Dynamics of Target DNA Binding and Cleavage by Staphylococcus aureus Cas9 as Revealed by High-Speed Atomic Force Microscopy

Dynamics of target DNA binding and cleavage by *Staphylococcus aureus* Cas9 as revealed by high-speed atomic force microscopy

Leonardo Puppulin[1,2,*], Junichiro Ishikawa[3], Ayumi Sumino[1,4], Arin Marchesi[5], Holger Flechsig[1], Kenichi Umeda[1], Noriyuki Kodera[1], Hiroshi Nishimasu[3,6,7,8], and Mikihiro Shibata[1, 4, *]

[1] WPI Nano Life Science Institute (WPI-NanoLSI), Kanazawa University, Kanazawa, Ishikawa 920-1192, Japan.

[2]Department of Molecular Sciences and Nanosystems, Università Ca' Foscari Venezia, Via Torino 155, I-30172 Mestre, Venice, Italy

[3]Structural Biology Division, Research Center for Advanced Science and Technology, The University of Tokyo, Meguro-ku, Tokyo 153-8904, Japan.

[4]Institute for Frontier Science Initiative, Kanazawa University, Kanazawa, Ishikawa 920-1192, Japan.

[5]Department of Physiology, Polytechnic University of Marche, Ancona, Italy.

[6]Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan.

[7]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan.

[8]Inamori Research Institute for Science, Shimogyo-ku, Kyoto 600-8411, Japan.

*Corresponding authors. Email: leonardo.puppulin@unive.it, msshibata@staff.kanazawa-u.ac.jp

Abstract

Programmable DNA binding and cleavage by CRISPR-Cas9 has revolutionized the life sciences. However, the off-target cleavage observed in DNA sequences with some homology to the target still represents a major limitation for a more widespread use of Cas9 in biology and medicine. For this reason, complete understanding of the dynamics of DNA binding, interrogation and cleavage by Cas9 is crucial to improve the efficiency of genome editing. Here, we use high-speed atomic force microscopy (HS-AFM) to investigate *Staphylococcus aureus* Cas9 (SaCas9) and its dynamics of DNA binding and cleavage. Upon binding to single-guide RNA (sgRNA), SaCas9 forms a close bilobed structure that transiently and flexibly adopts also an open configuration. The SaCas9-mediated DNA cleavage is characterized by release of cleaved DNA and immediate dissociation, confirming that SaCas9 operates as a multiple turnover endonuclease. According to present knowledge, the process of search for target DNA is mainly governed by three-dimensional diffusion. Independent HS-AFM experiments show a potential long-range attractive interaction between SaCas9-sgRNA and its target DNA. The interaction precedes the formation of the stable ternary complex and is observed exclusively in the vicinity of the protospacer-adjacent motif (PAM), up to distances of several nanometers. The direct visualization of the process by sequential topographic images suggests that SaCas9-sgRNA binds to the target sequence first, while the following binding of the PAM is accompanied by local DNA bending and formation of the stable complex. Collectively, our HS-AFM data reveal a potential and unexpected behavior of SaCas9 during the search for DNA targets.

Keywords

CRISPR/Cas9, genome editing, SaCas9, high-speed atomic force microscopy, single-molecule imaging

Introduction

In recent years, the life sciences have been profoundly marked by the discovery of clustered regularly interspaced short palindromic repeats (CRISPR) in bacteria genomes. Unveiling the existence of CRISPR allowed to clarify the bacterial antiviral defense system for selective recognition and neutralization of foreign DNA based on RNA-guided endonuclease CRISPR-associated proteins 9 (Cas9)[1-4]. As the first studies of CRISPR-Cas9 were disclosed to the scientific community, it was immediately clear that such a sophisticated system could be utilized in genetic engineering as a new formidable platform for nimble genome editing. In fact, Cas9 can be activated via association with an artificial single-guide RNA (sgRNA), which contains a short segment of 20 nucleotides (i.e., spacer) that complements the selected sequence of the target DNA to cleave[5]. The Cas9-sgRNA binary complex is capable of interrogating DNA, searching for complementarity between the spacer and the target sequence, and ultimately generating a double-strand break on the DNA. This endonuclease recognizes and cleaves the correct DNA target only in the presence of a defined sequence of nucleotides following the non-complementary strand of the target, which is known as the protospacer-adjacent motif (PAM)[6-7]. The ease of bioengineering sgRNA with spacer incorporating any desired sequence has opened the possibility of harnessing Cas9 for facile genome editing by programmable DNA cleavage or genome manipulation by programmable binding of a catalytically inactive Cas9 mutant[8-12]. Nonetheless, a main limitation for a more widespread use of Cas9 in biology and medicine is represented by the occurrence of off-target cleavage observed in DNA sequences with some homology to the target[13-16].

In light of these considerations, major efforts are currently devoted to elucidating the molecular dynamics of DNA binding, interrogation and cleavage by Cas9-sgRNA. Based on current knowledge from studies by single-molecule fluorescence spectroscopy, the process of search for target DNA is governed by Cas9-sgRNA three-dimensional diffusion and random collisions[7], presumably facilitated by one-dimensional sliding along DNA[6]. Structural studies showed that Cas9 possesses a multi-domain structure in which different domains moves and reorganize during the formation of the binary complex with sgRNA and, later, the ternary complex with target DNA[17-25]. In addition, single-molecule Förster resonance energy transfer microscopy (smFRET) analyses[16, 26] and molecular dynamics simulations[27-29] suggest high flexibility of Cas9 during interaction with the target DNA, from binding to the final step of hydrolysis of the two DNA strands. However, direct and conclusive experimental evidence has not been established to fully clarify the conformational dynamics of Cas9-sgRNA during the entire process of ternary complex formation and activation. Furthermore, small Cas9 orthologs, such as Cas9 from *Staphylococcus aureus* (SaCas9), have not yet been extensively investigated to unveil

possible peculiarities in their conformational dynamics or similarities with the Cas9 isolated from *Streptococcus pyogenes* (SpCas9), which is the most frequently used ortholog, to date.

High-speed atomic force microscopy (HS-AFM) is a powerful technique that enables real-space and real-time visualization of the functional dynamics of proteins, polynucleotides, lipid membranes and their complexes at the nanometer scale, which are not feasible by other techniques[30-40]. In the present investigation, we utilized HS-AFM to study SaCas9 and visualize the structural configurations of apo-SaCas9, its binary complex with sgRNA, the ternary complex with sgRNA and DNA and the dynamics of target cleavage. Next, we collected unique experimental observations of target DNA search by SaCas9-sgRNA and formation of the ternary complex. As a preliminary step to acquire clear evidence of molecular binding by HS-AFM, we calibrated the functionalization of mica surfaces by aminosilane to obtain a trade-off between sufficiently high molecular mobility to permit the reaction to occur during imaging and sufficiently strong interaction with the AFM substrate to slow down and visualize the molecules by HS-AFM. The engineered AFM substrate enabled us to capture the binding reaction within time frames accessible to HS-AFM, confirm the authenticity of the interaction and discern different sequential phases of the process, which reveal unsuspected behaviors of the endonuclease.

Results

The SaCas9-sgRNA binary complex adopts a flexible bilobed structure

First, we visualized by HS-AFM the structural configurations of apo-SaCas9 and the SaCas9-sgRNA binary complex added to mica surfaces treated with 3-aminopropyll-triethoxysilane (APTES), hereafter referred as AP-mica. Figures 1a and 1b show the domain structure of SaCas9 and the crystal structure of SaCas9-sgRNA-target DNA complex[21], respectively. Like the other Cas9 orthologs, SaCas9 is composed of multiple domains that organize into a closed bilobed architecture upon binding to sgRNA and target DNA. The recognition lobe (REC) and the nuclease lobe (NUC) are connected by an arginine-rich bridge helix (BH) and a linker loop. The REC lobe interacts with the sugar phosphate backbones of the target DNA strand and the PAM-distal region of sgRNA. The NUC lobe contains the PAM-interacting domain (PI) and the two endonuclease domains HNH and RuvC, which cleave the target and non-target DNA strands, respectively. These two domains are connected by two linkers (L1 and L2). The WED domain in the NUC lobe is responsible for the sgRNA recognition. As shown in the representative sequential HS-AFM images in Figure 1c and Movie S1, HS-AFM observations of apo-SaCas9 revealed a highly flexible modular conformation, which is similar to that visualized in our previous HS-AFM experiments on the ortholog SpCas9[36]. In contrast, as shown in the selected HS-AFM sequential images of Figure 1d and Movie S2,

the association of sgRNA to SaCas9 stabilizes the structure into a bilobed conformation. Although crystal structures of SaCas9-sgRNA have not yet been reported, the homology with SpCas9 suggests that the domain reorganization of the binary complex in SaCas9 may comprehend the REC lobe and the NUC lobe. The latter is formed by the interactions of RuvC domain with HNH, PI and WED domains that appear to be stabilized by sgRNA. In previous observations of the SpCas9-sgRNA binary complex by HS-AFM[36], the bilobed structure showed stability in a close conformation that resembles those observed in the crystal structures of sgRNA-bound SpCas9[20]. Nonetheless, more recent studies by cryo-electron microscopy (cryo-EM) revealed the existence of a new conformation of SpCas9-sgRNA, in which the REC lobe rotates away from the NUC lobe into an open-protein conformation[24]. Indeed, in SaCas9-sgRNA, we observed some flexibility during imaging, that is higher than the one observed in the SpCas9 ortholog and allows the complex to assume also transient open configurations in which the two lobes are even spaced of several nm. The cross-correlation analysis of the sequential HS-AFM images of apo-SaCas9 and SaCas9-sgRNA confirmed the high conformational flexibility of apo-SaCas9, that is drastically reduced but not completely suppressed by sgRNA binding (Figures 1e and S1). Tracking analysis of the distance between the mean center positions of the two lobes during HS-AFM imaging of SaCas9-sgRNA further highlighted that the binary complex can transiently adopt open configurations (e.g., frames at 4.0 s and 8.6 s in Figures 1d, 1f and S2), which would provide the dsDNA substrate with accessibility to the binding sites. The difference between the most compact and the most open bilobed structures was estimated to be equal to 6.5 nm. Flexibility of apo-SaCas9 may facilitate the incorporation of sgRNA to form the binary complex, while the newly observed preservation of flexibility in SaCas9-sgRNA may be instrumental for the formation of the ternary complex with DNA, as we will see and discuss later.

SaCas9-sgRNA cleaves target DNA and rapidly dissociates

Next, we performed HS-AFM experiments on the ternary complex with DNA added to an AP-mica surface. SaCas9 recognizes a 5'-NNGRRT-3' PAM (where N is any nucleotide and R represents a purine, A or G). We prepared a double-stranded DNA substrate (dsDNA) with a 5'-CAGAAT-3' sequence in the non-target strand (NTS) following a 20-nucleotides sequence equal to the spacer of the sgRNA (Supporting Information Data 1). Figure 2a reports the schematic of the designed dsDNA, which elucidates the positions of the target sequence and the PAM. Figure 2b shows a representative HS-AFM topographic image of the ternary complex and the relative surface representation of the molecular structure, from which we can infer the locations of each SaCas9 domain. The optimized molecular orientation of the structure was obtained by employing an automatized fitting routine[41-42] that identifies the best match between the experimental AFM image (Figure 2b) and the simulated one (Figure 2c). In the topographic images of this HS-AFM experiment, the

ternary complex lying on the AP-mica resembles the shape of a clover, in which the HNH endonuclease domain is located on the lower left leaf. In addition, the binding of SaCas9-sgRNA induced local bending in the target DNA, which is consistent with the crystal structure of the ternary complex[21]. The high correlation between the HS-AFM observation and the simulated AFM image indicates that also at room temperature and in liquid, the HNH domain is located above the RuvC domain, far from the target DNA strand, as in the crystal structure[21]. The ternary complex preparation and the initial HS-AFM imaging were performed in buffer solution without Mg^{2+}, which is required for DNA cleavage. Next, we observed the activity of SaCas9 by adding Mg^{2+} to the imaging buffer during HS-AFM experiments. Movie S3 and the selected sequential topographic images in Figure 2d show the ternary complex after Mg^{2+} addition (i.e., final concentration 2 mM), which indeed triggered DNA cleavage by SaCas9-sgRNA. Interestingly, the SaCas9-sgRNA complex released the DNA on the non-PAM side and immediately dissociated from the DNA on the PAM side, collapsing on the mica surface. In addition, we observed also fast dissociation of sgRNA that quickly moved away from the collapsing SaCas9. As shown in Figure S3a, using the same color scale as in Figure 1c, the topography of the apo-SaCas9 acquired after dissociation from sgRNA and DNA resembles the typical topography of the apo-SaCas9 added to an AP-mica shown in Figure S3b. The observation of the same structures with the loose and flexible domains suggests that SaCas9 was not disrupted by the force of the HS-AFM probe. In fact, when the dissociation of the SaCas9-sgRNA from the target DNA is caused by the HS-AFM tip, the structure of apo-Cas9 appears modified and damaged, as shown in Movie S4 and Figure S3c, in which the ternary complex was observed for long time in buffer solution without addition of Mg^{2+}. In addition, all of our HS-AFM observations of DNA cleavage (i.e., 32 cases) show simultaneous release of cleaved DNA and separation of SaCas9-sgRNA (see further examples in Movie S5). Accordingly, the post-cleavage rapid dissociation may be a peculiar behavior of this Cas9 ortholog. Next, we characterized the conformational change of the complex in Figure 2d by tracking the position of the HNH domain (indicated by magenta arrowheads). Figure 2e shows the in-plane trajectory (x,y) of the mean center and the variation of the maximum height (see *Materials and methods* for more details). We considered the four consecutive frames starting from 14.58 s as representative of the endonuclease activity of SaCas9. Unfortunately, the temporal resolution of HS-AFM is not sufficient to capture the entire process of HNH translocation to the cleavage site. Nonetheless, in the frame at which the cleaved DNA appears for the first time (i.e., at 15.13 s), the topography of the complex changes dramatically and shows the clear appearance of a distinct protrusion above the collapsing clover. We assigned this new topographic feature to the HNH domain that moved to the DNA cleavage site. The explanatory surface structures in Figure 2f are two tentative configurations in agreement to the experimental ternary complexes shown in the images at 14.58 s and 15.13 s (i.e.,

HNH tracking position 1 and 2 in Figure 2d). Accordingly, the structure of SaCas9-sgRNA rotated about 20° clockwise after DNA cleavage. The protrusion assigned to HNH domain is located between the REC and NUC lobes, exactly where the cleavage site is located. The variation of maximum height measured between the tracking positions 1 and 2 is 0.82 nm (Figure 2e). All these experimental evidences suggest that the HNH domain was captured by HS-AFM above the cleavage site of the target DNA strand immediately after cleavage and before SaCas9-sgRNA dissociation. In Figure S4, we report a further example of DNA cleavage, in which the analysis of the sequential HS-AFM images confirm the dynamics observed in Figure 2.

Ternary complex formation at the target site is promoted by potential selective long-range attractive forces

In the attempt of studying the formation of the ternary complex, we performed HS-AFM imaging on mica surfaces weakly functionalized with APTES (hereafter referred as wAP-mica). The SaCas9-sgRNA and the DNA substrate were added separately to the surface in order to locate DNA substrates in proximity to certain endonucleases at the beginning of HS-AFM imaging. The cartoon image in Figure 3a elucidates the peculiar experimental conditions designed for this specific type of HS-AFM observation. The functionalization of mica by APTES was calibrated to reduce but not suppress the mobility of the molecules during the observations. In fact, the surface of a wAP-mica is characterized by scattered distribution of negative and positive charges that transiently attract and repulse the biomolecules during imaging, as shown in the explanatory cartoon image of Figure 3b. In the optimized preparation, APTES was diluted 1:40000 in Milli-Q water, added to the mica surface and incubated for 3 min. In this experimental condition, we captured the formation of the ternary complex with sufficient temporal and spatial resolutions to unveil and discriminate different stages of the binding process. Movie S6 and the selected sequential images of Figure 3c show an example of ternary complex formation visualized by HS-AFM on a wAP-mica. At the beginning of the imaging, we observed a hopping SaCas9-sgRNA in proximity to a dsDNA. The direction of motion of the binary complex was initially parallel to the dsDNA, at a distance of a few tens of nm. Surprisingly, the binary complex abruptly changed its motion toward the DNA to form a molecular complex (see image at 3.5 s in Figure 3c). In order to confirm the formation of the stable complex at the expected target site, we measured the length of the shorter segment of free DNA in every frame following molecular engagement. Considering the ideal case of fully immobilized dsDNA with distance between two consecutive base pairs of 0.34 nm, we estimated the total length of our dsDNA and the length from the PAM to the 3' end of the non-target strand (NTS 3') to be equal to 204 nm and 65.3 nm, respectively (see also dsDNA design in Figure 2a). In Figure 3d, we report the experimental length of the shorter segment of dsDNA after ternary complex formation as a function of time. The measurements of length are only slightly shorter than the ideal length representing the position of PAM. These small deviations from the ideal case can be explained by the lack of tight immobilization of dsDNA on a wAP-mica. In fact,

dsDNA is not completely immobilized on the surface during HS-AFM imaging and it should be wavy as shown in Figure 3b. Overall, these results confirm the stable binding of SaCas9-sgRNA to dsDNA at the expected target site. In Figure 3e, we show the enlarged area of the frame at 3.5 s, in which the molecular binding occurred during AFM-tip scanning. In our HS-AFM set-up, the total topographic image is composed by consecutive line scans along the fast scan axis, that is the *x-axis*. Each line scan, after the retrace of the previous one, is shifted of a preselected step, Δy, along the *y-axis*, from the bottom to the top of the image. In the fast axis, data of height are acquired at preselected step, Δx. During scanning, for each *y* position, only the trace from the left to the right is considered to reconstruct the topographic image, while the retrace is neglected. In our measurements, we acquired images with square pixels, namely Δx and Δy are equal. Keeping this knowledge in mind, we can conclude that the discontinuous topography of SaCas9-sgRNA in the frame at 3.5 s is due to the rapid translocation of the binary complex to the target site of the DNA. Although the force applied by the HS-AFM tip moving from the left to the right may have caused the movement of the SaCas9-sgRNA to the target site with fortuitous accuracy, also the occurrence of a selective molecular attraction cannot be ruled out. In fact, typical nano-manipulations of molecules in contact with the surface by the AFM tip apex, such as kicking or dragging, produce peculiar topographic features that are not found in Figure 3e. As can be inferred from the enlarged area reported in Figure 3, the binding event occurred during two consecutive AFM-tip line scans. Point A belongs the first of these two line scans, while point B belongs to the successive line scan. The first line scan includes only the topography of SaCas9-sgRNA hopping on the substrate prior binding (i.e., on the left of the image), while, conversely, the successive line scan shows only the topography of the ternary complex already formed (i.e., on the right of the image). In case the AFM-tip had kicked SaCas9-sgRNA toward the target DNA, two possible scenarios may be the observation of SaCas9-sgRNA already bound to the DNA before point A or, alternatively, some transient trace of the SaCas9-sgRNA topography in point B (i.e., heights higher than 5 nm). In case SaCas9-sgRNA had been dragged by the AFM-tip toward the target DNA, the pixels acquired before point A or, alternatively, after point B should also show trace of the topography of SaCas9-sgRNA moving toward the DNA. Conversely, in Figure 3e, we can clearly observe the unperturbed topography of the DNA and the substrate before point A, as well as the unperturbed topography of the substrate after point B. It is worth mentioning that also the interaction of the AFM-tip with the SaCas9-sgRNA entirely lifted off the surface in point B (i.e., not detected in the topography image) may have caused the abrupt translocation to the binding site. If we consider the translocation triggered by a long range interaction, it should have occurred within the time spent by the AFM-tip to move from point A to point B, which we estimated equal to 3.6 ms (i.e., we measured an AFM-tip travelled distance of 345.8 nm, including one retrace of 200 nm, at speed of 96.8 nm/ms). Considering that SaCas9-sgRNA prior to binding appears to be separated by 11 nm from the dsDNA (Figure 3e), the speed at which SaCas9-sgRNA moved to the target site is at least higher than 3 nm/ms, that is two order of magnitude higher than the speed of translation on wAP-mica

during hopping, estimated in 0.02 nm/ms from the two sequential HS-AFM images prior binding (details in Figure S5). Based on these considerations, the sudden change of direction that led to the ternary complex formation may have been induced by a potential long-range and selective attractive force between SaCas9-sgRNA and the target DNA. The range of interaction of 11 nm measured from the experimental evidence in Figure 3e may be overestimated of few nanometers, due to possible hopping of the SaCas9-sgRNA toward right during the AFM-tip scan from A to B, before the onset of the attractive force (i.e., translation undetected by HS-AFM).

A more convincing evidence of this type of interaction and accurate estimate of the range are provided by the HS-AFM observations shown in Movie S7 and the selected sequential images of Figure 4a. In this case, at the beginning of the HS-AFM imaging, the dsDNA was hopping in proximity of a SaCas9-sgRNA complex. The green arrowheads in Figure 4a indicate the position of the PAM in each frame, which was estimated as explained later. The SaCas9-sgRNA remained in a stable position and did not even flinch at the approaching DNA until the PAM moved closer to the SaCas9-sgRNA and reached the distance of 8 nm (frame at 56.4 s in Figure 4a). At this point, we observed the onset of a clear long-range interaction that led to the following events: i) increase of SaCas9-sgRNA mobility; ii) dramatic opening and separation of the SaCas9-sgRNA into two protrusions (i.e., presumably NUC and REC lobes); iii) translocation of one of the protrusions to the opposite side of the dsDNA with the linker region above the target sequence; iv) final engagement of the DNA with the SaCas9-sgRNA to form a stable ternary complex. Interestingly, after engagement (i.e., from the frame at 60.8 s in Figure 4a), the ternary complex underwent to further structural modification. In fact, as shown in the selected sequential HS-AFM images of Figure 4b, after binding we clearly observed two distinct configurations of the complex: the first one shows a linear DNA bound to the endonuclease, while the second configuration is formed upon sudden bending and reorientation of the upper segment of DNA by the SaCas9-sgRNA (i.e., frame at 114.5 s in Figure 4b). Next, we calculated the length of this DNA segment in every HS-AFM image, starting from the first frame after molecular binding. In Figure 4c are reported the results of this analysis, in which data were divided into two groups, according to the two configurations previously described (i.e., hereafter referred as linear DNA (*) and bent DNA (**)). The lengths are shown as percentage of the mean length calculated for the group of bent DNA. Interestingly, in both groups the data are normally distributed ($p > 0.05$ for both groups, Shapiro-Wilk test) and clearly separated ($p < 0.01$, Welch's t-test). The mean length of bent DNA is shorter than the mean length of linear DNA of about 2.6 nm (i.e., 53.5 nm versus 56.1 nm, respectively). Considering the weak immobilization of the DNA on the wAP-mica (even weaker than the case shown in Figure 3), the values in both groups are shorter but comparable to the length from NTS-3' to PAM of our fully stretched dsDNA. Overall, these real-space and direct observations by HS-AFM suggest that the ternary complex was formed at the expected target position, but the engagement may have been occurred first at the target region proximal to the PAM. Successive binding of the PAM

coincided with DNA bending. The local bending in the target DNA observed at the end of the imaging is consistent with the crystal structure of the stable ternary complex (Figure S6). Based on this consideration, we utilized the mean length of the bent DNA to estimate and indicate the position of the PAM by green arrowheads in each frame of Figures 4a-b.

In Figure S7, we report a third direct evidence of the potential long-range and selective interaction of SaCas9-sgRNA with the target DNA, which was obtained using a different experimental procedure. First, we added only dsDNA to a wAP-mica surface. During HS-AFM observation of one selected dsDNA, we injected SaCas9-sgRNA in the imaging buffer in order to capture the eventual formation of the ternary complex in case of interaction at the target site. The selected sequential HS-AFM images in Figure S7b show an SaCas9-sgRNA that landed on the wAP-mica, hopped toward the dsDNA substrate and induced sudden attraction of the dsDNA exactly in close proximity of the estimated position of the PAM. Unfortunately, the perturbation of the HS-AFM-tip led to the collapse of the SaCas9-sgRNA on the wAP-mica surface after initial binding to dsDNA. In this HS-AFM imaging, the dsDNA is visualized in its entire length for many frames prior to engagement with SaCas9-sgRNA. Accordingly, we estimated the position of the PAM in each frame by considering 32% of the experimental full length of the dsDNA. If the NTS 3' end of the dsDNA was the one on the bottom of the images, the PAM should be located at the positions indicated by the green arrowheads in Figure S7b. Indeed, the dsDNA moved suddenly toward the SaCas9-sgRNA when the latter was at the closest distance from the PAM (i.e., 7 nm, frame at 34.4 s in Figure S7b). Cross-correlation analysis of the binary images containing only the isolated dsDNA confirmed the sudden change of morphology of the dsDNA starting from this frame (Figure S7c). The use of wAP-mica was clearly instrumental in visualizing the dynamic of binding using HS-AFM. However, the resolution and quality of the HS-AFM images are clearly affected by the presence of highly mobile molecules on the imaging buffer and the wAP-mica surface. Nonetheless, the enlargement of the frame at 36.8 s in Figure S7d demonstrates that, after the first translocation occurred at 34.4 s, only a short segment of the dsDNA was attracted and engulfed by the endonuclease. The measurements of the DNA full length before interaction with SaCas9-sgRNA are normally distributed, as shown in Figure S7e. The mean full length is shorter than the ideal length (i.e., 183.5 and 204 nm, respectively). Accordingly, the estimated mean length of the DNA segment from the NTS 3' end to the PAM is 58.6 nm, which is comparable to the experimental measurements of linear and bent DNA reported in Figure 4. Collectively, these results confirm that the weak immobilization on wAP-mica systematically but consistently affect the apparent length of the DNA measured in our HS-AFM images. Interestingly, in the HS-AFM image of Figure S7d, the estimated position of the PAM is located outside of SaCas9-sgRNA, in agreement with the observations reported in Figure 4.

Discussion

HS-AFM possesses the nanometer resolution suitable for visualizing single molecules and sufficient temporal resolution for meaningful analyses of dynamics in biological processes without molecular labelling[30-32, 34, 36, 39-40, 43] (i.e., scanning speed up to 20 milliseconds per frame[44]). In previous investigations, HS-AFM enabled the real-time visualization of the DNA cleavage mediated by SpCas9[36]. In those observations, the HNH nuclease domain of SpCas9 fluctuated upon DNA binding and after Mg^{2+} addition adopted the catalytically-active docked conformation to cleave the target DNA strand. The HNH was visualized as fluctuating between two states characterized by different heights, with the low-height state followed by release of cleaved DNA. These two states are in agreement with the I and D conformations suggested by bulk and smFRET studies[45-46]. The height differences of the HNH domain in the two states (0.8 ± 0.2 nm, n = 14, ref[36]) are likely to reflect the HNH displacement toward the target DNA for the cleavage reaction. In the present study, SaCas9 in the inactive form did not show fluctuations of the HNH domain, which, conversely, stably interacts with the RuvC domain, as in the crystal structure. Upon addition of Mg^{2+}, SaCas9 cleaved the DNA substrate and simultaneously dissociated from it. The rapid dynamics of cleavage and molecular dissociation prevented HS-AFM from clearly visualizing the translocation of the HNH domain toward the cleavage site. The results shown in Figures 2 and S4 are two examples of imaging in which, immediately after the release of the cleaved DNA and before complete dissociation, the SaCas9-sgRNA showed a distinct new topographic feature located above the channel between NUC and REC lobes, namely the expected cleavage site. The maximum height of the protrusion above the cleavage site in Figure 2d is 0.82 nm lower than the maximum height of HNH before DNA release, which is consistent with the height difference in HNH between the I and D states previously measured for SpCas9[36]. In addition, in every case of DNA cleavage captured by HS-AFM (total of 32 cases), we observed conformational change of the SaCas9-sgRNA only during DNA release and dissociation. Collectively, these findings support the conclusion of previous studies that, unlike SpCas9, SaCas9 operates as a multiple turnover endonuclease[47-48].

The target DNA search by SpCas9 was also previously studied by HS-AFM[36]. Those experiments were performed after absorption of dsDNA on a mica-supported lipid bilayer and addition of the preassembled binary complex to the buffer during imaging. In those experimental conditions, SpCas9-sgRNA could freely diffuse on the lipid surface and search the target DNA during imaging. However, due to the fast diffusion of the molecules, the temporal resolution of HS-AFM was barely sufficient to confirm the occurrence of binding events at the target site without discerning the actions performed by SpCas9-sgRNA to engage and selectively recognize the target DNA. On the basis of those observations, it was concluded that SpCas9-sgRNA searches the target sites via three-dimensional diffusion, confirming previous studies conducted by DNA-curtain assays and single-particle tracking analysis[7]. Conversely, it was not observed one-dimensional sliding on DNA competing with

three-dimensional diffusion, as previously suggested by smFRET assays[6]. In the present study, we successfully optimized an AFM substrate that reduced the fast diffusion of the binary complex otherwise observed on a lipid surface. In the newly designed experimental conditions, the interaction of the AFM substrate with dsDNA and SaCas9-sgRNA did not prevent the formation of the ternary complex but instead slowed the dynamics of binding within intervals of time accessible to HS-AFM. The unique experimental evidences presented here revealed a potential long-range selective interaction that facilitates binding of SaCas9-sgRNA to the target DNA.

The HS-AFM experiments of apo-SaCas9 and SaCas9-sgRNA still preserves sufficient flexibility for fluctuations from a more stable closed structure to an open conformation, in which the two lobes transiently separated of several nanometers. The ternary complex formation in Movie S7 and Figure 4a revealed an even more dramatic conformational change of the bilobed structure that was triggered by the long-range attractive interaction with the target DNA. The highly open conformation of the binary complex was probably overstretched by the immobilization of one lobe that did not follow the second one interacting to the target DNA. The AFM-substrate hampered also the rapid closing of the SaCas9-sgRNA after catching the dsDNA and following molecular rearrangement of the SaCas9 structure after engagement, which enabled us to unveil the process of binding and the sequential transition from linear to bent DNA. Recent cryo-EM studies revealed the SpCas9-sgRNA-dsDNA complexes trapped at different states in the interrogation pathway[24-25]. In some of these experiments, linear and bent DNA configurations were observed after covalent cross-linking of the protein and the DNA on the side of the PAM distal to the site of R-loop initiation[24]. The bent DNA is associated to a closed conformation of the protein, similar to the previous crystal structure of SpCas9-sgRNA, while linear DNA is associated to a new open conformation in which the REC lobe is rotated away from the NUC lobe. Bending of DNA induces base flipping that initiate the R-loop formation with the complementary spacer. In our HS-AFM experiment reported in Figure 4, we visualized a sequential transition from linear to bent DNA. Although the long-range interaction was selectively triggered by proximity to the PAM, the difference in length of the NTS-3' DNA segment before and after bending suggest that the initial binding did not occur at the PAM. Indeed, this conclusion is also supported by the topographic images of open SaCas9-sgRNA over the target DNA in Figure 4a (i.e., frames from 71.5 s to 75.5 s) and Figure S7d. Conversely, the successive PAM engagement clearly occurred after bending of the DNA. Accordingly, we hypothesize and propose the model reported in Figure 5 to describe the ternary complex formation. The binary complex exists in close and open bilobed configurations, with the former one more energetically favorable (step I). When the diffusing SaCas9-sgRNA approaches the target site, presumably the PI domain interacts with the PAM, which increases the probability of transition of SaCas9-sgRNA to the open configuration transiently observed in HS-AFM

experiments of SaCas9-sgRNA. The nature of this long-range selective interaction is still elusive. Next, the increased probability of the open structure in the vicinity of the PAM may favor the exposure of clusters of hydrophobic residues toward the neighboring target DNA, which triggers the long-range attractive forces (step II) leading to the engagement of linear dsDNA with SaCas9 (step III). We hypothesize that the attractive forces observed in our experiments may have hydrophobic nature based on the strength and the long range of interaction. Interpretation of hydrophobic forces and quantification of their magnitude and range in aqueous solutions is still under debate. In fact, experimental results are qualitatively different as the size of the hydrophobic species varies from a small single molecule to extended surfaces. Nonetheless, interactions between hydrophobic patches of biomolecules (i.e., surface clusters of amino acid residues) cannot be ruled out within a range of 10-20 nm[49-52]. In addition, recent spectroscopy data and optical tweezers experiments suggest that hydrophobic interactions play a crucial role in the dynamics and reactivity of DNA in water[53]. According to our proposed model, after engagement, conformational changes of the NUC and REC lobes lead to PI-PAM binding (step IV, the crystal structure with linear DNA in ref. [24]), followed by DNA bending for R-loop formation (step V, the crystal structure with bent DNA in ref. [24]). The latest two steps occurred presumably during the acquisition of the image at 114.5 s in Figure 4b, namely they were too fast to be discriminated by HS-AFM. Noteworthy is the fact that previous experiments on the ortholog SpCas9 suggest that at least two possible binding intermediates exist in the process of cognate target recognition[7]. Moreover, the off-target binding events of SpCas9 occur specifically at PAM loci and show insensitivity to ionic strength[7]. We hypothesize that the two binding intermediates showed in step III and IV of Figure 5 may be due to hydrophobic interaction between SaCas9-sgRNA and dsDNA and electrostatic interactions between the PAM and the PI domain, respectively. These intermediates precede R-loop formation in case of binding to the target DNA or rapid dissociation in case of off-target binding. As a matter of fact, hydrophobic forces are promoted by increasing salt concentration, while, conversely, electrostatic interactions tend to decrease with increasing ionic strength. Two binding intermediates governed by two type of forces with opposite sensitivity to the change of ionic strength may explain the double-exponential decay of the survival probability of off-target binding previously observed in Sp-Cas9[7], which was not appreciably affected by the variation of salt concentration. The proximity to PAM that maximizes the exposure of the hydrophobic residues may explain the selectivity inferred by the higher number of off-target binding events observed in PAM-rich regions of DNA substrates[7].

Conclusion

In summary, according to our HS-AFM observations, SaCas9 is a highly flexible modular structure in the apo-state, which is stabilized in a bilobed close structure upon binding with sgRNA. The binary complex still preserves some flexibility and transiently exists in conformations with two open lobes. We obtained independent experimental evidences that suggest the existence of a potential long-range selective interaction between SaCas9-sgRNA and its target DNA. The forces are attractive and selectively triggered by the proximity to the PAM, up to several nanometers. Remarkably, sequential topographic images of the molecular engagement suggest that the first binding event occurred at the target sequence. Conversely, the target DNA bending necessary for R-loop formation and activation of the endonuclease involved the binding of the PAM, which was visualized in a second phase of the process. Although the present HS-AFM observations represent a direct evidence of the long-range selective interaction between SaCas9-sgRNA and its target DNA, further investigations are necessary to confirm the mechanism of search for target DNA hypothesized and illustrated in the present study.

Methods

Sample preparation

Staphylococcus aureus Cas9 and the guide RNA were prepared as previously described[21], with minor modifications. Briefly, the SaCas9 protein was expressed in *Escherichia coli* Rosetta2 (DE3) (Novagen), and then purified by chromatography on Ni-NTA Superflow (QIAGEN) and HiTrap SP HP (GE Healthcare). The 98-nt guide RNA was transcribed *in vitro* using T7 RNA polymerase, and then purified by denaturing urea polyacrylamide gel electrophoresis. The 600-bp target DNA was PCR-amplified using the pUC19 plasmid containing the 21-nt target sequence and the CAGAAT PAM as the template, and then purified using a Wizard DNA Clean-Up System (Promega). The sequences of SaCas9, the guide RNA and the target DNA are provided in Supporting Information Data 1.

HS-AFM apparatus

HS-AFM experiments were performed using a laboratory-made high-speed atomic force microscope operating in tapping mode. The deflection of the cantilever was monitored by a laser at 670 nm with 0.7 mW. The laser was focused onto the back side of the cantilever through a 20× objective lens (CFI S Plan Fluor ELWD 20X, Nikon, Japan). The reflected laser was collected and detected with a two-segmented PIN photodiode (MPR-1, Graviton, Japan). We use a small cantilever with spring constant, resonant frequency and quality

factor of ~100 pN/nm, ~400 kHz and ~2, respectively (BL-AC10DS-A2, Olympus, Japan). The free oscillation amplitude of the cantilever was under 1 nm, and the setpoint amplitude was set to 80% of the free amplitude during HS-AFM observations. For high-resolution single-molecule imaging, an amorphous carbon tip was grown on the apex of the original AFM tip by electron beam deposition (EBD) using a scanning electron microscope (SEM) (Supra 40VP, Zeiss, Oberkochen, Germany). The length of the additional sharp AFM tip was ~500 nm, and the apex of the tip was ~1 nm in radius after further plasma etching by a plasma cleaner (Tergeo, PIC Scientific, USA).

HS-AFM observations

For HS-AFM observations of apo-SaCas9, SaCas9-sgRNA and SaCas9-sgRNA-dsDNA on AP-mica, a mica surface was treated for 3 min with 0.01% (3-aminopropyl)triethoxysilane (APTES) (Shin-Etsu Chemical, Japan) in MilliQ. For HS-AFM observations of SaCas9-sgRNA search and binding to target DNA on wAP-mica, a mica surface was treated for 3 min with 0.0025% APTES in Milli-Q water. The samples were added to the AP-mica or wAP-mica surface upon 3 min incubation of 2 µL at concentration of 10 nM. The binary and ternary complexes were pre-assembled (SaCas9:sgRNA:dsDNA = 1:1:1 mole ratio) in AFM-imaging buffer. HS-AFM observations of apo-SaCas9 and SaCas9-sgRNA were performed in buffer consisting of 20 mM Tris-HCl, pH 7.5, 100 mM KCl, 10% glycerol and 0.1 mM EDTA. HS-AFM observations of the ternary complex were performed in buffer consisting of 20 mM Tris-HCl, pH 7.5, 30 mM KCl, 10% glycerol and 0.1 mM EDTA. All HS-AFM experiments were performed at room temperature.

HS-AFM image processing

Image flattening was achieved by means of plane or second order polynomial surface fitting, as appropriate. Afterwards, median (0 order) line-by-line leveling was used to remove height offsets along the fast scan axis. These steps were performed using in-house software routines developed in MATLAB (MathWorks)[54]. Subsequently, HS-AFM images were processed using *Fiji* software (NIH, USA)[55]. A 0.7-pixel-radius mean filter was applied to each HS-AFM image to reduce noise.

Simulation of AFM images

The BioAFMviewer software was used for validation of measured SaCas9-sgRNA-dsDNA topographies. Simulated scanning was based on the non-elastic collisions of a rigid cone-shaped tip model with the rigid Van-der-Waals atomic model of the protein structure[42]. As a structural template PDB:5CZZ was used. Automatized fitting[41] was employed to obtain the simulated image best matching the HS-AFM target image (image correlation coefficients reported in each figure). In the simulation shown in Figure 2c, the tip shape parameters

were R=1.5 nm for the tip probe sphere radius and α=23° for the cone half-angle. In the simulation shown in Figure S4, the tip shape parameters were R=1.1 nm for the tip probe sphere radius and α=5° for the cone half-angle.

Data and statistical analyses

The TrackMate[56] plugin for *Fiji* was utilized to track the positions and the maximum height of the HNH domain in the ternary complex after Mg^{2+} addition to the imaging buffer. A circular area of diameter set to 5 nm was used for the tracking analysis. Cross-correlation analyses of HS-AFM sequential images were performed using the plugin coloc2 included in *Fiji*. For each HS-AFM experiment, we correlated each sequential image with the initial frame by means of the Pearson correlation coefficient. In the case of cross-correlation analysis of dsDNA, we compared the sequential binary images of the dsDNA after isolation of the dsDNA topography and thresholding. The length of dsDNA was measured from topographic images using segmented lines in *Fiji*. Normality of the data distributions for linear and bent DNA was evaluated by Shapiro-Wilk test (level of significance, α= 0.05), while the mean lengths of linear and bent DNA segments were compared by Welch's t-test (level of significance, α= 0.01). The statistical tests were performed using a commercially available software package (Origin 9.1, OriginLab Co., MA, USA).

Acknowledgment

We thank Prof. Toshio Ando at Kanazawa University for providing the HS-AFM apparatus, Prof. Takayuki Uchihashi at Nagoya University for providing HS-AFM software, Yoko Mikami and Yui Okada at Kanazawa University for laboratory management. This work was supported by the World Premier International Research Center Initiative (WPI), MEXT, Japan and JSPS KAKENHI (grant numbers 18H01836, 21H01771 and 21H00247 (M.S.)).

Author contributions

Leonardo Puppulin[1,†,*], Junichiro Ishikawa[2], Ayumi Sumino[1,3], Arin Marchesi[4], Holger Flechsig[1], Kenichi Umeda[1], Noriyuki Kodera[1], Hiroshi Nishimasu[2,5,6,7], and Mikihiro Shibata[1,]

L.P and M.S. conceived and designed the research. L.P. and M.S. performed HS-AFM experiments. J.I. and H.N. prepared materials (DNA, RNA and SaCas9). L.P., A.S., A.M., H.F. analyzed the data. K.U. and N.K. improved HS-AFM. L.P., A.S., A.M., H.F., K.U., N.K., H.N, and M.S. edited the manuscript. L.P. and M.S. wrote the manuscript.

References:

1. Barrangou, R.; Fremaux, C.; Deveau, H.; Richards, M.; Boyaval, P.; Moineau, S.; Romero, D. A.; Horvath, P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **2007,** *315* (5819), 1709-1712.
2. Brouns, S. J.; Jore, M. M.; Lundgren, M.; Westra, E. R.; Slijkhuis, R. J.; Snijders, A. P.; Dickman, M. J.; Makarova, K. S.; Koonin, E. V.; Van Der Oost, J. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **2008,** *321* (5891), 960-964.
3. Garneau, J. E.; Dupuis, M.-È.; Villion, M.; Romero, D. A.; Barrangou, R.; Boyaval, P.; Fremaux, C.; Horvath, P.; Magadán, A. H.; Moineau, S. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **2010,** *468* (7320), 67-71.
4. Wiedenheft, B.; Sternberg, S. H.; Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **2012,** *482* (7385), 331-338.
5. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J. A.; Charpentier, E. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science* **2012,** *337* (6096), 816-821.
6. Globyte, V.; Lee, S. H.; Bae, T.; Kim, J. S.; Joo, C. CRISPR/Cas9 searches for a protospacer adjacent motif by lateral diffusion. *EMBO J.* **2019,** *38* (4), e99466.
7. Sternberg, S. H.; Redding, S.; Jinek, M.; Greene, E. C.; Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **2014,** *507* (7490), 62-67.
8. Qi, L. S.; Larson, M. H.; Gilbert, L. A.; Doudna, J. A.; Weissman, J. S.; Arkin, A. P.; Lim, W. A. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **2013,** *152* (5), 1173-1183.
9. Larson, M. H.; Gilbert, L. A.; Wang, X.; Lim, W. A.; Weissman, J. S.; Qi, L. S. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat. Protoc.* **2013,** *8* (11), 2180-2196.
10. Sander, J. D.; Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* **2014,** *32* (4), 347-355.
11. Wang, H.; La Russa, M.; Qi, L. S. CRISPR/Cas9 in genome editing and beyond. *Annu. Rev. Biochem.* **2016,** *85* (1), 227-264.
12. Gillmore, J. D.; Gane, E.; Taubel, J.; Kao, J.; Fontana, M.; Maitland, M. L.; Seitzer, J.; O'Connell, D.; Walsh, K. R.; Wood, K. CRISPR-Cas9 in vivo gene editing for transthyretin amyloidosis. *N. Engl. J. Med.* **2021,** *385* (6), 493-502.
13. Duan, J.; Lu, G.; Xie, Z.; Lou, M.; Luo, J.; Guo, L.; Zhang, Y. Genome-wide identification of CRISPR/Cas9 off-targets in human genome. *Cell Res.* **2014,** *24* (8), 1009-1012.
14. Wang, X.; Wang, Y.; Wu, X.; Wang, J.; Wang, Y.; Qiu, Z.; Chang, T.; Huang, H.; Lin, R.-J.; Yee, J.-K. Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat. Biotechnol.* **2015,** *33* (2), 175-178.
15. Wu, X.; Scott, D. A.; Kriz, A. J.; Chiu, A. C.; Hsu, P. D.; Dadon, D. B.; Cheng, A. W.; Trevino, A. E.; Konermann, S.; Chen, S. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.* **2014,** *32* (7), 670-676.
16. Singh, D.; Sternberg, S. H.; Fei, J.; Doudna, J. A.; Ha, T. Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9. *Nat. Commun.* **2016,** *7* (1), 1-8.
17. Jinek, M.; Jiang, F.; Taylor, D. W.; Sternberg, S. H.; Kaya, E.; Ma, E.; Anders, C.; Hauer, M.; Zhou, K.; Lin, S. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* **2014,** *343* (6176), 1247997.

18. Anders, C.; Niewoehner, O.; Duerst, A.; Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **2014,** *513* (7519), 569-573.
19. Nishimasu, H.; Ran, F. A.; Hsu, P. D.; Konermann, S.; Shehata, S. I.; Dohmae, N.; Ishitani, R.; Zhang, F.; Nureki, O. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **2014,** *156* (5), 935-949.
20. Jiang, F.; Zhou, K.; Ma, L.; Gressel, S.; Doudna, J. A. A Cas9–guide RNA complex preorganized for target DNA recognition. *Science* **2015,** *348* (6242), 1477-1481.
21. Nishimasu, H.; Cong, L.; Yan, W. X.; Ran, F. A.; Zetsche, B.; Li, Y.; Kurabayashi, A.; Ishitani, R.; Zhang, F.; Nureki, O. Crystal structure of Staphylococcus aureus Cas9. *Cell* **2015,** *162* (5), 1113-1126.
22. Jiang, F.; Taylor, D. W.; Chen, J. S.; Kornfeld, J. E.; Zhou, K.; Thompson, A. J.; Nogales, E.; Doudna, J. A. Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science* **2016,** *351* (6275), 867-871.
23. Jiang, F.; Doudna, J. A. CRISPR-Cas9 structures and mechanisms. *Annu. Rev. Biophys.* **2017,** *46* (1), 505-529.
24. Cofsky, J. C.; Soczek, K. M.; Knott, G. J.; Nogales, E.; Doudna, J. A. CRISPR–Cas9 bends and twists DNA to read its sequence. *Nat. Struct. Mol. Biol.* **2022,** *29* (4), 395-402.
25. Pacesa, M.; Loeff, L.; Querques, I.; Muckenfuss, L. M.; Sawicka, M.; Jinek, M. R-loop formation and conformational activation mechanisms of Cas9. *Nature* **2022**, 1-6.
26. Osuka, S.; Isomura, K.; Kajimoto, S.; Komori, T.; Nishimasu, H.; Shima, T.; Nureki, O.; Uemura, S. Real-time observation of flexible domain movements in CRISPR–Cas9. *EMBO J.* **2018,** *37* (10), e96941.
27. Palermo, G.; Miao, Y.; Walker, R. C.; Jinek, M.; McCammon, J. A. Striking plasticity of CRISPR-Cas9 and key role of non-target DNA, as revealed by molecular simulations. *ACS Cent. Sci.* **2016,** *2* (10), 756-763.
28. Zuo, Z.; Liu, J. Cas9-catalyzed DNA cleavage generates staggered ends: evidence from molecular dynamics simulations. *Sci. Rep.* **2016,** *6* (1), 1-9.
29. Zheng, W. Probing the structural dynamics of the CRISPR-Cas9 RNA-guided DNA-cleavage system by coarse-grained modeling. *Proteins: Struct. Funct. Bioinfor.* **2017,** *85* (2), 342-353.
30. Ando, T.; Kodera, N.; Takai, E.; Maruyama, D.; Saito, K.; Toda, A. A high-speed atomic force microscope for studying biological macromolecules. *Proc. Natl. Acad. Sci. U. S. A.* **2001,** *98* (22), 12468-12472.
31. Shibata, M.; Yamashita, H.; Uchihashi, T.; Kandori, H.; Ando, T. High-speed atomic force microscopy shows dynamic molecular processes in photoactivated bacteriorhodopsin. *Nat. Nanotechnol.* **2010,** *5* (3), 208-212.
32. Kodera, N.; Yamamoto, D.; Ishikawa, R.; Ando, T. Video imaging of walking myosin V by high-speed atomic force microscopy. *Nature* **2010,** *468* (7320), 72-76.
33. Uchihashi, T.; Iino, R.; Ando, T.; Noji, H. High-speed atomic force microscopy reveals rotary catalysis of rotorless F1-ATPase. *Science* **2011,** *333* (6043), 755-758.
34. Ando, T.; Uchihashi, T.; Scheuring, S. Filming biomolecular processes by high-speed atomic force microscopy. *Chem. Rev.* **2014,** *114* (6), 3120-3188.
35. Preiner, J.; Horner, A.; Karner, A.; Ollinger, N.; Siligan, C.; Pohl, P.; Hinterdorfer, P. High-speed AFM images of thermal motion provide stiffness map of interfacial membrane protein moieties. *Nano Lett.* **2015,** *15* (1), 759-763.

36. Shibata, M.; Nishimasu, H.; Kodera, N.; Hirano, S.; Ando, T.; Uchihashi, T.; Nureki, O. Real-space and real-time dynamics of CRISPR-Cas9 visualized by high-speed atomic force microscopy. *Nat. Commun.* **2017,** *8* (1), 1-9.
37. Lyubchenko, Y. L. Direct AFM visualization of the nanoscale dynamics of biomolecular complexes. *J. Phys. D Appl. Phys.* **2018,** *51* (40), 403001.
38. Nievergelt, A. P.; Banterle, N.; Andany, S. H.; Gönczy, P.; Fantner, G. E. High-speed photothermal off-resonance atomic force microscopy reveals assembly routes of centriolar scaffold protein SAS-6. *Nat. Nanotechnol.* **2018,** *13* (8), 696-701.
39. Marchesi, A.; Gao, X.; Adaixo, R.; Rheinberger, J.; Stahlberg, H.; Nimigean, C.; Scheuring, S. An iris diaphragm mechanism to gate a cyclic nucleotide-gated ion channel. *Nat. Commun.* **2018,** *9* (1), 1-11.
40. Puppulin, L.; Kanayama, D.; Terasaka, N.; Sakai, K.; Kodera, N.; Umeda, K.; Sumino, A.; Marchesi, A.; Weilin, W.; Tanaka, H. Macrocyclic Peptide-Conjugated Tip for Fast and Selective Molecular Recognition Imaging by High-Speed Atomic Force Microscopy. *ACS Appl. Mater. Interfaces* **2021,** *13* (46), 54817-54829.
41. Amyot, R.; Marchesi, A.; Franz, C. M.; Casuso, I.; Flechsig, H. Simulation atomic force microscopy for atomic reconstruction of biomolecular structures from resolution-limited experimental images. *PLoS Comput. Biol.* **2022,** *18* (3), e1009970.
42. Amyot, R.; Flechsig, H. BioAFMviewer: An interactive interface for simulated AFM scanning of biomolecular structures and dynamics. *PLoS Comput. Biol.* **2020,** *16* (11), e1008444.
43. Fukuda, S.; Ando, T. Faster high-speed atomic force microscopy for imaging of biomolecular processes. *Rev. Sci. Instrum.* **2021,** *92* (3), 033705.
44. Shimizu, M.; Okamoto, C.; Umeda, K.; Watanabe, S.; Ando, T.; Kodera, N. An ultrafast piezoelectric Z-scanner with a resonance frequency above 1.1 MHz for high-speed atomic force microscopy. *Rev. Sci. Instrum.* **2022,** *93* (1), 013701.
45. Sternberg, S. H.; LaFrance, B.; Kaplan, M.; Doudna, J. A. Conformational control of DNA target cleavage by CRISPR–Cas9. *Nature* **2015,** *527* (7576), 110-113.
46. Dagdas, Y. S.; Chen, J. S.; Sternberg, S. H.; Doudna, J. A.; Yildiz, A. A conformational checkpoint between DNA binding and cleavage by CRISPR-Cas9. *Sci. Adv.* **2017,** *3* (8), eaao0027.
47. Yourik, P.; Fuchs, R. T.; Mabuchi, M.; Curcuru, J. L.; Robb, G. B. Staphylococcus aureus Cas9 is a multiple-turnover enzyme. *RNA* **2019,** *25* (1), 35-44.
48. Zhang, S.; Zhang, Q.; Hou, X. M.; Guo, L.; Wang, F.; Bi, L.; Zhang, X.; Li, H. H.; Wen, F.; Xi, X. G. Dynamics of Staphylococcus aureus Cas9 in DNA target Association and Dissociation. *EMBO Rep.* **2020,** *21* (10), e50184.
49. Israelachvili, J.; Pashley, R. The hydrophobic interaction is long range, decaying exponentially with distance. *Nature* **1982,** *300* (5890), 341-342.
50. Zhang, X.; Zhu, Y.; Granick, S. Softened hydrophobic attraction between macroscopic surfaces in relative motion. *J. Am. Chem. Soc.* **2001,** *123* (27), 6736-6737.
51. Meyer, E. E.; Rosenberg, K. J.; Israelachvili, J. Recent progress in understanding hydrophobic interactions. *Proc. Natl. Acad. Sci. U. S. A.* **2006,** *103* (43), 15739-15746.
52. Donaldson Jr, S. H.; Røyne, A.; Kristiansen, K.; Rapp, M. V.; Das, S.; Gebbie, M. A.; Lee, D. W.; Stock, P.; Valtiner, M.; Israelachvili, J. Developing a general interaction potential for hydrophobic and hydrophilic interactions. *Langmuir* **2015,** *31* (7), 2051-2064.
53. Feng, B.; Sosa, R. P.; Mårtensson, A. K.; Jiang, K.; Tong, A.; Dorfman, K. D.; Takahashi, M.; Lincoln, P.; Bustamante, C. J.; Westerlund, F. Hydrophobic catalysis and a potential biological role of DNA

unstacking induced by environment effects. *Proc. Natl. Acad. Sci. U. S. A.* **2019,** *116* (35), 17169-17174.
54. Zuttion, F.; Redondo-Morata, L.; Marchesi, A.; Casuso, I. High-resolution and high-speed atomic force microscope imaging. In *Nanoscale Imaging*, Springer: 2018; pp 181-200.
55. Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **2012,** *9* (7), 676-682.
56. Tinevez, J.-Y.; Perry, N.; Schindelin, J.; Hoopes, G. M.; Reynolds, G. D.; Laplantine, E.; Bednarek, S. Y.; Shorte, S. L.; Eliceiri, K. W. TrackMate: An open and extensible platform for single-particle tracking. *Methods* **2017,** *115*, 80-90.

Figures and captions:

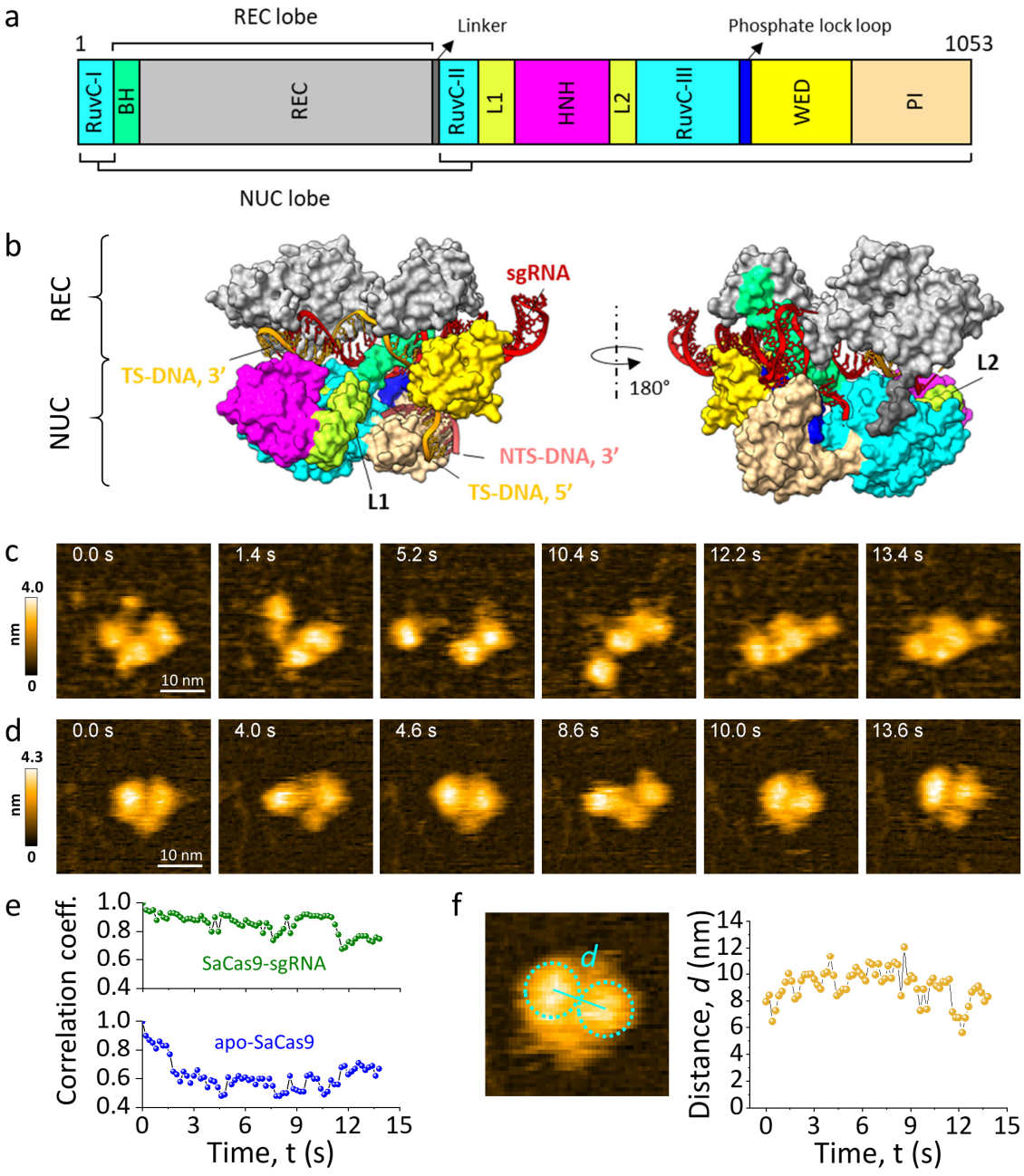

Figure 1: HS-AFM observations of apo-SaCas9 and SaCas9-sgRNA. a Domain structure of *Staphylococcus aureus* Cas9 (SaCas9). BH, Bridge helix; L1 and L2, Linker region 1 and 2. **b** Crystal structure of SaCas9 in complex with sgRNA and target DNA (PDB: 5CZZ). The domains are colored as

in **a**. **c, d** Sequential HS-AFM images of apo-SaCas9 and SaCas9-sgRNA, respectively, on the AP-mica surface. Imaging parameters in **c**: scanning area = 60 × 48 nm^2 (120 × 96 pixels); x-scan velocity = 59.9 µm/s; frame time = 200 ms. In **d:** scanning area = 60 × 60 nm^2 (120 × 120 pixels); x-scan velocity = 75.2 µm/s; frame time = 200 ms. The displayed area in **c** and **d** is 35 × 35 nm^2. **e** Time courses of correlation coefficients between the sequential HS-AFM images of apo-SaCas9 and SaCas9-sgRNA. Every image is correlated to the image at t=0. **f** Time course of the distance between the mean center positions of the two lobes of SaCas9-sgRNA observed during HS-AFM imaging. Each lobe was tracked by a circular area of diameter set to 5 nm.

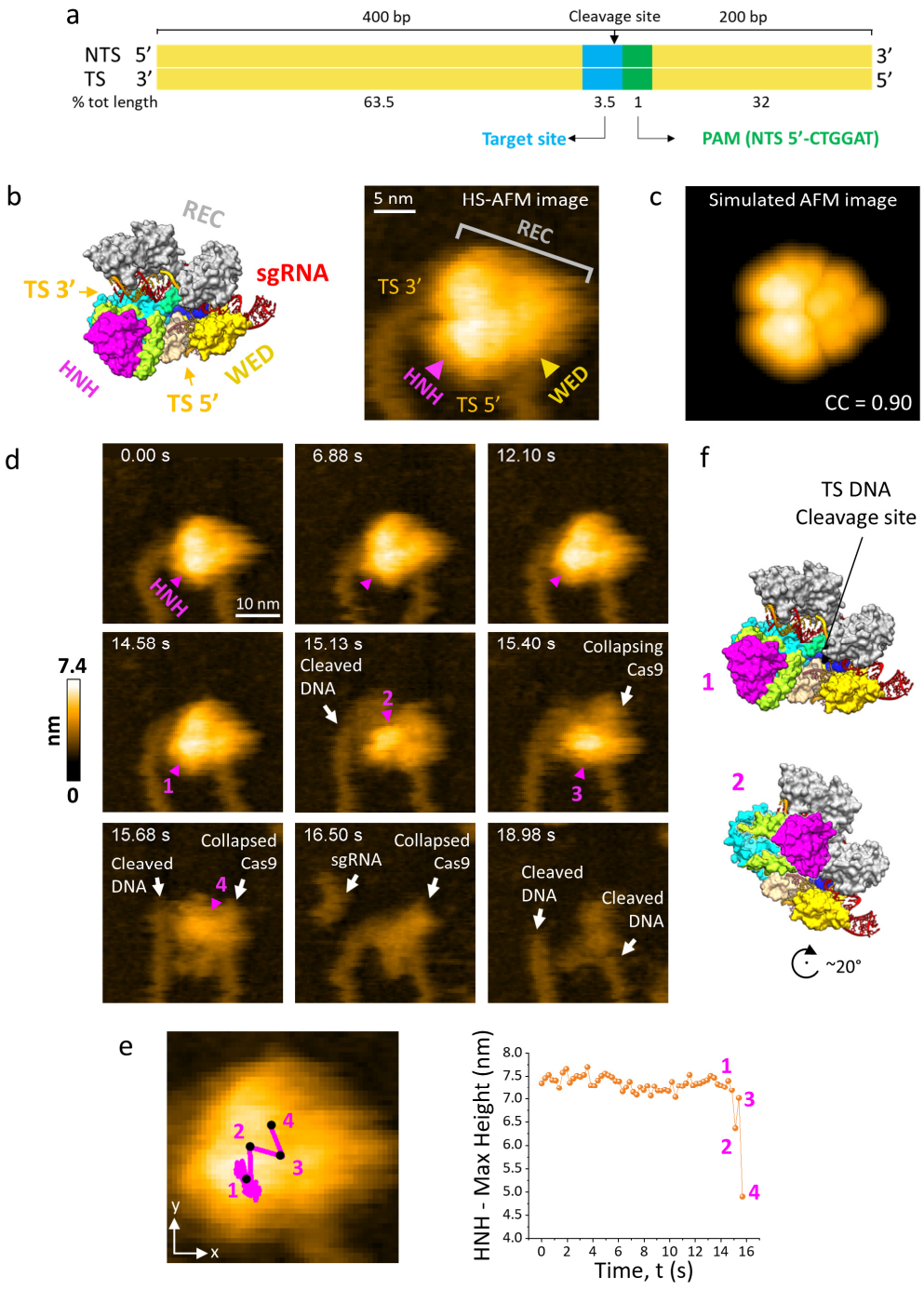

Figure 2: HS-AFM observations of DNA cleavage by SaCas9-sgRNA. a Schematic of the dsDNA substrate. The target site and the PAM are colored blue and green, respectively. The lengths of each dsDNA segment are indicated as numbers of base-pair (bp) and percentages of the total length. TS, target strand; NTS, non-target strand. **b** HS-AFM topographic image of the ternary complex and the

relative surface representation, in which are indicated the visible domains, the sgRNA and the target DNA. TS, target strand. The optimized orientation of the structure was obtained by an automatized fitting routine that matches the experimental HS-AFM image to the simulated AFM image. The best simulated AFM image and the coefficient of correlation (CC) are reported in **c**. **d** Sequential HS-AFM images of SaCas9-sgRNA-dsDNA in the presence of $MgCl_2$. The HNH domain in each frame is indicated by magenta arrowhead. Imaging parameters: scanning area = 60 × 60 nm^2 (140 × 140 pixels); x-scan velocity = 65.3 µm/s; frame time = 275 ms. The displayed area is 40 × 40 nm^2. **e** In-plane trajectory of the HNH domain and the variation of its maximum height as obtained from the tracking analysis. The points labelled from 1 to 4 correspond to the HNH numbered in the topographic images of **d**. **f** Two tentative configurations of the ternary complex in agreement to the two consecutive topographic images at 14.58 s and 15.13 s in **d**, representing the translocation of HNH to the expected cleavage site.

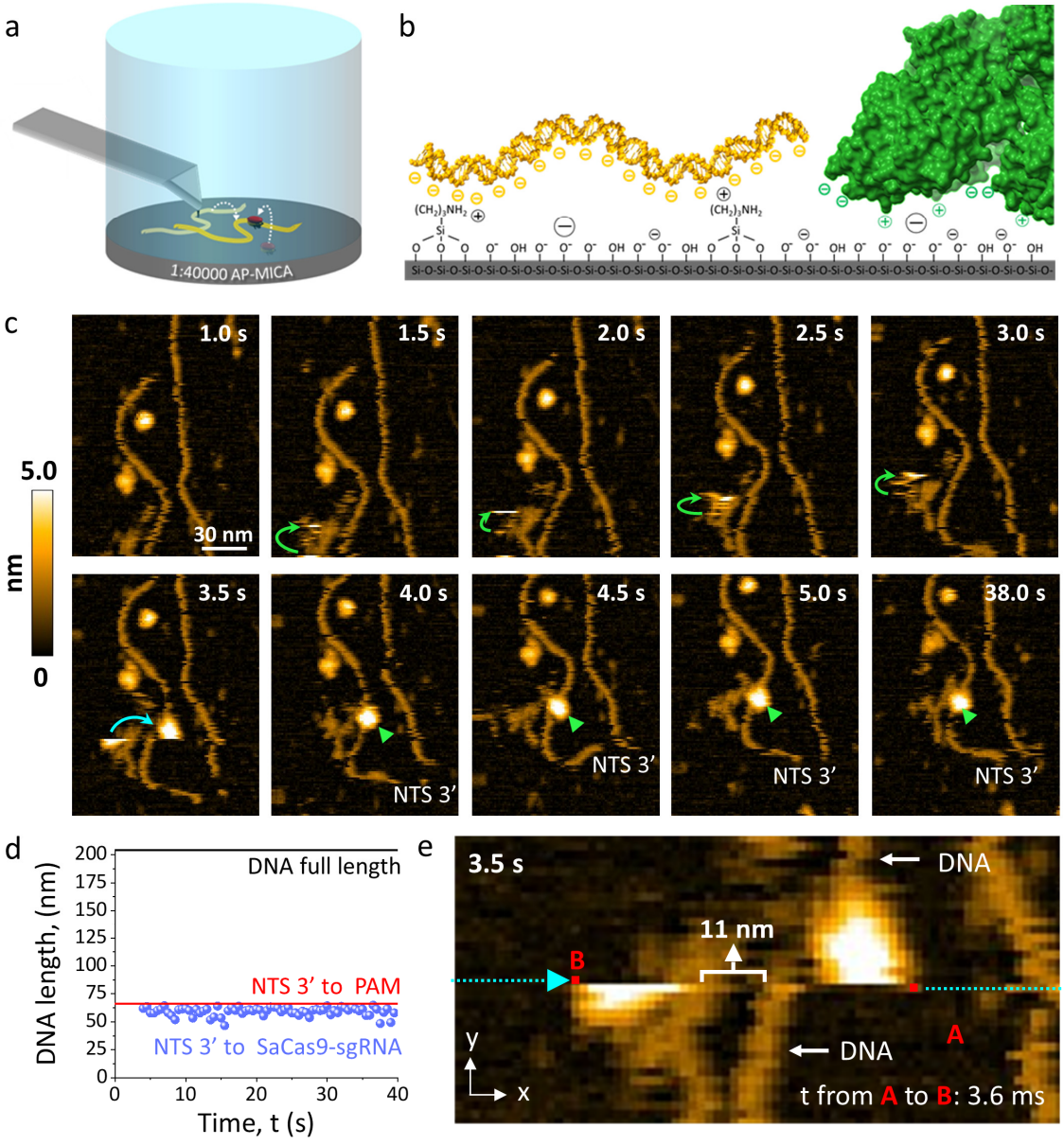

Figure 3: Direct evidence of long-range selective interaction between SaCas9-sgRNA and target DNA. a Explanatory cartoon image of the HS-AFM experimental conditions. **b** Cartoon image of a weakly functionalized mica by APTES: scattered negative and positive charges transiently attract and repulse the biomolecules, reducing their mobility without impeding their mutual interaction. **c** Selected sequential HS-AFM images of the ternary complex formation. Green curved arrows and cyan curved arrow indicate SaCas9-sgRNA hopping on the wAP-mica surface and SaCas9-sgRNA suddenly translocating to the target DNA, respectively. In the HS-AFM frame at 3.5 s, the binary complex abruptly changes its direction of motion to form the stable molecular complex indicated by green arrowheads in the following frames. Imaging parameters: scanning area = 200 × 160 nm^2

(150 × 120 pixels); x-scan velocity = 96.8 µm/s; frame time = 500 ms. The displayed area is 125 × 160 nm^2. **d** Experimental length of the shorter segment of dsDNA after ternary complex formation as a function of time. The length of a fully stretched dsDNA substrate and the ideal length representing the position of PAM are shown by black and red lines, respectively. **e** Enlarged area of the frame at 3.5 s in which the complex is formed. The displayed area is 96 × 46 nm^2. The discontinuity of the topography of SaCas9-sgRNA is due to the rapid translocation of the endonuclease to the target site of the DNA, which is faster than the distance travelled by the AFM-tip from point A to point B (i.e., faster than 3.6 ms). The cyan dotted line following point A and the cyan dotted arrow preceding point B, which is in the successive and upper pixel line, indicate the trajectory of the AFM-tip from A to B. The travelled distance includes also the AFM-tip retrace line of 200 nm (not shown).

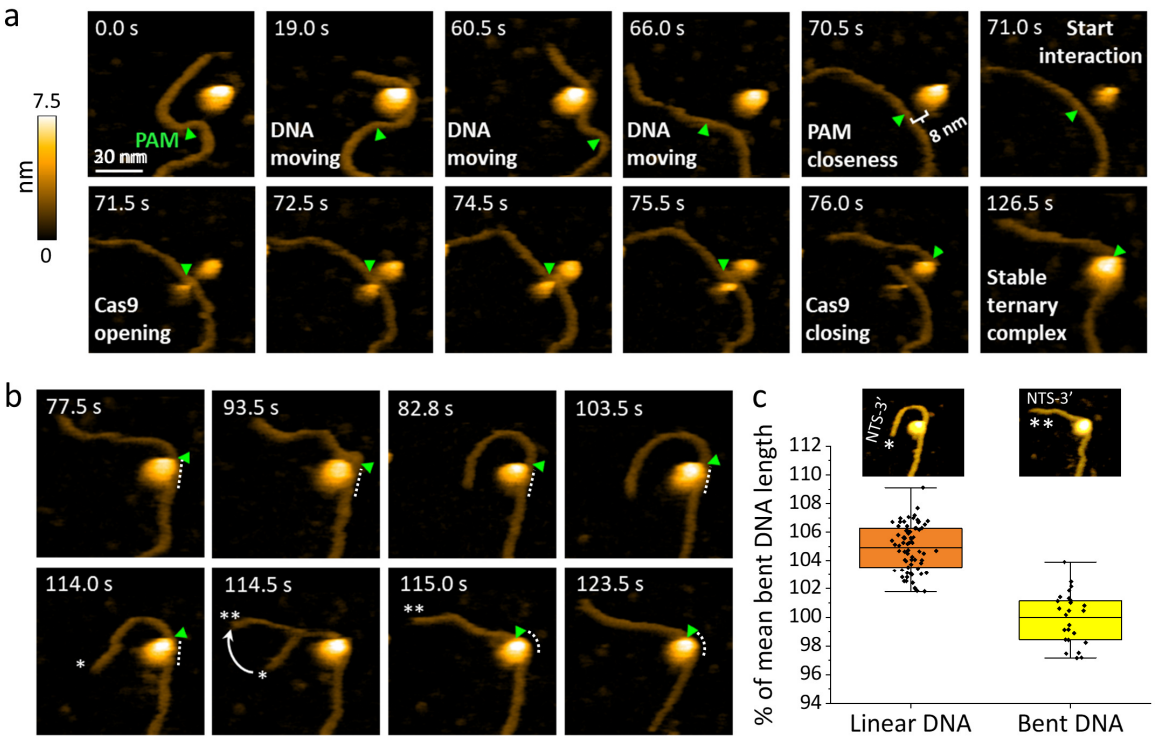

Figure 4: HS-AFM direct evidence of the luring, capturing and bending of the target DNA by SaCas9-sgRNA. a Sequential HS-AFM images of dsDNA hopping in proximity of SaCas9-sgRNA. When SaCas9-sgRNA is facing the PAM from a distance of 8 nm, it interacts with the dsDNA and forms the ternary complex. Green arrowheads indicate the estimated position of the PAM, that is the mean length of bent DNA shown in **b** and **c**. Imaging parameters: scanning area = 100 × 80 nm^2 (200 × 160 pixels); x-scan velocity = 64.8 µm/s; frame time = 500 ms. The displayed area is 80 × 80 nm^2. **b** Sequential HS-AFM images of the ternary complex after binding. Two distinct sequential configurations of the complex are observed, that is SaCas9-sgRNA bound to linear DNA (white * and linear dotted line) followed by SaCas9-sgRNA bound to bent DNA (white ** and curved dotted line). The sudden bending and reorientation of the upper segment of DNA by the endonuclease occurred in the frame at 114.5 s. Green arrowheads indicate the estimated position of the PAM. **c** Distributions of the length of the dsDNA from NTS 3' end to the endonuclease measured after binding. The results are divided into two groups, according to the two configurations previously described (74 and 25 frames for linear and bent DNA, respectively). Box plots display mean, 25-75 percentiles and outliers (whiskers). The difference of mean DNA length between the two groups is statistically significant ($p<0.01$, Welch's t-test).

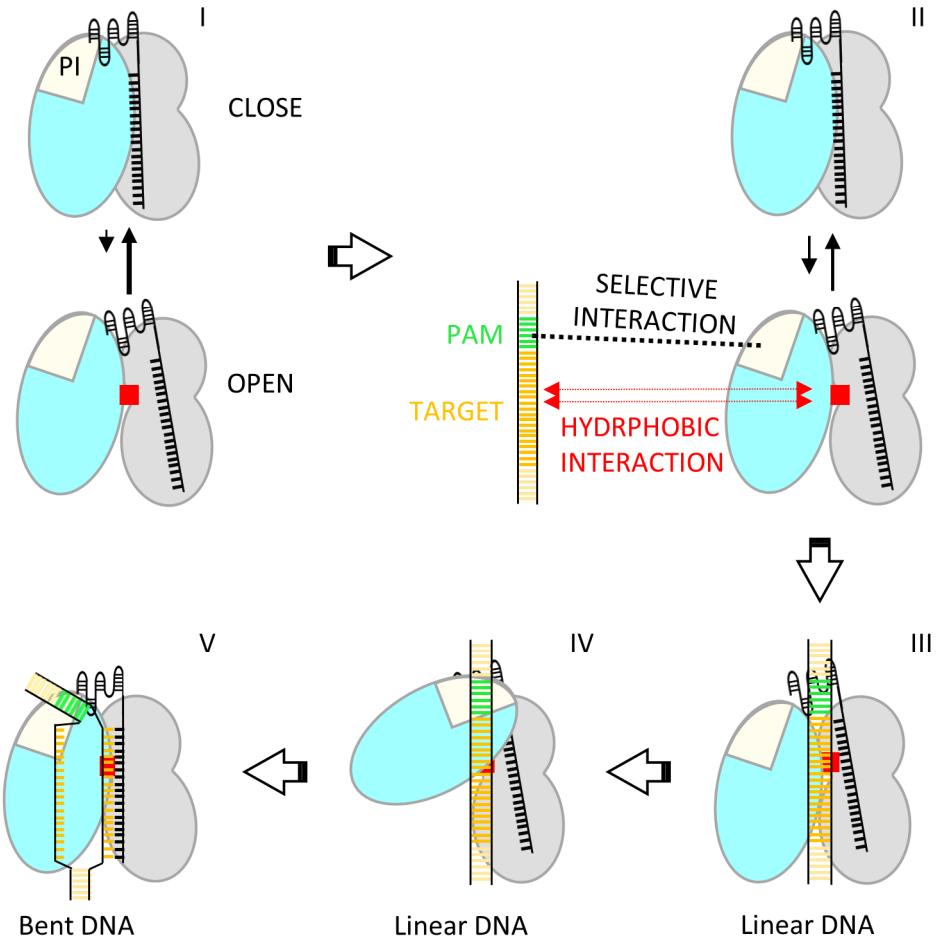

Figure 5: Model for SaCas9-sgRNA target search facilitated but selective long-range interaction. Cartoon images describing the proposed mechanism of binding and ternary complex formation between SaCas9-sgRNA and its target dsDNA. Steps I to III depict states inferred by the results of the present study. Red square represents a hydrophobic patch exposed to water in the open configuration of SaCas9-sgRNA. Steps IV and V are based on ref. [24]. In steps I and II the rate of transition from open to close SaCas9-sgRNA structures is indicated by the size of the arrows.