Lessons from Enzyme Kinetics Reveal Specificity Principles for RNA-Guided Nucleases in RNA Interference and CRISPR-Based Genome Editing

Namita Bisaria,1,5,6 Inga Jarmoskaite,1,6 and Daniel Herschlag1,2,3,4,7,*
1Department of Biochemistry
2Department of Chemical Engineering
3Department of Chemistry
4Stanford ChEM-H
Stanford University, Stanford, CA 94305, USA
5Present address: Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA
6Co-first author
7Lead Contact
*Correspondence: herschla@stanford.edu
http://dx.doi.org/10.1016/j.cels.2016.12.010

RNA-guided nucleases (RGNs) provide sequence-specific gene regulation through base-pairing interactions between a small RNA guide and target RNA or DNA. RGN systems, which include CRISPR-Cas9 and RNA interference (RNAi), hold tremendous promise as programmable tools for engineering and therapeutic purposes. However, pervasive targeting of sequences that closely resemble the intended target has remained a major challenge, limiting the reliability and interpretation of RGN activity and the range of possible applications. Efforts to reduce off-target activity and enhance RGN specificity have led to a collection of empirically derived rules, which often paradoxically include decreased binding affinity of the RNA-guided nuclease to its target. We consider the kinetics of these reactions and show that basic kinetic properties can explain the specificities observed in the literature and the changes in these specificities in engineered systems. The kinetic models described provide a foundation for understanding RGN targeting and a necessary conceptual framework for their rational engineering.

Introduction
Over the past decades, RNA-guided nucleases (RGNs) have emerged as powerful and versatile tools for genome editing and for manipulation of gene expression. RNA interference (RNAi), CRISPR-based genome editing, and ribozyme targeting utilize short oligonucleotides to guide enzymatic activity to particular DNA and RNA sequences through base-pairing interactions. The established rules for Watson-Crick base-pairing and their predictable energetics have led to the prevailing perspective that RGNs and related systems are readily programmable and inherently specific. Nevertheless, experimental studies have exposed limitations to RGN specificity, revealing persistent targeting of unintended sequences that resemble the target RNA or DNA sequence (“off-targets”) (Jackson et al., 2003; Tsai et al., 2016). This limited specificity is one of several key challenges that must be overcome to allow RGN-based technologies to achieve their full impact in the laboratory, in engineering, and ultimately as therapies.

A series of successes in reducing off-target cleavage have been reported, most recently in the CRISPR-Cas9 field, by making changes in the nuclease or the RNA guide (Doench et al., 2016; Fu et al., 2014; Kleinstiver et al., 2016; Mali et al., 2013; Ran et al., 2013; Slaymaker et al., 2016; Tsai et al., 2014). A common theme that emerged from studies of CRISPR-Cas9 targeting is that specificity of targeting increases upon weakening interactions of the Cas9-RNA complex with the target DNA—e.g., by using truncated RNA guides (Fu et al., 2014; Tsai et al., 2015) or by introducing mutations to the protein-DNA interface (Kleinstiver et al., 2016; Slaymaker et al., 2016). This counterintuitive observation has been rationalized in terms of an “excess energy” model, postulating that binding energy beyond a certain threshold stabilizes binding to both correct and incorrect targets (Fu et al., 2014; Hu et al., 2016; Kleinstiver et al., 2016; Tsai et al., 2015; Tsai and Joung, 2016; Wyvekens et al., 2015). Similar observations have previously been made for RNAi, RNase H, and ribozyme targeting and in the related fields of TALEN and zinc-finger-nuclease targeting (Gulingier et al., 2014a; Herschlag, 1991; Pattanayak et al., 2011; Pedersen et al., 2014; Østergaard et al., 2013). However, a mechanistic explanation for the “excess energy” phenomenon in CRISPR-Cas9 and RNAi has not been presented. Such a mechanistic explanation that takes into account the kinetics of individual steps of the targeting process will help transform isolated successes into generalizable approaches that will, in turn, direct the development of new rational strategies for improving the specificity of RGNs.

Here, we demonstrate the importance—and power—of kinetic considerations for determining and improving RGN specificity. RGN targeting can be described by a simple kinetic model that includes binding followed by a rapid downstream step such as cleavage, a model that is supported by published biochemical studies of RGNs and is consistent with the strategies utilized in the CRISPR-Cas9 and RNAi field to increase specificity. The concepts described here provide a conceptual framework for...
future engineering to enhance the specificity of RGN systems and for accurate predictions of RGN targets in vivo.

**Kinetic Regimes of Enzyme Specificity**

An intuitive approach to understanding and predicting RGN specificity is to use formalisms from enzyme kinetics. RGN specificity can be defined as the ratio of cleavage efficiencies for competing substrates (see Box 1). The prevailing perspective of the reaction catalyzed by RGNs, explicitly or implicitly, assumes rapid equilibrium of binding before the cleavage step (or other downstream steps (Wang et al., 2016) (Box 1)), an assumption also made in the original Michaelis-Menten kinetic formulation (Michaelis and Menten, 1913; Michaelis et al., 2011). In this “rapid equilibrium” kinetic regime, cleavage is slower than RGN dissociation (see Box 1), which allows RGN binding to reach equilibrium. Consequently, the difference in cleavage efficiencies (i.e., specificity) between a target and potential off-target is equal to the difference in binding affinity between the two sites (if the rates of cleavage of the two targets are the same). In an alternative kinetic regime, which we refer to as “sticky,” dissociation of the RGN from the target is slow relative to the downstream rate of cleavage (Briggs and Haldane, 1925) (Box 1). In this regime, binding equilibrium is not established, and the difference in cleavage efficiencies is less than the difference.
Figure 1. Kinetic Model for RGN-Targeting Reaction

Top: two-step kinetic model illustrating a generic RGN-targeting reaction with a matched (black) and mismatched (red) target. $k_{on}$, $k_{off}$, $k_{cat}$: rate constants for target binding, dissociation, and cleavage, respectively. Bottom: prediction for specificity for the matched target, assuming that the mismatched target is bound 100-fold less stably than the mismatched target and that this difference is based on the dissociation rate ($k_{off}$), while the association and cleavage rate constants are the same for both targets ($k_{on}$). The association and dissociation rate constants are $k_{on} = 1 \times 10^7$ M$^{-1}$ s$^{-1}$, $k_{off} = 3.5 \times 10^{-1}$ s$^{-1}$, based on values for mammalian AGO2 and the let7a miRNA and target (Salomon et al., 2015). Specificity is defined as the ratio of $k_{on}/k_{off}$ values for matched versus mismatched target (Box 1). The $K_D$ values on the horizontal axis correspond to the equilibrium dissociation constants for the correct target ($K_D = k_{off}/k_{on}$), where $K_D$ varies from $10^{-6}$ to $10^{-8}$.

### RGNs in the Sticky Kinetic Regime

Extensive biochemical studies have focused on the target binding and cleavage by Argonaute enzymes, and kinetic measurements of fly Ago2 and mammalian AGO2 enzymes support the model that these RGNs likely exist in the sticky kinetic regime (Salomon et al., 2015; Wee et al., 2012). In vitro, fly Ago2 cleaves a fully complementary target with a rate constant that is almost three orders of magnitude faster than dissociation (Table 1; $k_{cat}$ versus $k_{off}$) (Wee et al., 2012). Thus, for every 1,000 fully complementary targets that bind, ~999 are cleaved, and only one is expected to dissociate. The difference between cleavage ($k_{cat}$) and dissociation ($k_{off}$) is so large that even for targets with multiple mismatches, cleavage may occur before the target has an opportunity to dissociate.

Recent single-molecule measurements indicate that mouse AGO2 is also in the sticky enzyme kinetic regime for substrates that are fully complementary to the guide RNA, and the same likely holds for the 99% identical human AGO2 (Table 1, Salomon et al., 2015). Moreover, in vitro kinetic measurements of *Streptococcus pyogenes* Cas9 strongly suggest that the enzyme operates in a sticky regime: the observed cleavage rate constants are fast (reported values range from ~1–10 min$^{-1}$ (Gasiunas et al., 2012; Jinek et al., 2012; Sternberg et al., 2014, 2015; Szczelkun et al., 2014)), while dissociation from fully matched targets has been slower than observed timescales, estimated to occur over hours (Knight et al., 2015; Sternberg et al., 2014; Singh et al., 2016; Richardson et al., 2016).

Operating in the sticky enzyme regime may be of value to RNAi and CRISPR systems in vivo. For example, fly Ago2 has been implicated in siRNA-mediated silencing for defense against viral infection, whereby foreign RNA is processed by the RNAi machinery to generate RNA guides against additional copies of the intruder RNA (Wang et al., 2006). When under a viral onslaught, maximizing cleavage rates may be more important than maximizing specificity, and a certain degree of promiscuity may be beneficial in RNAi and CRISPR defense systems to accommodate rapidly mutating sequences of invading nucleic acids (Carroll, 2013; Paez-Espino et al., 2015; Sun et al., 2013). Furthermore, minimizing off-target effects may not be a major selective pressure for bacterial RGNs (such as CRISPR-Cas9), in part because of the small sizes of prokaryotic genomes (e.g., the genome of *Streptococcus pyogenes*, the source of the widely used Cas9 protein, is only ~2 Mb, more than 1000-fold smaller than the human genome).

Individual RNA-guided targeting systems evolved under different sets of selective pressures that presumably have optimized specificities and turnover rates for particular natural contexts. Consequently, it may not be surprising that initial efforts to apply these systems for laboratory applications revealed specificities that fall short of those required to target unique sites in mammalian transcriptomes or genomes (Jackson et al., 2003; Tsai and Joung, 2016). Developing RGN systems to be highly...
been utilized in similar fashion to the RGNs discussed herein. Additional effective strategies to increase specificity will involve increasing dissociation becomes faster than cleavage. Thus, effective strategies to increase specificity will involve increasing $k_{\text{off}}$ or decreasing $k_{\text{cat}}$, and the greatest discrimination will be achieved upon simultaneously changing both rate constants. Indeed, many of the successful approaches reported in the literature appear to follow these strategies (Table 2). Additional effective strategies to increase specificity that are distinct from the two present below are discussed in Box 2.

### Strategies for Engineering RGNs to Increase Specificity

We present two general strategies for improving specificity based on the kinetic considerations laid out in the previous sections. The unifying premise is that the ability of an RGN to discriminate between correct and mismatched targets increases as dissociation becomes faster than cleavage. Thus, effective strategies to increase specificity will involve increasing $k_{\text{off}}$ or decreasing $k_{\text{cat}}$, and the greatest discrimination will be achieved upon simultaneously changing both rate constants. Indeed, many of the successful approaches reported in the literature appear to follow these strategies (Table 2). Additional effective strategies to increase specificity that are distinct from the two present below are discussed in Box 2.

**Strategy A: Increasing $k_{\text{off}}$**

The RGN literature is rich in examples of increased specificity upon perturbing RGN-target interactions, and many of these examples could be explained by increased $k_{\text{off}}$. The most straightforward strategy to weaken target-guide interactions involves shortening the guide to decrease the number of base-pairs formed with the target sequence. For example, truncated CRISPR guide RNAs showed reduced off-target activity, in some cases—with no observed reduction of cleavage of the intended target (Fu et al., 2014; Tsai et al., 2015). This observation is consistent with a switch from the sticky regime to a rapid-equilibrium regime (Figure 2, Strategy A), resulting in enhanced specificity. In practice, RNA guides can only be shortened by a few bases—both to maintain sufficient number of contacts for robust on-target activity and because recognizing individual genomic sites imposes minimum length requirements on the guide. Nevertheless, even truncating guide RNAs by two to three nucleotides provided substantial specificity enhancements in CRISPR-Cas9 targeting (Fu et al., 2014). Target binding can be weakened without sacrificing the number of residues recognized by lowering the average base-pair stability. As was originally described in the context of ribozyme targeting (Herschlag, 1991), the decreased average stability of a base pair (e.g., in an AU-rich versus a GC-rich sequence) leads to weakened target-guide interactions, and many of these examples could be explained by increased $k_{\text{off}}$. The most straightforward strategy to weaken target-guide interactions involves shortening the guide to decrease the number of base-pairs formed with the target sequence. For example, truncated CRISPR guide RNAs showed reduced off-target activity, in some cases—with no observed reduction of cleavage of the intended target (Fu et al., 2014; Tsai et al., 2015). This observation is consistent with a switch from the sticky regime to a rapid-equilibrium regime (Figure 2, Strategy A), resulting in enhanced specificity. In practice, RNA guides can only be shortened by a few bases—both to maintain sufficient number of contacts for robust on-target activity and because recognizing individual genomic sites imposes minimum length requirements on the guide. Nevertheless, even truncating guide RNAs by two to three nucleotides provided substantial specificity enhancements in CRISPR-Cas9 targeting (Fu et al., 2014). Target binding can be weakened without sacrificing the number of residues recognized by lowering the average base-pair stability. As was originally described in the context of ribozyme targeting (Herschlag, 1991), the decreased average stability of a base pair (e.g., in an AU-rich versus a GC-rich sequence) leads

<table>
<thead>
<tr>
<th>System</th>
<th>Modification</th>
<th>Hypothesized consequence</th>
<th>Strategy A/B</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNAi</td>
<td>Chemical modifications of guide RNA</td>
<td>Weaker guide-target interactions/ Perturbed active site</td>
<td>A/B</td>
<td>(Jackson et al., 2006; Vaish et al., 2011)</td>
</tr>
<tr>
<td></td>
<td>Introduction of DNA nucleotides in the seed region</td>
<td>Weaker guide-target interactions</td>
<td>A</td>
<td>(Ji-Tei et al., 2008b)</td>
</tr>
<tr>
<td></td>
<td>Introduction of G:U base-pairs in seed region</td>
<td>Weaker guide-target interactions, weaker protein interactions with the guide-target helix</td>
<td>A</td>
<td>(Ji-Tei et al., 2008a)</td>
</tr>
<tr>
<td></td>
<td>Decreased GC content in guide RNA</td>
<td>Weaker guide-target interactions</td>
<td>A</td>
<td>(Gu et al., 2014)</td>
</tr>
<tr>
<td></td>
<td>Introduction of seed mismatches to discriminate between alleles</td>
<td>Weaker guide-target interactions</td>
<td>A</td>
<td>(Dua et al., 2011; Ohnishi et al., 2008; Pfister et al., 2009)</td>
</tr>
<tr>
<td></td>
<td>Introduction of mismatches near the catalytic site</td>
<td>Perturbed active site</td>
<td>B</td>
<td>(Pfister et al., 2009)</td>
</tr>
<tr>
<td>CRISPR-Cas9</td>
<td>Truncated guide RNA</td>
<td>Weaker guide-target interactions</td>
<td>A</td>
<td>(Fu et al., 2014; Tsai et al., 2015)</td>
</tr>
<tr>
<td></td>
<td>Cas9 mutations neutralizing positive charge in the DNA binding grooves</td>
<td>Weaker Cas9-target interactions</td>
<td>A</td>
<td>(Kleinstiver et al., 2016; Slaymaker et al., 2016)</td>
</tr>
<tr>
<td>RNaseH</td>
<td>Chemical modifications of guide DNA</td>
<td>Perturbed active site</td>
<td>B</td>
<td>(Ostergaard et al., 2013)</td>
</tr>
</tbody>
</table>

aStrategies A and B correspond to increasing $k_{\text{off}}$ and decreasing $k_{\text{cat}}$, respectively, as described in the main text.

bRNase H is technically a DNA-guided nuclease that cleaves the complementary RNA target. We include it because it is conceptually similar to and has been utilized in similar fashion to the RGNs discussed herein.
to more discrimination in a random pool of target sequences, as more residues can be used in recognition before dissociation becomes slower than cleavage of mismatched targets. This concept can be extended to targeting by any RGN; for instance, specificity improvements in RNAi were observed after increasing the AU content in the guide (Gu et al., 2014), and the lower stability of DNA/RNA compared to RNA/RNA base pairs (Sugimoto et al., 1995) has been used to enhance specificity by replacing parts of the RNA guide with DNA (Ui-Tei et al., 2008b). It remains to be established to what extent the observed increases in specificity resulted from effects of these perturbations on \( k_{\text{on}} \) rather than other steps in the maturation and loading of the RNA-induced silencing complex (Gu et al., 2014; Khvorova et al., 2003; Schwarz et al., 2003). In cases involving DNA/RNA base pairs, the effects of increased AU content may be especially large because of the exceptionally low stability of dA-rU base pairs (Martin and Tinoco, 1980). Certain chemical modifications of the guide could also destabilize target binding by lowering base-pair stability or weakening contacts between target and the protein.

Target binding also can be weakened by introducing mutations into the nuclease itself. Significant improvements in CRISPR-Cas9 specificity have been achieved by substitutions of uncharged residues for positively charged ones in the DNA-binding grooves of Cas9, hypothesized to weaken electrostatic interactions with the DNA backbone (Kleinstiver et al., 2016; Slaymaker et al., 2016). It remains to be established whether the effect is primarily on \( k_{\text{off}} \) or on other reaction steps.

### Strategy B: Decreasing \( k_{\text{cat}} \)

Decreasing the rate of target cleavage can increase the specificity of a sticky enzyme by promoting binding equilibration prior to cleavage (Figure 2, Strategy B). Although a decrease in on-target cleavage is generally viewed as a flaw and corresponding mutations and constructs are dismissed, it may be fruitful to re-examine nuclease and guide variants that give slower cleavage, as they may be used to further optimize specificity.

### How can \( k_{\text{cat}} \) be manipulated?

In vitro measurements have shown that mismatches in the miRNA-target duplex in proximity of the cleavage site reduce \( k_{\text{cat}} \) (Wee et al., 2012), and the same likely applies to other RGNs, presumably as a result of active site perturbations. Chemical modifications of the guide have been reported to improve RNAi specificity, and some of these modifications may primarily affect \( k_{\text{cat}} \) rather than the dissociation rate or may affect both (Jackson et al., 2006; Vaish et al., 2011). Finally, nuclease domains with varying cleavage rates may be tethered to a catalytically inactive RGN to achieve a range of activities.

There will be a point at which these manipulations will make cleavage too slow to act on the required timescale, or dissociation will be too fast to efficiently target the desired molecules. Thus, there will be a kinetic “sweet spot” for a given RGN and application that can be identified through systematic studies. Moreover, some manipulations could potentially introduce detrimental effects on other steps (such as nuclease loading with the guide RNA), which will need to be accounted for. Given these complexities, one might be tempted to leave aside kinetic logic. We suggest instead that conceptual strategies such as those presented herein are even more important in complex, less-intuitive systems and will be needed to guide rational re-engineering of RGNs, to identify the limits of improved RGN specificity, and to determine the in vivo factors that adjust RGN reaction properties.

### A Practical Example: Discrimination between Single-Nucleotide Polymorphisms

RGNs have long been recognized as promising therapeutics because of their potential to specifically target mutated genes associated with disease, especially autosomal dominant disorders, and several RNAi-based therapies are currently in clinical trials (Sullenger and Nair, 2016; Wittrup and Lieberman, 2015; Xiong et al., 2016). Among potential applications, targeting single-nucleotide polymorphisms (SNPs)—which are linked to hereditary diseases, such as Huntington’s disease, hypertrophic...
cardiomyopathy, and amyotrophic lateral sclerosis—has proven to be a particularly challenging problem because the mutated target loci or associated SNPs differ from the wild-type (WT) allele by only a single nucleotide.

Early efforts toward therapeutic applications used short interfering RNA (siRNA) guides with full complementarity to the disease allele and, thus, a single mismatch to the WT allele at the site of disease-associated polymorphism (Figure 3). However, cell culture studies indicated little or no discrimination in silencing between WT and disease alleles (e.g., Miller et al., 2003; Pfister et al., 2009; Schwarz et al., 2006; Yu et al., 2012). Subsequently, much greater discrimination in siRNA-mediated silencing was found upon introduction of an additional mismatch against both the WT and disease allele (Dahlgren et al., 2008; Miller et al., 2003; Ohnishi et al., 2008; Pfister et al., 2009). This effect can be understood based on kinetic considerations. The first mismatch, against the WT allele only, is located in the central region of the guide RNA and, therefore, likely reduces $k_{cat}$ (Ameres et al., 2007; Haley and Zamore, 2004; Wee et al., 2012). The second mismatch, against both the WT and the disease allele, is located in the seed region and is thus expected to increase $k_{off}$ (Salomon et al., 2015; Wee et al., 2012). Although neither individual mismatch is sufficient to allow discrimination (sticky regime is maintained), the two concomitant changes—slowing cleavage and speeding dissociation—together allow the system to cross a threshold from the sticky regime to the rapid-equilibrium regime, where binding equilibrates prior to cleavage and dissociation occurs (Figure 1). Here, we have rationalized results that were arrived at through trial and error; in the future, kinetic frameworks can be used to guide analogous successful engineering efforts.

**Conclusions**

RGN systems hold tremendous technological and therapeutic promise, and recent advances in increasing their delivery, stability, and specificity have markedly lowered the barriers to wide application of these potent and versatile tools (Haussecker and Kay, 2015; Hendel et al., 2015; Hu et al., 2016; Sullenger and Nair, 2016; Wittrup and Lieberman, 2015; Xiong et al., 2016). Here, we approached the widespread problem of off-target effects. We explain previous observations of RGN specificity by considering the kinetics of the targeting reaction, and we codify these observations in terms of a simple and generalizable model. This model leads to rational strategies for minimizing off-target activity and will help guide efforts to design more specific RGN systems.

The most obvious culprit in the widespread inability of RGN systems to efficiently discriminate between targets and very similar sequences is the inherent high stability of base-pairing interactions that underlie target recognition (Herschlag, 1991). Very slow dissociation of stable target-guide duplexes leads to sticky kinetics, where the RGN cleaves matched and partially mismatched targets before having an opportunity to discriminate between them. This limitation appears to be alleviated, in part, by RGN proteins that destabilize guide-target interactions by deforming the duplex and blocking helix propagation along the full length of the guide (Bartel, 2009; Elkayam et al., 2012; Nakarnish et al., 2013, 2012; Schirle and MacRae, 2012; Schirle et al., 2014) and by conformational changes that dissipate some of the binding energy (Alber, 1981; Herschlag, 1988; Jiang et al., 2015; Johnson, 2008; Schirle et al., 2014; Sternberg et al., 2015). Effective general strategies to improve RGN specificity are through further weakening of the RGN-target interactions and slowing of cleavage rates.

Our simple kinetic framework highlights fundamental commonalities between different RGN systems, and we hope these similarities will generate cross-talk between RNAi, CRISPR, and related RGN fields. Although it might appear that the complexities and idiosyncrasies of each RGN system render descriptions from the standpoint of basic principles hopelessly naive, we believe just the opposite to be true. To understand and ultimately tame a complex system, a stepwise approach is needed. For each RGN, it will be important to build out from this simple, common framework to account for the individual reaction steps for that RGN and the kinetics and thermodynamics of each step.

To our knowledge, cleavage and dissociation kinetics are not currently taken into account in any prediction algorithm for RGN target sites and off-target effects. Quantitative studies of RGN systems (Chandradoss et al., 2015; De et al., 2013; Hale and Zamore, 2004; Jo et al., 2015; Salomon et al., 2015; Sternberg
destabilized by the mismatch at the SNP location. This additional destabilization leads to a change in $k_{cat}/K_{M}$ and, thus, an increase in discrimination between the two alleles by allowing the system to cross the threshold from a sticky kinetic regime to a rapid-equilibrium kinetic regime (Figure 1).

et al., 2014; Wee et al., 2012; Sterenberg et al., 2015; Singh et al., 2016) inaugurate the types of analyses that will allow physical models to be developed and predictions to be made for specificity and efficacy. Careful, in-depth studies that build on these assays and results will lead to frameworks that provide quantitative predictions for each RGN system and identify steps whose modulation will improve specificity. With this knowledge, cellular factors that interact with these systems and alter their kinetic and thermodynamic properties can be identified. We hope that this perspective will stimulate these mechanistic studies of RNAs, which in turn will enable more accurate predictions and rational approaches to RGN re-engineering and application.

ACKNOWLEDGMENTS

We thank Liang Wee, members of the Herschlag and Bartel labs, Matt Porteus, Andy Fire, Stanley Qi, and Mark Kay for helpful discussions, comments, and for critical advice throughout this work. N.B. was supported by an NSF graduate fellowship, and this work was supported by a grant to D.H. from the National Institutes of Health [P01GM066275].

REFERENCES


