bhaya D, Lambowitz AM, Fire AZ (2016) Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase–Cas1 fusion protein. *Science* 351(6276): aad4234. <https://doi.org/10.1126/science.aad4234>
- Singh D, Ha T (2018) Understanding the molecular mechanisms of the CRISPR toolbox using single molecule approaches. *ACS Chem Biol* 13(3): 516–526
- Singh D, Sternberg SH, Fei JY, Doudna JA, Ha T (2016) Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9. *Nat Commun* 7(1): 12778. <https://doi.org/10.1038/ncomms12778>
- Singh D, Wang Y, Mallon J, Yang O, Fei J, Poddar A, Ceylan D, Bailey S, Ha T (2018) Mechanisms of improved specificity of engineered Cas9s revealed by single-molecule FRET analysis. *Nat Struct Mol Biol* 25(4): 347–354
- Smith SB, Cui Y, Bustamante C (1996) Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science* 271(5250): 795–799
- Staals RH, Agari Y, Maki-Yonekura S, Zhu Y, Taylor DW, Van Duijn E, Barendregt A, Vlot M, Koehorst JJ, Sakamoto K (2013) Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of *Thermus thermophilus*. *Mol Cell* 52(1): 135–145
- Staals RH, Zhu Y, Taylor DW, Kornfeld JE, Sharma K, Barendregt A, Koehorst JJ, Vlot M, Neupane N, Varossieau K (2014) RNA targeting by the type III-A CRISPR-Cas Csm complex of *Thermus thermophilus*. *Mol Cell* 56(4): 518–530
- Sternberg SH, LaFrance B, Kaplan M, Doudna JA (2015) Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature* 527(7576): 110–113
- Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507(7490): 62–67
- Sternberg SH, Richter H, Charpentier E, Qimron U (2016) Adaptation in CRISPR-Cas systems. *Mol Cell* 61(6): 797–808
- Strick T, Allemand J-F, Bensimon D, Croquette V (1998) Behavior of supercoiled DNA. *Biophys J* 74(4): 2016–2028
- Strick TR, Allemand J-F, Bensimon D, Bensimon A, Croquette V (1996) The elasticity of a single supercoiled DNA molecule. *Science* 271(5257): 1835–1837
- Sun B (2019) Rescuing replication from barriers: mechanistic insights from single-molecule studies. *Mol Cell Biol* 39(10): e00576–18
- Sun B, Wang MD (2016) Single-molecule perspectives on helicase mechanisms and functions. *Crit Rev Biochem Mol Biol* 51(1): 15–25
- Sung K, Park J, Kim Y, Lee NK, Kim SK (2018) Target specificity of Cas9 nuclease via DNA rearrangement regulated by the REC2 domain. *J Am Chem Soc* 140(25): 7778–7781
- Szczelkun MD, Tikhomirova MS, Sinkunas T, Gasiunas G, Karvelis T, Pschera P, Siksnys V, Seidel R (2014) Direct observation of R-loop formation by single RNA-guided Cas9 and cascade effector complexes. *Proc Natl Acad Sci USA* 111(27): 9798–9803
- Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD (2014) A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* 159(3): 635–646
- Thakore PI, D'ippolito AM, Song L, Safi A, Shivakumar NK, Kabadi AM, Reddy TE, Crawford GE, Gersbach CA (2015) Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat Methods* 12(12): 1143–1149
- Vrtis KB, Dewar JM, Chistol G, Wu RA, Graham TG, Walter JC (2021) Single-strand DNA breaks cause replisome disassembly. *Mol Cell* 81(6): 1309–1318
- Wang AS, Chen LC, Wu RA, Hao Y, McSwiggen DT, Heckert AB, Richardson CD, Gowen BG, Kazane KR, Vu JT (2020) The histone chaperone FACT induces Cas9 multi-turnover behavior and modifies genome manipulation in human cells. *Mol Cell* 79(2): 221–233
- Wang MD, Schnitzer MJ, Yin H, Landick R, Gelles J, Block SM (1998) Force and velocity measured for single molecules of RNA polymerase. *Science* 282(5390): 902–907
- Wang Y, Mallon J, Wang H, Singh D, Jo MH, Hua B, Bailey S, Ha T (2021) Real-time observation of Cas9 postcatalytic domain motions. *Proc Natl Acad Sci USA* 118(2): e2010650118. <https://doi.org/10.1073/pnas.2010650118>
- Wei Y, Terns RM, Terns MP (2015) Cas9 function and host genome sampling in Type II-A CRISPR–Cas adaptation. *Genes Dev* 29(4): 356–361
- Whinn KS, van Oijen AM, Ghodke H (2019) Spy-ing on Cas9: single-molecule tools reveal the enzymology of Cas9. *Curr Opin Biomed Eng* 12: 25–33
- Yang M, Peng S, Sun R, Lin J, Wang N, Chen C (2018) The conformational dynamics of Cas9 governing DNA cleavage are revealed by single-molecule FRET. *Cell Rep* 22(2): 372–382
- Yang M, Sun R, Deng P, Yang Y, Wang W, Liu J-JG, 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(38): 12776–12784
- Zeng Y, Cui Y, Zhang Y, Zhang Y, Liang M, Chen H, Lan J, Song G, Lou J (2018) The initiation, propagation and dynamics of CRISPR-SpyCas9 R-loop complex. *Nucleic Acids Res* 46(1): 350–361
- Zhang F (2019) Development of CRISPR-Cas systems for genome

- editing and beyond. *Quart Rev Biophys* 52: e6. <https://doi.org/10.1017/S0033583519000052>
- Zhang Q, Chen Z, Wang F, Zhang S, Chen H, Gu X, Wen F, Jin J, Zhang X, Huang X, Shen B, Sun B (2021) Efficient DNA interrogation of SpCas9 governed by its electrostatic interaction with DNA beyond the PAM and protospacer. *Nucleic Acids Res* 49(21): 12433–12444
- Zhang Q, Wen F, Zhang S, Jin J, Bi L, Lu Y, Li M, Xi XG, Huang X, Shen B, Sun B (2019) The post-PAM interaction of RNA-guided spCas9 with DNA dictates its target binding and dissociation. *Sci Adv* 5(11): eaaw9807. <https://doi.org/10.1126/sciadv.aaw9807>
- Zhang S, Zhang Q, Hou XM, Guo L, Wang F, Bi L, Zhang X, Li HH, Wen F, Xi XG, Huang X, Shen B, Sun B (2020) Dynamics of *Staphylococcus aureus* Cas9 in DNA target association and dissociation. *EMBO Rep* 21(10): e50184. <https://doi.org/10.15252/embr.202050184>
- Zhang X, Ma L, Zhang Y (2013a) High-resolution optical tweezers for single-molecule manipulation. *Yale J Biol Med* 86(3): 367–383
- Zhang Y, Heidrich N, Ampattu BJ, Gunderson CW, Seifert HS, Schoen C, Vogel J, Sontheimer EJ (2013b) Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. *Mol Cell* 50(4): 488–503