Critical role of DNA unwinding and Cas9 conformational changes in modelling off-target effects

Aset Khakimzhan¹ and Vincent Noireaux¹

¹School of Physics and Astronomy, University of Minnesota, 115 Union Street SE, Minneapolis, MN 55455, USA

E-mail: khaki005@umn.edu; noireaux@umn.edu

Abstract

CRISPR-Cas9 off-target effects interfere with the ability to accurately perform genetic edits. To predict off-target effects CRISPR-Cas9 researchers perform high throughput guide RNA mismatch and bulge experiments and then use the data to fit thermodynamic binding models. While impactful from an engineering perspective such models are not based on the experimentally observed target interrogation process and thus incorrectly measure the energetic effects mismatches have on the system. In this work we convert an experimentally deduced qualitative model of target interrogation to a linear ODE model and demonstrate that the mismatch tolerance patterns observed in experiments do not need to be caused by differences in energetic penalties of mismatches but rather are emergent effects of the timing and coordination of target DNA unwinding and Cas9 conformational changes.

Introduction

A major bottleneck in the use of CRISPR-Cas9 lies in the off-target effects [1]. Off-target effects are caused by a CRISPR-Cas9 complex binding and cleaving an unintended site on the genome. One strategy to decrease off-target effects is the development of thermodynamic guide RNA optimizing models trained on large mismatch and bulge guide RNA datasets [2–6]. By comparing the cleavage/binding/indel rates between fully complementary guides and mismatched guides researchers inferred the energetic costs accrued by each spacer mutation. While such models are widely used for guide RNA development for Cas9 (and Cas12 [7]) systems they have one major flaw: they ignore much of the latest mechanistic experimental results of CRISPR-Cas9 target interrogation. This approach is understandable for the early stages of CRISPR-Cas9 models but now that the field has acquired a wide range of mechanistic experimental evidence it is important that off-target effect models account for them as well.

In this work we develop a CRISPR-Cas9 model that is based on the latest experimental target interrogation results. An important distinction of this model is the accounting for the discrete steps in which Cas9 unwinds the target DNA. This element of the model is based on structural experiments [8–12], single molecule Forster resonance energy transfer (smFRET) experiments [13–17], molecular dynamics simulations [18–20], and magnetic tweezers experiments [21]. In addition to nucleic acid dynamics the model also accounts for the strategy by which CRISPR-Cas9 searches for the target DNA [22–24]. We demonstrate that the PAM-proximal mismatch tolerance is largely controlled by CRISPR-Cas9 rebinding conditions, while the middle-target and PAM-distal mismatch patterns are dependent of the rates of CRISPR-Cas9 conformational changes. The message of the model is that the mismatch tolerance profiles observed across multiple experiments are emergent of the topology and timing of CRISPR-Cas9's target recognition process rather than the energies of bonds themselves. In addition to

providing a qualitative picture of different mismatch profiles our results also demonstrate how Cas9 conformational changes can be thought of nature leveraging a speed-accuracy tradeoff. With all the parameters of the model explained we test the model against available mismatch data. To reduce the number of parameters, we convert our model into a previously described energy landscape scheme [25] and then utilize this energy landscape to fit data from 30 targets in HEK293T cells [26] using realistic parameter constraints.

Model of CRISPR-Cas9 target interrogation

The first steps of CRISPR-Cas9 cleavage are target search. Our model relies on the hypothesis that CRISPR-Cas9 search is a mix of 3D diffusion across the larger experimental space and 1D-like sliding between neighboring PAM-sites [23,24,27]. Previous theoretical works studying the dynamics of nucleic acid binding proteins [28] and specifically CRISPR-Cas9 [29] demonstrates that an element of 1D sliding/PAM localization is an important contribution to the search mechanism. The system starts in the overall experimental space (Fig 1A) and via 3D diffusion enters an arbitrary vicinity (Fig 1B) of the target at rate of k_S , with the rate of exiting that vicinity to the overall experimental space being k_{Lost} . The 1D motion corresponds to the binding and unbinding to the target DNA once in the arbitrary vicinity, labeled k_{ON} and k_{OFF} respectively. For our model we set k_S to a realistic for cells value of 10^{-2} s^{-1} , which correspond to an average target search time of 100 seconds. The rate of escaping the target vicinity to the general experimental space, k_{Lost} is set to 1 s^{-1} , which corresponds to a residence time of 1s. That is a realistic value based on theoretical analysis of CRISPR-Cas9 localization [29]. The measured residence time of non-target is on the order of 30ms [22–24] and therefore the rate of breaking a PAM bond, k_{OFF} , can be approximated as 30 s^{-1} . The neighboring PAM sites will also have a

residence time of approximately 30ms and considering a 1D sliding model a ½ probability of returning to the target site. Therefore an expected timeframe for which CRISPR-Cas9 will rebind to the correct target is on the scale of a 100ms and thus we approximate k_{ON} as 10 s^{-1} .

Upon forming a PAM bond, the CRISPR-Cas9 system starts displacing the dsDNA DNA:DNA bonds with spacer-target RNA:DNA bonds starting from the PAM-proximal base pairs (Fig 1C). In our model the rate of displacing a single base pair with a matching RNA:DNA bond happens at a rate of k_{fc} , while the reverse process (displacing the RNA:DNA with a DNA:DNA bond) happens at a rate of k_{rc} . Similarly, rate to form an RNA:DNA mismatch is k_{fm} , and the rate to reverse the RNA:DNA mismatch is k_{rm} . The strand displacement rates are approximated by nucleic acid energetics:

$$\begin{aligned} k_{fc} &= k_{bp} \exp(-E_C/k_b T) \\ k_{rc} &= k_{bp} \exp(E_C/k_b T) \\ k_{fm} &= k_{bp} \exp(-E_{MM}/k_b T) \\ k_{rm} &= k_{bp} \exp(E_{MM}/k_b T) \end{aligned} \tag{1}$$

Where k_{bp} is a normalizing term to set the rate for nucleic acid dynamics, E_C is the average free energy acquired for a matched displacement, and E_{MM} is the free energy acquired for a mismatched displacement. During the CRISPR-Cas9 strand displacement, the limiting rate is the melting of dsDNA. Experiments studying the dynamics of base pairing approximate the timescales of melting a single DNA:DNA is on the order of 3ms [30] and thus we approximate k_{bp} as equal to 300 s^{-1} . The average energies acquired or lost when replacing a DNA:DNA bond with an RNA:DNA match or an RNA:DNA mismatch are calculated by comparing the energies of DNA:DNA bonds with the energies of RNA:DNA bonds. The average match displacement

energy E_C is on the order of $0.5k_bT$, while the average mismatch displacement energy is on the scale of $5k_bT$.

However, the formation of the spacer-target hybrid cannot be modeled by only considering the nucleic acid dynamics. Structural [8,12], smFRET [14,17], and rotor bead tracking experiments [21] demonstrated that during target recognition the Cas9 enzyme undergoes conformational changes that are coordinated by the number of RNA:DNA bonds formed. Therefore, our model accounts for two major conformational changes: the seed to intermediate transition (Fig 1D to Fig 1E) and the intermediate to active conformational change (Fig 1F to Fig 1G). These transitions can only occur when a sufficient number of RNA:DNA bonds are formed and the more of the target DNA can be unwound. The number needed for a seed-intermediate conformational change is L_S and the number of RNA:DNA bonds needed for an intermediate-active conformational change is L_I . We set the lengths for L_S and L_I to 10bp and 18bp based on the unwinding of DNA in the magnetic bead study [21] and the rates of the transitions k_{SI} , k_{IS} , k_{IA} , k_{AI} are all approximated at 1 s^{-1} based on the rates of FRET experiments [17]. However, the rates of the conformational changes could also be on the order of tens of seconds based on results from AFM experiments [31]. CRISPR-Cas9 undergoes conformational changes to enter its catalytically active conformation in which it can cleave the target DNA. In the model the cleavage can happen from any of the states within the active conformation (between Fig 1G and Fig 1H) at a rate of k_{cat} . From kinetic experiments of CRISPR-Cas9, k_{cat} is on the order of 1 s^{-1} but is highly dependent on the Mg^{+2} of the experimental system [32]. Once the DNA is cut, the system enters the final and irreversible product state (Fig 1I).

From previous CRISPR-Cas9 cleavage kinetics experiments we know that probability of cutting target DNA in CRISPR-Cas9 experiments is approximated with the following formula:

$$p = (1 - \exp(-k_H t)) \quad (1)$$

Where k_H is the global cutting rate in a given CRISPR experiment. We can calculate k_H from our model by converting the system of ODEs into a Markov chain (difference equation). We convert the system by assuming that in some timeframe t_s that we set to 1us, the probabilities of forward transitions, p_i , and reverse transitions, q_i , can be approximated as:

$$p_i = t_s k_{i,i+1} \quad (2)$$

$$q_i = t_s k_{i,i-1} \quad (3)$$

To find the global cleavage rate by taking the inverse of the mean absorption time τ_H . Markovian absorption time, μ_0 , can be found by solving the following system of linear equations:

$$\mu_i = 1 + q_i \mu_{i-1} + r_i \mu_i + p_i \mu_{i+1} \quad (4)$$

and then scaled to real time using t_s as follows:

$$\tau_H = t_s \mu_0 \quad (5)$$

$$k_H = \tau_H^{-1} \quad (6).$$

To demonstrate with most clarity the effect conformational changes and target search carry on the mismatch tolerance of CRISPR-Cas9, we calculate the k_H for targets with a single mismatch in each of the 20 possible RNA:DNA bonds (assuming the spacer is 20bp long).

Parameter	Value	Description
k_S	0.01 s^{-1}	Rate of entering the target vicinity
k_{Lost}	1 s^{-1} [27,29]	Rate of escaping target vicinity
k_{ON}	10 s^{-1} [estimated]	Rate of binding to the target when in the target vicinity
k_{OFF}	30 s^{-1} [22–24]	Rate of breaking a PAM-bond
k_{bp}	300 s^{-1} [30]	Normalizing strand displacement rate

E_C	-0.5k _B T [33–35]	Average free energy acquired by replacing a DNA:DNA match with an RNA:DNA match
E_{MM}	5k _B T [36]	Average free energy lost by replacing a DNA:DNA match with an RNA:DNA mismatch
L_S	10bp [21]	An average required length of the spacer-target hybrid to unwind the intermediate segment of the target DNA
L_I	18bp [12,21]	An average required length of the spacer-target hybrid to unwind the active segment of the target DNA
$k_{SB}, k_{IS}, k_{IA}, k_{AI}$	1 s ⁻¹ [13,16,17]	Rates of transitions between conformations of CRISPR-Cas9
k_{cat}	1 s ⁻¹ [31,32]	Cleavage rate once in the active conformation

Table 1. The rates of the CRISPR-Cas9 target interrogation model.

Figure 1. Diagram of the CRISPR-Cas9 target interrogation dynamical system. **(a)** The system starts with CRISPR-Cas9 in the general experimental space, searching for the target. The search rate is the rate of entering the region where the target is located and is defined as k_{search} . **(b)** When in proximity with the target, CRISPR-Cas9 can either exit the area of proximity at a rate of k_{lost} or bind to the target PAM with a rate of k_{on} . **(c)** Upon forming a PAM bond, CRISPR-Cas9 can initiate the strand displacement of 'seed' dsDNA at a rate of k_{fc} for a complementary bond and k_{fm} for a non-complementary bond. The reversal rates for a complementary bond and a non-complementary bond are k_{rc} and k_{rm} respectively. **(d)** Upon forming all the 'seed' RNA:DNA bonds the CRISPR-Cas9 system is able to unwind the intermediate segment of the target dsDNA at a rate of k_{SI} . **(e)** Winding back the intermediate segment of the target dsDNA occurs at the rate k_{IS} and only when zero intermediate RNA:DNA bonds are formed. Again, upon unwinding the

system can form new RNA:DNA bonds at the rates k_{fc} and k_{fm} depending on the complementarity of the nucleotides. The diagram presents an example when the 12th RNA:DNA bond is not complementary and thus the formation of the 12th RNA:DNA occurs at a rate of k_{fm} instead of k_{fc} . **(f)** When CRISPR-Cas9 forms all the intermediate bonds, it can unwind the PAM-distal dsDNA and at the same enter the cleavage active conformation. **(g)** In the cleavage active conformation CRISPR-Cas9 can form more RNA:DNA bonds or proceed to cut the target DNA at a rate of k_{cat} . **(h)** The fully formed R-loop. **(i)** We approximate that cutting is irreversible and is thus the ultimate state of our model.

Results and discussion

Nucleic acid dynamics during CRISPR-Cas9 target interrogation

Since different target sequences might have different average match and mismatch energies, we observed the effect varying the nucleic acid free energies E_C and E_{MM} carried on the mismatch patterns of the model. Predictably, the model demonstrates that decreasing the average free energy acquired when displacing a single dsDNA bond with a matching RNA:DNA bond causes the system to be more tolerant of mismatches (Fig 2A). This is due to the increased stability of the partial RNA:DNA hybrids prior to the mismatch locations, which in turn increase the overall likelihood of surpassing the mismatched site. Similarly, decreasing the energetic penalty E_{MM} increases the likelihood of tolerating a mismatch (Fig 2B), because the relative rate of passing through a mismatch is accelerated in comparison to the rate of unbinding from the DNA. Surprisingly, in this model, the mismatch tolerance patterns are neither uniform, as might be expected for a naïve thermodynamic model, nor are the patterns monotonically increasing as might be expected for a strand displacement kinetic model [25]. To test whether this effect is caused by picking a specific value of k_{bp} , we varied the rates from slow to fast (Fig 2C). While

changing the rates of k_{bp} affects the mismatch tolerance of the system, the overall pattern remains the same. Another set of nucleic acid parameter are the average conformational change (or unwinding) lengths, which are hypothesized to depend on the sequence of the target DNA [21]. By varying the lengths L_S and L_I in the model, we could control how critical a certain RNA:DNA bond becomes in the target recognition process. When increasing L_S by *Int* the shape location of the maxima of both mid-spacer peaks also increases by *Int*. (Fig 2D). However, when increasing the value of L_I by *Int* only the shape of the second peak changes (Fig 2E).

Figure 2. Emergence of single mismatch tolerance in CRISPR-Cas9 systems. The mismatch tolerance patterns of mechanistically faithful CRISPR-Cas9 model are not a monotonic, but instead have multiple local maxima points dependent on the conformational change lengths. **(a)** The energetic parameter E_C corresponds to the average free energy acquired for each DNA:DNA bond replaced with an RNA:DNA bond. Decreasing E_C makes the system more stable per bond, which in turn increases the likelihood of tolerating a mismatch. **(b)** The other energetic

parameter, E_{MM} , corresponds to the energetic penalty of a single DNA:DNA matching bond replaced with an RNA:DNA mismatch. As the penalty energy is increased the system becomes more specific, since the energy acquired by matches cannot compensate for the mismatch as effectively. **(c)** The nucleic acid basal rate parameter, k_{bp} , plays an important role in the mismatch patterns of CRISPR-Cas9 binding. As we increase k_{bp} the mismatch tolerance also increases due to 2 reasons: (i) a larger k_{bp} increases the relative probability of forming an initial RNA:DNA bond instead of unbinding from the target DNA and (ii) the speed of overcoming a mismatch, k_{fm} , becomes relatively large in comparison to the rates of conformational changes and thus are ignored during target recognition. **(d) and (e)** demonstrate the effect changing the conformational change lengths L_S and L_I respectively have on the single mismatch profiles of CRISPR-Cas9. We notice that mismatches after conformational changes are tolerated the least, since those cause the probability density of the system to be concentrated RNA:DNA hybrid lengths in which the system can revert to the previous conformation.

Target search and re-binding of CRISPR-Cas9

Normally when modeling the search of a DNA binding enzyme, we approximate the rate of finding the target by some constant rate, k_{search} . Upon unbinding from the DNA, the enzyme returns to its unbound state and will re-bind to the target DNA with the same rate k_{search} . However, this approach does not take into consideration that when an enzyme re-binds to the target, the rate must be accelerated since the initial location is proximal to the target itself [28]. In the case of Cas9, this effect is amplified by the fact that target search for CRISPR-Cas9 is divided into 1-D diffusion between non-target PAM sites and 3-D diffusion in the cell/tube/microfluidic device (Fig 3A) [23,27]. If Cas9 unbinds from its target PAM site, it is likely to return to the starting PAM site much faster than would be proposed by k_{search} . For this

reason, in the model we split the target search and binding process into two steps: (i) searching for the general area of the target DNA and (ii) binding to the target DNA when in that region.

We then tested the effect the diffusion and binding parameters had on the global cleavage rates from equation (6) for single mismatch experiments. Depending on the site, the rate of escaping the proximity of the target might vary. To see how the escape rate affects the single mismatch patterns in the model we varied the rate k_{lost} (Fig 3B). We notice that k_{lost} mostly affects the mismatch tolerances of the PAM-proximal RNA:DNA bonds, showing that for smaller rates of 1-D escape, the tolerance for mismatches increases. We then varied the rate of unbinding from a target site can k_{off} and noticed that it again affects only the PAM-proximal binding to a significant level, while the middle and PAM-distal RNA:DNA bonds appear to maintain the same mismatch tolerance (Fig 3C). We noticed that in a system with a quicker 1-D rebinding rate, k_{on} the mismatch tolerance also increases only for the PAM-proximal RNA:DNA bonds (Fig 3D).

Figure 3. Separation of CRISPR-Cas9 target search in 2 steps. **(a)** A schematic of the difference between 3D diffusion space and the 1D sliding space during CRISPR-Cas9 target search. The search for the target-proximal region occurs with a rate of k_{search} , while escaping the target-proximal region occurs with a rate of k_{lost} . Within the target proximal region, the motion of CRISPR-Cas9 can be approximated as 1-D sliding between PAM sites. The rate of rebinding to the target within the region is approximated with a single on-rate k_{on} and the unbinding from the target site occurs with some rate k_{off} defined by the energetics of the bond. **(b)** Decreasing k_{lost} , causes not only an emergence of PAM-proximal mismatch tolerance, but an increase in the

mismatch tolerance as a whole. **(c)** Decreasing k_{off} causes the system to be more tolerance of PAM-proximal mismatches but does not significantly affect the rates of the targets with mismatches past L_S . **(d)** Increasing the rate k_{on} strongly decreases the specificity of the PAM-proximal bonds, while decreasing the specificity of the mid-target and PAM-distal bond to a lesser extent.

Role of DNA unwinding rates in CRISPR-Cas9 target interrogation

When we increase the seed-intermediate transition rate, k_{SI} , we observe that the system becomes overall more error prone in comparison to the fully matched target (Fig 4A). Increasing the rate of the reverse process, k_{IS} , slows down the interrogation process and thus decreases the rate of cutting for mismatched targets, without significantly affecting the cleavage rate of a fully matched target (Fig 4B). For the intermediate-active transition, increasing k_{IA} also reduced the specificity of the system but only for the base pairs in the region between L_S and L_I (Fig 4C). Varying the rate of the reverse process, k_{AI} , did not affect the mismatch profile significantly (Fig 4D). This is because the rate of forming the base-pairs after L_I is faster than k_{AI} and thus the fully formed target-spacer hybrid serves as a sink in the dynamical system. Unsurprisingly, when increasing the cleavage rate k_{Cat} the model demonstrates that CRISPR-Cas9 becomes less specific (Fig 4E) as demonstrated previously in experiments with the enhanced specificity mutants of Cas9.

We can notice that as we increase the ‘gatekeeping’ rates, k_{SI} and k_{IA} , the system reaches the product state faster, but also loses specificity. The mechanism of CRISPR-Cas9 leverages a speed-accuracy tradeoff: the slower the conformational changes, the larger the expected number times a mismatched nucleotide is interrogated. With current directed evolution and rational protein design techniques it might be possible to re-engineer Cas9 (or other Cas enzymes) to

have slower or faster transition rates. Therefore, it might be of interest for future work to derive the optimal transition rates as a function of the CRISPR enzymes concentration, spatially dependent diffusion coefficient, and sequences of off-target sites.

Figure 4. Rates of CRISPR-Cas9 conformational changes/unwinding on mismatch patterns. (a), (b), (c), and (d) are the mismatch profiles produced by the model when varying the rates k_{SI} , k_{IS} , k_{IA} , and k_{AI} , respectively. (e) The mismatch profiles produced by the model for different rates of k_{cat} .

4 Conclusion

Our model demonstrates that it is unnecessary to assign energetic costs to each separate mismatch of a target to explain the patterns seen in CRISPR-Cas9 off-target experiments. The mismatch patterns can be explained as an emergent effect of the dynamics and topology of the system. However, this model uses only approximate rates and models the target recognition

process of CRISPR-Cas9 without considering the spatial effects of the Cas9 enzyme and the nucleic acids. Thus, our model only serves as a semi-quantitative picture of the CRISPR-Cas9 target interrogation to convince CRISPR-Cas9 community to incorporate the dynamics and mechanism of the Cas9 enzyme into off-target prediction software [37].

References

1. Keep off-target effects in focus. *Nature Medicine*. 2018. p. 1081. doi:10.1038/s41591-018-0150-3
2. Farasat I, Salis HM. A Biophysical Model of CRISPR/Cas9 Activity for Rational Design of Genome Editing and Gene Regulation. *PLoS Computational Biology*. 2016;12. doi:10.1371/journal.pcbi.1004724
3. Jung C, Hawkins JA, Jones SK, Xiao Y, Rybarski JR, Dillard KE, et al. Massively Parallel Biophysical Analysis of CRISPR-Cas Complexes on Next Generation Sequencing Chips. *Cell*. 2017;170: 35-47.e13. doi:10.1016/j.cell.2017.05.044
4. Jones SK, Hawkins JA, Johnson N v., Jung C, Hu K, Rybarski JR, et al. Massively parallel kinetic profiling of natural and engineered CRISPR nucleases. *Nature Biotechnology*. 2021;39: 84–93. doi:10.1038/s41587-020-0646-5
5. Boyle EA, Becker WR, Bai HB, Chen JS, Doudna JA, Greenleaf WJ. Quantification of Cas9 binding and cleavage across diverse guide sequences maps landscapes of target engagement. *bioRxiv*. 2020; 1–43. doi:10.1101/2020.09.09.290668
6. Alkan F, Wenzel A, Anthon C, Havgaard JH, Gorodkin J. CRISPR-Cas9 off-targeting assessment with nucleic acid duplex energy parameters. *Genome Biology*. 2018;19: 1–13. doi:10.1186/s13059-018-1534-x
7. Specht DA, Xu Y, Lambert G. Massively parallel CRISPRi assays reveal concealed thermodynamic determinants of dCas12a binding. *Proceedings of the National Academy of Sciences of the United States of America*. 2020;117: 25–27. doi:10.1073/pnas.1918685117
8. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*. 2014;156: 935–949. doi:10.1016/j.cell.2014.02.001
9. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science*. 2014;343. doi:10.1126/science.1247997
10. Jiang F, Taylor DW, Chen JS, Kornfeld JE, Zhou K, Thompson AJ, et al. Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science*. 2016;351: 867–871. doi:10.1126/science.aad8282

11. Josephs EA, Kocak DD, Fitzgibbon CJ, McMenemy J, Gersbach CA, Marszalek PE. Structure and specificity of the RNA-guided endonuclease Cas9 during DNA interrogation, target binding and cleavage. *Nucleic Acids Research*. 2015;43: 8924–8941. doi:10.1093/nar/gkv892
12. Sternberg SH, Lafrance B, Kaplan M, Doudna JA. Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature*. 2015;527: 110–113. doi:10.1038/nature15544
13. Singh D, Sternberg SH, Fei J, Doudna JA, Ha T. Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9. *Nature Communications*. 2016;7: 1–8. doi:10.1038/ncomms12778
14. Dagdas YS, Chen JS, Sternberg SH, Doudna JA, Yildiz A. A conformational checkpoint between DNA binding and cleavage by CRISPR-Cas9. *Science Advances*. 2017;3: 1–9. doi:10.1126/sciadv.aao0027
15. Chen JS, Dagdas YS, Kleinstiver BP, Welch MM, Sousa AA, Harrington LB, et al. Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature*. 2017;550: 407–410. doi:10.1038/nature24268
16. Osuka S, Isomura K, Kajimoto S, Komori T, Nishimasu H, Shima T, et al. Real-time observation of flexible domain movements in CRISPR–Cas9. *The EMBO Journal*. 2018;37: 1–11. doi:10.15252/embj.201796941
17. Yang M, Peng S, Sun R, Lin J, Wang N, Chen C. The Conformational Dynamics of Cas9 Governing DNA Cleavage Are Revealed by Single-Molecule FRET. *Cell Reports*. 2018;22: 372–382. doi:10.1016/j.celrep.2017.12.048
18. Palermo G, Miao Y, Walker RC, Jinek M, McCammon JA. CRISPR-Cas9 conformational activation as elucidated from enhanced molecular simulations. *Proceedings of the National Academy of Sciences of the United States of America*. 2017;114: 7260–7265. doi:10.1073/pnas.1707645114
19. Ricci CG, Chen JS, Miao Y, Jinek M, Doudna JA, McCammon JA, et al. Deciphering Off-Target Effects in CRISPR-Cas9 through Accelerated Molecular Dynamics. *ACS Central Science*. 2019;5: 651–662. doi:10.1021/acscentsci.9b00020
20. East KW, Newton JC, Morzan UN, Narkhede YB, Acharya A, Skeens E, et al. Allosteric Motions of the CRISPR-Cas9 HNH Nuclease Probed by NMR and Molecular Dynamics. *Journal of the American Chemical Society*. 2020;142: 1348–1358. doi:10.1021/jacs.9b10521
21. Ivanov IE, Wright A v., Cofsky JC, Palacio Aris KD, Doudna JA, Bryant Z. Cas9 interrogates DNA in discrete steps modulated by mismatches and supercoiling. *Proceedings of the National Academy of Sciences of the United States of America*. 2020;117: 5853–5860. doi:10.1073/pnas.1913445117
22. Jones DL, Leroy P, Unoson C, Fange D, Ćurić V, Lawson MJ, et al. Kinetics of dCas9 target search in *Escherichia coli*. *Science*. 2017;357: 1420–1424. doi:10.1126/science.aah7084
23. Yang M, Zhang L, Sun R, Zhong W, Yang Y, Chen C. *Streptococcus pyogenes* Cas9 displays biased one-dimensional diffusion on dsDNA to search for a target. *bioRxiv*. 2019; 1–14. doi:10.1101/754564

24. Martens KJA, van Beljouw SPB, van der Els S, Vink JNA, Baas S, Vogelaar GA, et al. Visualisation of dCas9 target search in vivo using an open-microscopy framework. *Nature Communications*. 2019;10: 1–11. doi:10.1038/s41467-019-11514-0
25. Klein M, Eslami-Mossallam B, Arroyo DG, Depken M. Hybridization Kinetics Explains CRISPR-Cas Off-Targeting Rules. *Cell Reports*. 2018;22: 1413–1423. doi:10.1016/j.celrep.2018.01.045
26. Kim N, Kim HK, Lee S, Seo JH, Choi JW, Park J, et al. Prediction of the sequence-specific cleavage activity of Cas9 variants. *Nature Biotechnology*. 2020;38: 1328–1336. doi:10.1038/s41587-020-0537-9
27. Globyte V, Lee SH, Bae T, Kim J, Joo C. CRISPR /Cas9 searches for a protospacer adjacent motif by lateral diffusion . *The EMBO Journal*. 2019;38. doi:10.15252/embj.201899466
28. Berg OG, Winter RB, von Hippel PH. Diffusion-Driven Mechanisms of Protein Translocation on Nucleic Acids. 1. Models and Theory. *Biochemistry*. 1981;20: 6929–6948. doi:10.1021/bi00527a028
29. Lu Q, Bhat D, Stepanenko D, Pigolotti S. Search and localization dynamics of the CRISPR/Cas9 system. 2021; 1–5. Available: <http://arxiv.org/abs/2103.10667>
30. He G, Li J, Ci H, Qi C, Guo X. Direct Measurement of Single-Molecule DNA Hybridization Dynamics with Single-Base Resolution. *Angewandte Chemie - International Edition*. 2016;55: 9036–9040. doi:10.1002/anie.201603038
31. Shibata M, Nishimasu H, Kodera N, Hirano S, Ando T, Uchihashi T, et al. Real-space and real-Time dynamics of CRISPR-Cas9 visualized by high-speed atomic force microscopy. *Nature Communications*. 2017;8: 1–9. doi:10.1038/s41467-017-01466-8
32. Gong S, Yu HH, Johnson KA, Taylor DW. DNA Unwinding Is the Primary Determinant of CRISPR-Cas9 Activity. *Cell Reports*. 2018;22: 359–371. doi:10.1016/j.celrep.2017.12.041
33. Sugimoto N, Nakano S ichi, Katoh M, Matsumura A, Nakamuta H, Ohmichi T, et al. Thermodynamic Parameters To Predict Stability of RNA/DNA Hybrid Duplexes. *Biochemistry*. 1995;34: 11211–11216. doi:10.1021/bi00035a029
34. Natsume T, Ishikawa Y, Dedachi K, Tsukamoto T, Kurita N. Hybridization energies of double strands composed of DNA, RNA, PNA and LNA. *Chemical Physics Letters*. 2007;434: 133–138. doi:10.1016/j.cplett.2006.12.017
35. Ouldrige TE, Šulc P, Romano F, Doye JPK, Louis AA. DNA hybridization kinetics: Zippering, internal displacement and sequence dependence. *Nucleic Acids Research*. 2013;41: 8886–8895. doi:10.1093/nar/gkt687
36. Watkins NE, Kennelly WJ, Tsay MJ, Tuin A, Swenson L, Lee HR, et al. Thermodynamic contributions of single internal rAdA, rCdC, rGdG and rUdT mismatches in RNA/DNA duplexes. *Nucleic Acids Research*. 2011;39: 1894–1902. doi:10.1093/nar/gkq905
37. Willard J, Jia X, Xu S, Steinbach M, Kumar V. Integrating Physics-Based Modeling with Machine Learning: A Survey. 2020;1: 1–34. Available: <http://arxiv.org/abs/2003.04919>