Why Argonaute is needed to make microRNA target search fast and reliable

Klein, Misha; Chandradoss, Stanley D.; Depken, Martin; Joo, Chirlmin

DOI
10.1016/j.semcdb.2016.05.017

Publication date
2017

Document Version
Accepted author manuscript

Published in
Seminars in Cell and Developmental Biology

Why Argonaute is needed to make microRNA target search fast and reliable.

Misha Klein[1,#], Stanley D. Chandradoss[1,#], Martin Depken[1,*], Chirlmin Joo[1,*]

[1] Kavli Institute of NanoScience and Department of BioNanoScience, Delft University of Technology, Delft, 2629HZ, The Netherlands
[#] Co-first authors
*Correspondence: s.m.depken@tudelft.nl , c.joo@tudelft.nl

Summary

MicroRNA (miRNA) interferes with the translation of cognate messenger RNA (mRNA) by finding, preferentially binding, and marking it for degradation. To facilitate the search process, Argonaute (Ago) proteins come together with miRNA, forming a dynamic search complex. In this review we use the language of free-energy landscapes to discuss recent single-molecule and high-resolution structural data in the light of theoretical work appropriated from the study of transcription-factor search. We suggest that experimentally observed internal states of the Ago-miRNA search complex may have the explicit biological function of speeding up search while maintaining specificity.

1. Introduction

Eukaryotes regulate gene expression post-transcriptionally through the RNA interference (RNAi) pathway. This pathway begins with the transcription of non-coding RNA and its subsequent maturation into microRNA (miRNA). To facilitate search and suppression of target messenger RNA (mRNA), Argonaute (Ago) proteins join together with the miRNA molecule, forming an efficient search complex (1,2). In the pool of cellular RNA, the search complex finds mRNA cognate to its miRNA and primes its degradation. As the search relies on thermal motion, the functioning of the search complex can be understood in terms of diffusion and the binding-energy landscape of mRNA-Ago-miRNA interactions. In this Review, we discuss recent single-molecule and structural data on Ago, and borrow free-energy considerations and theory from transcription-factor search, highlighting how several of the observed Ago conformations could function to speed up the search process.

2. Target search in 1D and 3D

Ever since the initial observations of an astonishingly high association rate of the *E. coli Lac* repressor to the *lac* operon (3), researchers have been trying to understand general mechanisms that could speed up target search on nucleic-acid templates. In their seminal work (4), Berg, Winter and von Hippel proposed a facilitated diffusion mechanism by which the protein combines three-dimensional diffusion through the cytoplasm with lateral diffusion along the DNA (see **Figure 1**) (5). We here qualitatively summarize the theoretical arguments behind this suggestion and review the experimental evidence for lateral diffusion by various search complexes.

2.1 Facilitated diffusion enables rapid target search of miRNA

Though facilitated diffusion was originally aimed at transcription-factor search on DNA, the same arguments apply to any searcher along a nucleic acid sequence, including Ago-miRNA search on RNA. The benefit of

employing both 3D and 1D search can be qualitatively understood as follows: To find the next sequence to probe, it will always be faster to diffuse a short distance laterally along the RNA (through hopping and sliding; **Figure 1**) than to diffuse a long distance through the cytosol. As lateral diffusion brings you to close-by sites, there exists a point beyond which the search complex starts predominantly probing sites already visited. At this point it becomes favorable to move to an unprobed RNA neighborhood by diffusing through the cytosol. Minimizing redundancy of the one-dimensional (1D) search thus comes at the cost of employing the slower 3D search, and there exists an optimum partitioning between the two (4,6-9).

2.2 Experimental evidence for lateral diffusion during target search

Single-molecule fluorescence studies brought direct evidence of lateral diffusion during molecular target search, including sliding of transcription factors (10,11), DNA repair proteins (12-14) zinc-finger proteins (15), and the DNA recombination protein RecA (16). Like Argonaute, RecA makes a nucleoprotein complex (a RecA—single-stranded DNA filament) that is ready to basepair for target search (17-21). In order to investigate lateral diffusion of Ago-miRNA on RNA, we adopted an *in vitro* single-molecule FRET assay that was developed for studying RecA-mediated target search (16). We placed two identical binding sites on a single target RNA strand, each of which led to a different FRET efficiency with Ago-miRNA bound (22). We observed that a substantial fraction of the binding events (>50%) shuttled between two strong binding positions via rapid lateral diffusion. When using a volume-occupying reagent (PEG) to mimic physiological conditions, most binding events (>90%) displayed shuttling by the same Ago-miRNA complex. This suggests that lateral diffusion could also be important for *in vivo* microRNA search.

3. Multiple protein configurations for fast lateral diffusion and stable target recognition

While target search is sped up by facilitated diffusion, Slutsky and Mirny (8,23) argued that it is not possible to have both fast lateral diffusion and stable/preferential binding to the target using a single nucleoprotein conformation. The more stable binding to the target is, the more stable binding to similar sequences also becomes, and the lateral diffusion slows down as it gets increasingly trapped at non-target sites. To understand what is needed for the resolution of this apparent paradox, we now follow Slutsky and Mirny (8,23) and consider the statistical variation of binding energies along the substrate (which for us is mRNA).

3.1 Resolving the speed-stability paradox by utilizing multiple binding modes

Apart from the target, the sequences being searched through can be considered as essentially random and uncorrelated (8,24). A substantially preferential binding to the target requires that a correct match has a considerable energetic difference (ΔE, for definition see **Figure 2A**) to all partial matches. Slutsky and Mirny assume that the search complex has a binding energy roughly proportional to the degree of sequence homology between probed and target sequence. Under the assumption that the binding energy comes only from individual nucleotide-basepairing energies, a large energetic difference between target and non-target positions can only be achieved by large differences in pairing for each nucleotide. A general increase of basepairing energies results in a larger standard deviation among binding energies at different positions (compare $\sigma_{\mathbf{R}}$ of the “recognition”

landscape and σ_S of the "search" landscape in **Figures 2B** and **C** respectively), and the diffusion constant along the mRNA can be shown to decrease sharply (8,25). In **Figure 2B** we illustrate how a large recognition energy will generally imply large barriers to lateral diffusion ($\Delta E^{\dagger}$, for definition see **Figure 2A**), resulting in a slow search process. Reversely, in **Figure 2C** we illustrate how small barriers to diffusion implies poor recognition. Slutsky and Mirny proposed that the coupling between recognition energy and diffusion barrier ($\Delta E^{\dagger}$ being proportional to ΔE) can be broken if the search complex can stochastically switch between two internal modes with different binding energy strength (**Figure 2D**):

A search (S) mode: small affinity differences and fast diffusion ($\sigma_S \lesssim 2k_BT$; Ref. (8))

A recognition (R) mode: large affinity differences and slow diffusion ($\sigma_R \gtrsim 5k_BT$; Ref. (8))

An efficient searcher must have evolved the ability to combine the search and recognition modes. Thereby, the non-specific (average) energies (dashed lines in **Figure 2B-D**) are arranged such that all energies of the search mode lie between the energies of all non-target sites and the target in the recognition mode (see **Figure 2D**). Such systems predominantly move according to the search mode when not at the target site, but predominantly occupy the recognition mode once at the target (see states with orange dots in **Figure 2D**). The effective search barriers are now set by the search mode ($\Delta E^{\dagger} \approx \Delta E_S^{\dagger}$) while the recognition energies are set by the recognition mode ($\Delta E \approx \Delta E_R$). Both fast search and stable recognition is thus in principle possible if the searching protein possesses at least two distinct binding modes, and the above case represents the theoretical ideal scenario (for more general cases see (6,26-29)).

3.2 Experimental evidence for two initial binding modes of Ago-miRNA

Both recent structural and single-molecule data of eukaryotic Ago proteins suggest that the hybridization between guide and target is gradual and is coupled to structural changes in the search complex. We here discuss these studies in the light of a search-stability paradox for Ago-miRNA.

Biochemical, structural and computational analyses suggest that Argonaute divides its miRNA guides into five functional domains (5'anchor, seed, mid region, 3' supplementary region, and the tail region) (**Figure 3).** The seed region (nt 2–8) is crucial for gene suppression (1,17-19,30-32), and it was shown that protein mediated interactions stabilize nt 2–6 into an A-form-helix that exposes nt 2–4 (or 2–5) for base paring with the target (**Figure 4A**) (33). Based on this observation, Schirle *et al* (33) proposed a step-wise target recognition for human Argonaute-2 (hAgo2), in which the initial recognition of the target occurs in the 5' part of the miRNA. Two subsequent single-molecule studies showed that Ago-miRNA indeed uses this so-called sub-seed for the initial weak recognition. Solomon *et al* designed di-nucleotide mutation constructs for mouse Ago-miRNA and measured the unbinding rate from the target RNA (34). We have also shown that, when the paired region was gradually shrunk from the full seed (nt 2-8) to only the first three nucleotides (nt 2-4), no difference in the binding rate was noticeable (22). These two results showed that it is only the first three nucleotides of the seed that are used to maintain weak interaction during the initial search.

The two single-molecule works also suggested that Ago-miRNA exhibits a sharp increase in the binding affinity when the number of paired nucleotides changes from 6 to 7 (22,34). Comparison of crystal structures suggests that this property originates from the fact that Argonaute makes the guide kink away from the A-form stacked structure in several places (17-20). The most prominent kink disrupting the helical arrangement of the guide is between nt 6 and 7 (**Figure 4B**). Base paring to the target, therefore, requires a shift of the helix-7 that clashes with the incoming target. After pairing of nt 2-4, hAgo2 undergoes a conformational change leading to a 4Å displacement of the helix-7 loop and allowing base pairing of nt 6–8 (**Figure 4C**). It was hypothesized that the sharp increase in the time bound between having 6 and 7 nt matching is caused by the conformational change of the helix-7 motif (33). We here suggest that Ago makes a change from a weak binding (search) mode using nt 2-4 to a strong binding (recognition) mode using a full seed through the conformational change of the helix-7.

3.3 The experimental evidence for additional binding modes of Ago-miRNA

In addition to the helix-7 movement, more conformational changes take place after seed pairing is achieved, and before the bound Ago-miRNA complex becomes cleavage competent.

First, binding of the supplementary region (nt 13-16) ensuing the seed pairing enhances the binding stability of Ago-miRNA (34). But the pairing beyond nt 8 is restricted by a physical constraint (33)(**Figure 5A**).
Widening up of a channel between PAZ and N-terminus domains allows for a rearrangement of the disordered supplementary region (nt 13-16) of the miRNA into a helical A-form, preparing it for pairing with the target RNA (**Figure 5B**)(33). It remains to be seen whether target recognition is enhanced by this additional checkpoint.

Second, biochemical and single-molecule studies have shown that the base paring in the mid region is necessary for cleavage of target RNA (32,35). But Jo *et al* also observed that a significant portion of Ago-miRNAs were not able to cleave the target RNAs in spite of their perfect complementarity (32,35). The unsuccessful cleavage of perfect complementary target might be the resultant of a failure to induce an additional conformational change needed for cleavage that involves positioning of Ago's catalytic residues residing near nt 9-10 of the miRNA.

Third, Ago uses its PAZ domain to preclude miRNA from being tightly associated with target RNA. An earlier biochemical study reported that bare RNA as short as 12bp is long enough for stable hybridization (~a year of life time) (36). But it was observed that Ago-miRNA (or Ago-guide DNA) often dissociated from its target within seconds to minutes after binding (32). This reversible binding, which is speculated to reduce off-targeting (37), is possible because the 3' end of guide RNA is anchored to the PAZ domain and this lowers the binding affinity of Ago-miRNA (especially at the 3' end) to target RNA (20,37-41).

In addition to the complex interactions between Ago and a guide strand, a direct interaction between Ago and target RNA also contributes to the target selection. Schirle *et al* (42) showed that hAgo2 interacts with the adenine nucleotide of the target when it is opposite to the 1st nucleotide of the guide. Through a water network, the residues in the MID domain (**Figure 3A**) specifically recognize the t1A anchoring the Ago-miRNA complex to the target. Using a single-molecule assay they showed that t1A does not influences initial target recognition

but increases the residence time of Ago-miRNA on to the target RNA, which might enhance its cleavage efficiency (42).

4. Energy landscape of miRNA target search

Having discussed the evidence that a series of conformational changes are needed to initiate stable binding and cleavage of target mRNA, we now discuss how conformational changes effect the binding-energy landscape. When Ago initially scans the target RNA it exposes only nucleotides 2-4 of the miRNA, termed the sub-seed. In this search mode it does not discriminate strongly based on RNA sequence, and lateral diffusion is likely rapid. A complete match of the sub-seed stabilizes a conformational change that exposes the remainder of the seed (nt 2-8) for base pairing, and, once paired, it slows down the diffusion in this recognition mode (**Figure 6A**). Upon encountering a sequence bearing complementarity to the entire seed, the helix-7 is displaced to allow miRNA to fully pair with the target, and the Ago-miRNA complex arrives in this more stable recognition state (**Figure 6A** and **B**). We suggest that the function of these various states is analogous to the function of internal states in transcription-factor search (**Figure 2D**).

In **Figure 6B** we sketch a free-energy landscape of the dominant configuration at varying degrees of base pairing for a perfect match. Transitions requiring conformational changes cost energy, increasing barriers to further base pairing. We construct a sketch of the landscape based on a single-molecule study that reported the existence of various pathways even when the full sequence of miRNA matches with a target (35): a significant fraction of the population showed transient binding (~10%) and stable binding with no cleavage (~30%).

Assuming that the largest barrier to further basepairing originates from the required movement of helix-7, the substantial fraction of transiently binding proteins indicates that this barrier must come close to the barrier to unbind. Further, the even larger fraction of stable but non-cleaving complexes indicates that the average binding energy past helix-7 is strong, and that the cleavage rate is slow compared to experimental times, but fast compared to unbinding.

With these general considerations, we conclude that the free-energy landscape of **Figure 6B** captures at least one search mode (pre-seed pairing) and at least one recognition mode (post-seed pairing). These two modes could be further split up, e.g. the seed pairing into sub-seed and full seed pairing. Still, the general principle behind resolving the speed-stability paradox should apply. To determine the quantitative effects of this energy landscape will require additional theoretical work accounting for gradual base pairing and a series of conformational changes. Using single-molecule techniques and high resolution structural studies, it will also be possible to test the effect of Ago's conformational changes on target search by analysing mutated proteins or directly observe conformational switching (for instance by using FRET such as done for Cas9 in (43)).

5. Outlook

We have reviewed the principles behind facilitated diffusion and the speed-stability paradox in general target search processes, as well as the experimental evidence for facilitated diffusion in miRNA target search. We further discussed the evidence for multiple search states in the Ago-miRNA search complex, which could help resolve the speed-stability paradox—simultaneously enabling the search to be fast and the binding to the target to be strong.

5.1 Further insight into Ago-miRNA target search can improve microRNA target prediction algorithms.

Due to the complex nature of the mRNA targeting process, it is far from straightforward to predict what genes are silenced by a particular miRNA. Experimentally, mRNA targets have been found by analysing the effect of miRNA expression on protein production or by performing binding assays(44) . For such approaches to work, one needs to know what target gene should be considered from the outset. Using bioinformatics algorithms, potential target sites are scored, and high scoring targets are subsequently tested in experiment.

Simple sequence homology between the mRNA to the guiding miRNA does not by itself give an accurate prediction of targets. Presently, typical prediction algorithms are largely phenomenological in nature, for example, assigning higher scores to sequences that fully match the seed of the miRNA and/or are evolutionary conserved. Additionally, accounting for the secondary structure of mRNA and the sequence outside of the targeted 3'-UTR further improves predictions (44,45). A recent combined bioinformatics and *in vivo* study showed that there are at least 14 additional sequence features (for example the length 3'-UTR region and the predicted structural accessibility of the RNA) of the mRNA that improve microRNA target prediction algorithms (46). Yet, despite much effort, prediction algorithms often point to many target sites that cannot be validated experimentally or fail to pick out targets that have been previously validated.

Single-molecule studies allow one to study how Ago-miRNA's interaction with RNA binding proteins effects target affinity. Synthesising such molecular level understanding into the free-energy landscapes that we have discussed in this review should help improving the scoring functions of target prediction algorithms by taking the non-equilibrium features of the system into account. Additionally, prediction algorithms can potentially be improved by taking sequences neighbouring the target into account (30,47-49). Chandradoss *et al.* showed that, when two identical targets are neighbouring each other, the total retention time was substantially larger than what can be expected on theoretical grounds for two non-interacting targets (22). This synergistic effect might also be observed when a target is neighboured by sub-seed sequences. It will be interesting to determine whether this putative effect exists *in vivo*. Possibly, modelling the physical interaction with neighbouring sites, and accordingly assigning higher scores to those mRNA sequences with a high-density of sub-seed sequences, could then improve target prediction algorithms.

5.2 Implications for other target search systems

In the cell, multiple nucleic acid-mediated target search processes take place. Among them, RecA-mediated target search is the most thoroughly studied system. Qi et al (50) selectively observed stable interactions between a RecA-ssDNA homologue and DNA in a DNA curtain experiment, in which single-molecule signals were only

observed when ssDNA and dsDNA matched with each other for at least 8 nucleotides. Furthermore, using single-molecule FRET, Ragunathan et al (16) observed short-lived interactions (1-10 s) between RecA-ssDNA and target DNA that had 5-7 matching nucleotides. The difference between having 7 or 8 matches suggests there exists a rate limiting step hampering RecA-ssDNA filaments to extend base pairing beyond the 7th nucleotide (similar to the barrier representing the movement of the helix-7 motif in **Figure 6B**).

Recently, great attention has been brought to the CRISPR/Cas system, an adaptive immune system in bacteria, which uses RNA as a guide to target foreign DNA or RNA (51). CRISPR's target search involves a protein-DNA interaction (recognition of a 3-nt sequence, so-called PAM sequence) and RNA-DNA interactions. Biochemical studies suggested that it is the PAM recognition that occurs prior to the seed recognition (52-54). Recently, a structural study showed that the first 8 nucleotides of Cas9's guide are pre-organized in a helical A-form, similar to the seed sequence of microRNA in Argonaute (55). A recent FRET study indicated that there is another mode that follows binding of the seed recognition (43). The authors showed that only when the guide RNA of Cas9 makes extensive base pairing (~16nt out of the 20nt guide), a nuclease domain (HNH) migrates towards the target DNA. Altogether, the findings imply that CRISPR/Cas9, similar to Argonaute, uses more than two binding modes to overcome the speed-stability paradox ('PAM only' to 'PAM+seed' to 'cleavage competent'). Whereas a DNA curtain assay ruled out long distance lateral diffusion, it will be interesting to find out whether the CRISPR-Cas system makes any local lateral excursions when searching for the PAM sequence. Similarly, no large scale lateral diffusion has been observed for RecA/Rad51 systems using DNA curtain assays (>100nm resolution) (56), while short-range lateral diffusion was observed in single-molecule FRET experiments (nanometer resolution) (16) .

Finally, it will be interesting to find out how much the search mechanism of human Argonaute-2 is shared with other target search systems such as those mentioned in this review and different classes of Ago proteins that use DNA to target DNA (57,58) and RNA to target DNA (59) as well as PIWI proteins (60).

ACKNOWLEDGEMENTS

C.J. was funded by European Research Council under the European Union's Seventh Framework Programme [FP7/2007-2013] / ERC grant agreement n° [309509]. M.D. was supported by a TU Delft start up grant. This work was supported by the Netherlands Organization for Scientific Research (NWO/OCW), as part of the Frontiers in Nanoscience program. We thank Aafke van den Berg and Tao Ju (Thijs) Cui for critically reading and commenting on the manuscript.

REFERENCES

1. Bartel, D.P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell*, **136**, 215-233.
2. Kim, V.N., Han, J. and Siomi, M.C. (2009) Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol*, **10**, 126-139.
3. Riggs, A.D., Bourgeois, S. and Cohn, M. (1970) The lac repressor-operator interaction. 3. Kinetic studies. *J Mol Biol*, **53**, 401-417.
4. Berg, O.G., Winter, R.B. and von Hippel, P.H. (1981) Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry*, **20**, 6929-6948.
5. von Hippel, P.H. and Berg, O.G. (1989) Facilitated Target Location in Biological-Systems. *Journal of Biological Chemistry*, **264**, 675-678.
6. Bauer, M. and Metzler, R. (2012) Generalized facilitated diffusion model for DNA-binding proteins with search and recognition states. *Biophysical Journal*, **102**, 2321-2330.
7. Coppey, M., Bénichou, O., Voituriez, R. and Moreau, M. (2004) Kinetics of target site localization of a protein on DNA: a stochastic approach. *Biophysical Journal*, **87**, 1640-1649.
8. Slutsky, M. and Mirny, L.A. (2004) Kinetics of protein-DNA interaction: facilitated target location in sequence-dependent potential. *Biophys J*, **87**, 4021-4035.
9. Zhou, H.-X. (2011) Rapid search for specific sites on DNA through conformational switch of nonspecifically bound proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 8651-8656.
10. Hammar, P., Leroy, P., Mahmutovic, A., Marklund, E.G., Berg, O.G. and Elf, J. (2012) The lac repressor displays facilitated diffusion in living cells. *Science*, **336**, 1595-1598.
11. Tafvizi, A., Huang, F., Fersht, A.R., Mirny, L.A. and van Oijen, A.M. (2011) A single-molecule characterization of p53 search on DNA. *Proc Natl Acad Sci U S A*, **108**, 563-568.
12. Blainey, P.C., van Oijen, A.M., Banerjee, A., Verdine, G.L. and Xie, X.S. (2006) A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. *Proc Natl Acad Sci U S A*, **103**, 5752-5757.
13. Gorman, J., Chowdhury, A., Surtees, J.A., Shimada, J., Reichman, D.R., Alani, E. and Greene, E.C. (2007) Dynamic basis for one-dimensional DNA scanning by the mismatch repair complex Msh2-Msh6. *Mol Cell*, **28**, 359-370.
14. Jeong, C., Cho, W.K., Song, K.M., Cook, C., Yoon, T.Y., Ban, C., Fishel, R. and Lee, J.B. (2011) MutS switches between two fundamentally distinct clamps during mismatch repair. *Nature structural & molecular biology*, **18**, 379-385.
15. Zandarashvili, L., Esadze, A., Vuzman, D., Kemme, C.A., Levy, Y. and Iwahara, J. (2015) Balancing between affinity and speed in target DNA search by zinc-finger proteins via modulation of dynamic conformational ensemble. *Proc Natl Acad Sci U S A*, **112**, E5142-5149.
16. Ragunathan, K., Liu, C. and Ha, T. (2012) RecA filament sliding on DNA facilitates homology search. *Elife*, **1**, e00067.
17. Elkayam, E., Kuhn, C.D., Tocilj, A., Haase, A.D., Greene, E.M., Hannon, G.J. and Joshua-Tor, L. (2012) The structure of human argonaute-2 in complex with miR-20a. *Cell*, **150**, 100-110.
18. Nakanishi, K., Weinberg, D.E., Bartel, D.P. and Patel, D.J. (2012) Structure of yeast Argonaute with guide RNA. *Nature*, **486**, 368-+.
19. Schirle, N.T. and MacRae, I.J. (2012) The crystal structure of human Argonaute2. *Science*, **336**, 1037-1040.
20. Wang, Y., Sheng, G., Juranek, S., Tuschl, T. and Patel, D.J. (2008) Structure of the guide-strand-containing argonaute silencing complex. *Nature*, **456**, 209-213.
21. Chen, Z., Yang, H. and Pavletich, N.P. (2008) Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature*, **453**, 489-484.

22. Chandradoss, S.D., Schirle, N.T., Szczepaniak, M., MacRae, I.J. and Joo, C. (2015) A Dynamic Search Process Underlies MicroRNA Targeting. *Cell*, **162**, 96-107.
23. Mirny, L.A., Slutsky, M., Wunderlich, Z., Tafvizi, A., Leith, J.S. and Kosmrlj, A. (2009) How a protein searches for its site on DNA: the mechanism of facilitated diffusion. *J. Phys. A: Math. Theor.*, **42**, 434013.
24. Gerland, U., Moroz, J.D. and Hwa, T. (2002) Physical constraints and functional characteristics of transcription factor-DNA interaction. *Proc Natl Acad Sci U S A*, **99**, 12015-12020.
25. Zwanzig, R. (1988) Diffusion in a rough potential. *Proceedings of the National Academy of Sciences*, **85**, 2029-2030.
26. Bénichou, O., Kafri, Y., Sheinman, M. and Voituriez, R. (2009) Searching fast for a target on DNA without falling to traps. *Physical Review Letters*, **103**, 1-4.
27. Murugan, R. (2010) Theory of site-Specific DNA-Protein interactions in the presence of conformational fluctuations of DNA binding domains. *Biophysical Journal*, **99**, 353-359.
28. Reingruber, J. and Holcman, D. (2011) Transcription factor search for a DNA promoter in a three-state model. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, **84**, 1-4.
29. Yu, S., Wang, S. and Larson, R.G. (2013) Proteins searching for their target on DNA by one-dimensional diffusion: Overcoming the "speed-stability" paradox. *Journal of Biological Physics*, **39**, 565-586.
30. Ameres, S.L., Martinez, J. and Schroeder, R. (2007) Molecular basis for target RNA recognition and cleavage by human RISC. *Cell*, **130**, 101-112.
31. Khorshid, M., Hausser, J., Zavolan, M. and van Nimwegen, E. (2013) A biophysical miRNA-mRNA interaction model infers canonical and noncanonical targets. *Nat Methods*, **10**, 253-255.
32. Wee, L.M., Flores-Jasso, C.F., Salomon, W.E. and Zamore, P.D. (2012) Argonaute divides its RNA guide into domains with distinct functions and RNA-binding properties. *Cell*, **151**, 1055-1067.
33. Schirle, N.T., Sheu-Gruttadauria, J. and MacRae, I.J. (2014) Structural basis for microRNA targeting. *Science*, **346**, 608-613.
34. Salomon, W.E., Jolly, S.M., Moore, M.J., Zamore, P.D. and Serebrov, V. (2015) Single-Molecule Imaging Reveals that Argonaute Reshapes the Binding Properties of Its Nucleic Acid Guides. *Cell*, **162**, 84-95.
35. Jo, M.H., Shin, S., Jung, S.R., Kim, E., Song, J.J. and Hohng, S. (2015) Human Argonaute 2 Has Diverse Reaction Pathways on Target RNAs. *Mol Cell*, **59**, 117-124.
36. Herschlag, D. (1991) Implications of ribozyme kinetics for targeting the cleavage of specific RNA molecules in vivo: more isn't always better. *Proc Natl Acad Sci U S A*, **88**, 6921-6925.
37. Jung, S.R., Kim, E., Hwang, W., Shin, S., Song, J.J. and Hohng, S. (2013) Dynamic anchoring of the 3'-end of the guide strand controls the target dissociation of Argonaute-guide complex. *J Am Chem Soc*, **135**, 16865-16871.
38. Deerberg, A., Willkomm, S. and Restle, T. (2013) Minimal mechanistic model of siRNA-dependent target RNA slicing by recombinant human Argonaute 2 protein. *Proc Natl Acad Sci U S A*, **110**, 17850-17855.
39. Sasaki, H.M. and Tomari, Y. (2012) The true core of RNA silencing revealed. *Nature structural & molecular biology*, **19**, 657-660.
40. Zander, A., Holzmeister, P., Klose, D., Tinnefeld, P. and Grohmann, D. (2014) Single-molecule FRET supports the two-state model of Argonaute action. *RNA Biol*, **11**, 45-56.
41. Kawamata, T. and Tomari, Y. (2010) Making RISC. *Trends in biochemical sciences*, **35**, 368-376.
42. Schirle, N.T., Sheu-Gruttadauria, J., Chandradoss, S.D., Joo, C. and MacRae, I.J. (2015) Water-mediated recognition of t1-adenosine anchors Argonaute2 to microRNA targets. *Elife*, **4**.

43. Sternberg, S.H., LaFrance, B., Kaplan, M. and Doudna, J.A. (2015) Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature*, **527**, 1-14.
44. Thomas, M., Lieberman, J. and Lal, A. (2010) Desperately seeking microRNA targets. *Nature structural & molecular biology*, **17**, 1169-1174.
45. Bartel, D.P. (2015) http://www.targetscan.org/. *TargetScan 7.0*.
46. Agarwal, V., Bell, G.W., Nam, J.W. and Bartel, D.P. (2015) Predicting effective microRNA target sites in mammalian mRNAs. *Elife*, **4**.
47. Broderick, J.A., Salomon, W.E., Ryder, S.P., Aronin, N. and Zamore, P.D. (2011) Argonaute protein identity and pairing geometry determine cooperativity in mammalian RNA silencing. *Rna*, **17**, 1858-1869.
48. Grimson, A., Farh, K.K., Johnston, W.K., Garrett-Engele, P., Lim, L.P. and Bartel, D.P. (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell*, **27**, 91-105.
49. Saetrom, P., Heale, B.S., Snove, O., Jr., Aagaard, L., Alluin, J. and Rossi, J.J. (2007) Distance constraints between microRNA target sites dictate efficacy and cooperativity. *Nucleic acids research*, **35**, 2333-2342.
50. Qi, Z., Redding, S., Lee, J.Y., Gibb, B., Kwon, Y., Niu, H., Gaines, W.A., Sung, P. and Greene, E.C. (2015) DNA sequence alignment by microhomology sampling during homologous recombination. *Cell*, **160**, 856-869.
51. Wiedenheft, B., Sternberg, S.H. and Doudna, J.A. (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, **482**, 331-338.
52. Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C. and Doudna, J.A. (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, **507**, 62-67.
53. Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., Westra, E.R., Wanner, B., van der Oost, J., Brouns, S.J. and Severinov, K. (2011) Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc Natl Acad Sci U S A*, **108**, 10098-10103.
54. Westra, E.R., Semenova, E., Datsenko, K.A., Jackson, R.N., Wiedenheft, B., Severinov, K. and Brouns, S.J. (2013) Type I-E CRISPR-cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. *PLoS genetics*, **9**, e1003742.
55. Jiang, F., Zhou, K., Ma, L., Gressel, S. and Doudna, J.A. (2015) A Cas9-guide RNA complex preorganized for target DNA recognition. *Science*, **348**, 1477-1481.
56. Graneli, A., Yeykal, C.C., Robertson, R.B. and Greene, E.C. (2006) Long-distance lateral diffusion of human Rad51 on double-stranded DNA. *Proc Natl Acad Sci U S A*, **103**, 1221-1226.
57. Swarts, D.C., Jore, M.M., Westra, E.R., Zhu, Y., Janssen, J.H., Snijders, A.P., Wang, Y., Patel, D.J., Berenguer, J., Brouns, S.J. *et al.* (2014) DNA-guided DNA interference by a prokaryotic Argonaute. *Nature*, **507**, 258-261.
58. Gao, F., Shen, X.Z., Jiang, F., Wu, Y. and Han, C. (2016) DNA-guided genome editing using the Natronobacterium gregoryi Argonaute. *Nat Biotechnol*.
59. Olovnikov, I., Chan, K., Sachidanandam, R., Newman, D.K. and Aravin, A.A. (2013) Bacterial argonaute samples the transcriptome to identify foreign DNA. *Mol Cell*, **51**, 594-605.
60. Ross, R.J., Weiner, M.M. and Lin, H. (2014) PIWI proteins and PIWI-interacting RNAs in the soma. *Nature*, **505**, 353-359.

Figure 1: Facilitated diffusion

Four different modes of search can in principle be distinguished. 1) 3D search: An Argonaute protein probes a new sequence by first unbinding, then diffusing through the cytosol, and finally binding to probe a new uncorrelated site. 2) Sliding: A non-specifically bound protein laterally diffuses along the mRNA to probe a new site, probing every potential intermediate site from the start to the new site. 3) Hopping: A non-specifically bound protein unbinds, but quickly rebinds again to a site close by (along the RNA) from where it unbound, but not necessarily probing every site in between. 4) Intersegmental transfer: a hopping mechanism where unbinding and binding positions are correlated in 3D space, but far apart along the RNA. This is possible due to the coiled conformation RNA adapts in vivo.

Figure 2: Search-stability paradox

(A) Energies of the binding sites are shown as short black horizontal markers. Being a sum of base pairing energies, binding energies are (approximately) Gaussian distributed with a standard deviation σ. The target site is separated from the other binding sites by an energy of about ΔE. When diffusing laterally, the minimal barrier towards diffusion is set by the energetic difference between neighbouring sites ($\Delta E^{\dagger}$). In reality there are intervening barriers, as depicted by the dashed line. With little loss of generality, we will ignore these additional contributions to the barriers and focus on the best-case scenario.

(B) Recognition mode - Stable binding, but slow search: A larger difference between target and non-target energies comes at the cost of having larger barriers towards diffusion. The right panel shows the complete distribution of energetic states (standard deviation σ_R) of which a subset is plotted in the left panel. The typical (minimal) barrier towards diffusion ($\Delta E_R^{\dagger}$) and differential binding energy (ΔE_R) are indicated.

(C) Search mode - Fast search, but no stable binding: Decreasing the barriers also decreases the difference between target and non-target energy, which hampers the ability of the search complex to selectively bind to the target. The right panel shows the complete distribution of energetic states (standard deviation σ_S) of which a subset is plotted in the left panel. The typical (minimal) barrier towards diffusion ($\Delta E_S^{\dagger}$) and differential binding energy (ΔE_S) are indicated.

(D) Search + Recognition: Fast search and stable binding: If the search complex possesses (at least) two distinct binding modes, it becomes possible to combine the landscapes of figures B (blue) and C (green) to enable rapid diffusion ($\Delta E^{\dagger} \approx \Delta E_S^{\dagger}$) towards the target without loss of selectivity ($\Delta E \approx \Delta E_R$) (orange).

Figure 3: Structural and domain overview of hAgo2 and miRNA

(A) The binary structure of hAgo2-miRNA showing four well conserved domains among Argonaute proteins (snapshot of the structure 4W5N taken in pymol).

(B) Argonaute proteins divide miRNA(orange) in to several domains. The 5'phosphate and nt 1 of miRNA (anchor) is bound to the pocket in the MID domain. The nt 2–8 are known as seed sequence, as they are crucial for initial targeting. The nt 9–10 have the least significant role in target recognition and are known as the mid region. The 3' supplementary region is comprised of nt 13–16, they also have considerable role in stabilizing miRNA-target interaction. The nucleotides beyond the 16^{th} do not base pair with the target and are called the tail region. The 3' OH is bound to the binding pocket in PAZ domain making it as a 3' anchor. The t1 Adenosine (t1A) in the target RNA (pink) binds to the binding pocket in MID domain.

Figure 4: Seed of miRNA and hAgo2-helix7

(A) Nucleotides 2–4 (green) of the guide RNA are well exposed by residues in the PIWI domain (golden surface) possibly for initial target recognition (snapshot of the structure 4W5N taken in pymol).

(B) The access to nt 5–7 of the guide (green) is blocked by the helix-7 motif (red). The base paring of target to guide nt 5–7 would require displacement of helix-7 (snapshot of the structure 4W5N taken in pymol).

(C) Upon base paring with the target (grey) the helix-7 motif is displaced by 4Å compared to guide-only structure. The displacement of helix-7 removes the constraints from nt 6 and 7 (yellow) compared to guide only structure (green) making nt 6 and 7 available for base paring (see the close-up view in the right panel). (snapshot of the structures 4W5N (guide only) and 4W5O (guide and target) taken in pymol).

Figure 5: Cleavage competent state

(A) Structure showing the base pairing between a guide strand (green) and a target strand (red). The base pairing beyond nt 8(g8) is blocked by a residue F811 in a helix of the PIWI domain (snapshot of the structure 4W5O taken in pymol).

(B) A binary structure of hAgo2-miRNA showing the disordered 3' supplementary region of guide RNA (green) passing through a channel between N domain (blue) and PAZ domain (purple) (snapshot of the structure 4W5N taken in pymol).

(C) A ternary structure of hAgo2-miRNA and its target showing an A-form helical arrangement of the 3’ supplementary region of guide (green) in ternary structure (snapshot of the structure 4W5O taken in pymol).

Figure 6: Target search process by hAgo2

(A) A model summarizing conformational changes during target search by hAgo2-miRNA. In light of the search-stability paradox discussed in **Figure 2** we identify a two search modes (pink + green) and a recognition mode (blue). Alternating between search and recognition modes is enabled through the movement of the helix-7 motif (orange).

(B) Schematic free-energy diagram for Ago-microRNA target recognition. Forming bonds between target and guide (horizontal axis) makes the complex more stable (vertical axis). In light of the search-stability paradox, as proposed by Slutsky and Mirny and discussed in **Figure 2**, we identify at least 1 search mode (pre-seed pairing, green arrow) and at least one recognition mode (post-seed pairing, blue arrow). To resolve the paradox, Argonaute can use the movement of its helix-7 motif to switch between search and recognition modes (orange arrow). Potentially, additional modes can be distinguished, such as sub-seed pairing (pink arrow).

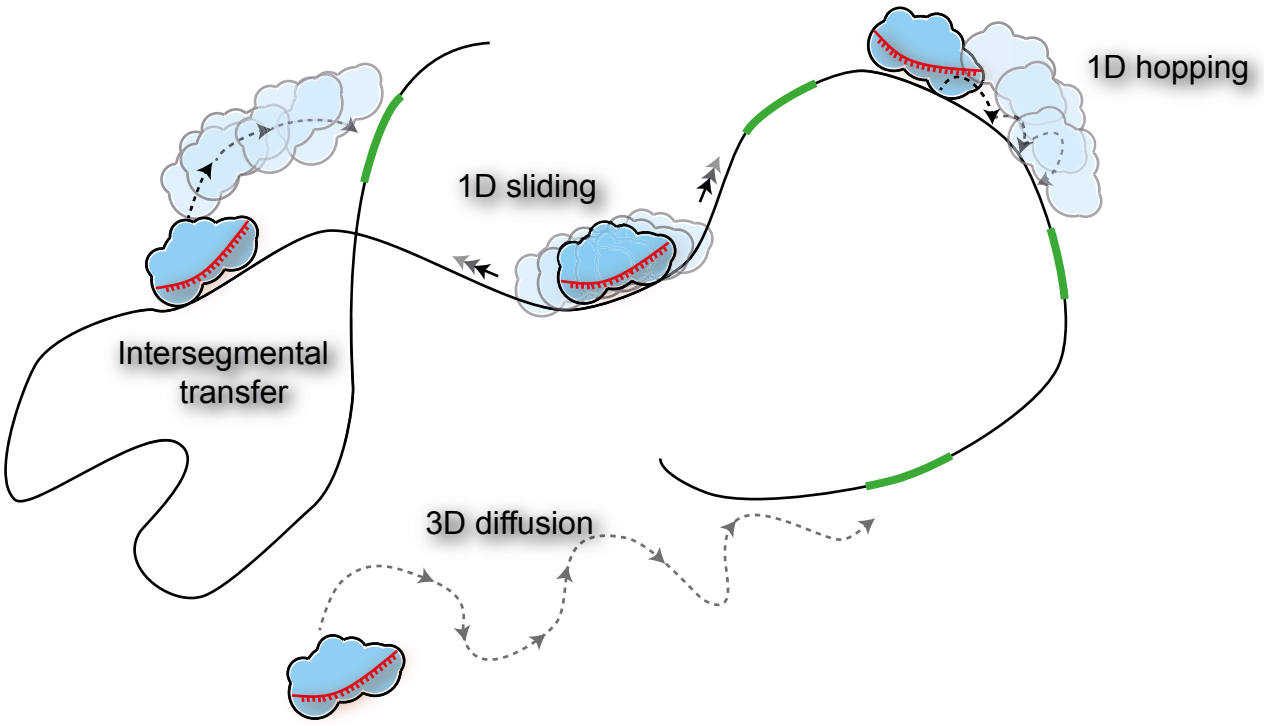

1D hopping
1D sliding
Intersegmental
transfer
3D diffusion

Figure: 1

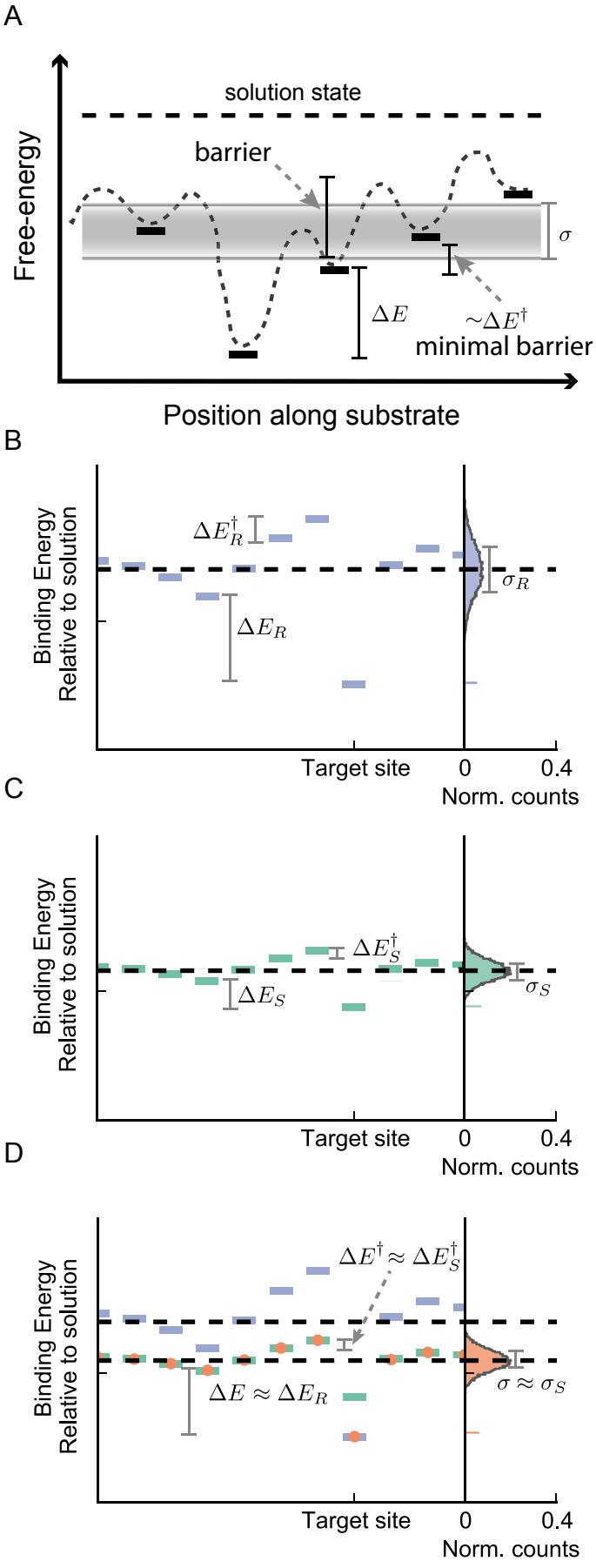
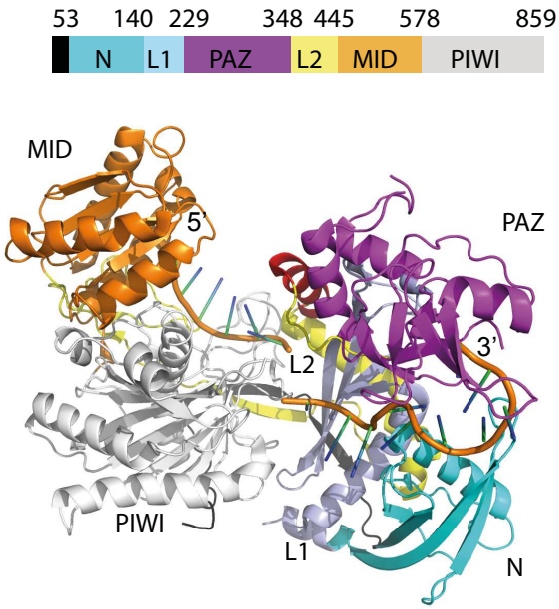

A
solution state
barrier
Free-energy
σ
ΔE
$\sim\Delta E^{\dagger}$
minimal barrier
Position along substrate
B
Binding Energy
Relative to solution
$\Delta E_R^{\dagger}$
σ_R
ΔE_R
Target site
0
0.4
Norm. counts
C
Binding Energy
Relative to solution
$\Delta E_S^{\dagger}$
σ_S
ΔE_S
Target site
0
0.4
Norm. counts
D
Binding Energy
Relative to solution
$\Delta E^{\dagger} \approx \Delta E_S^{\dagger}$
$\sigma \approx \sigma_S$
$\Delta E \approx \Delta E_R$
Target site
0
0.4
Norm. counts

Figure: 2

A


53
140
229
348
445
578
859
N
L1
PAZ
L2
MID
PIWI
MID
5'
PAZ
3'
L2
PIWI
L1
N

B

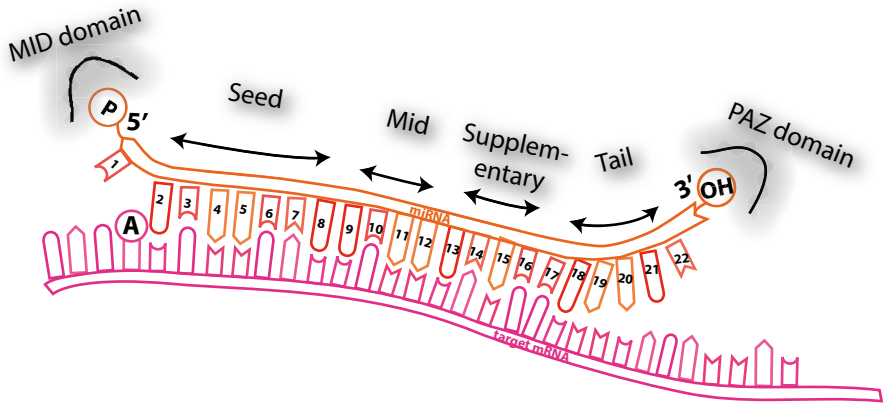

MID domain
P
5'
Seed
Mid
Supplem-
entary
Tail
3'
OH
PAZ domain
A
miRNA
target mRNA

Figure: 3

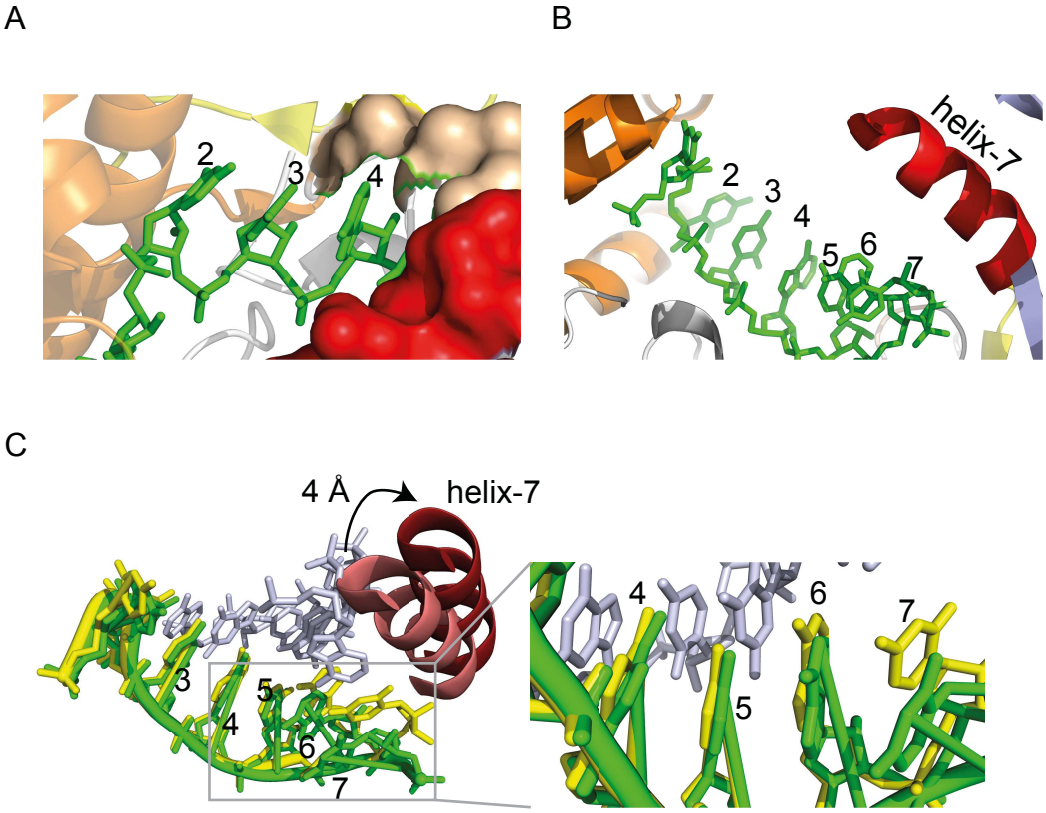


Figure: 4

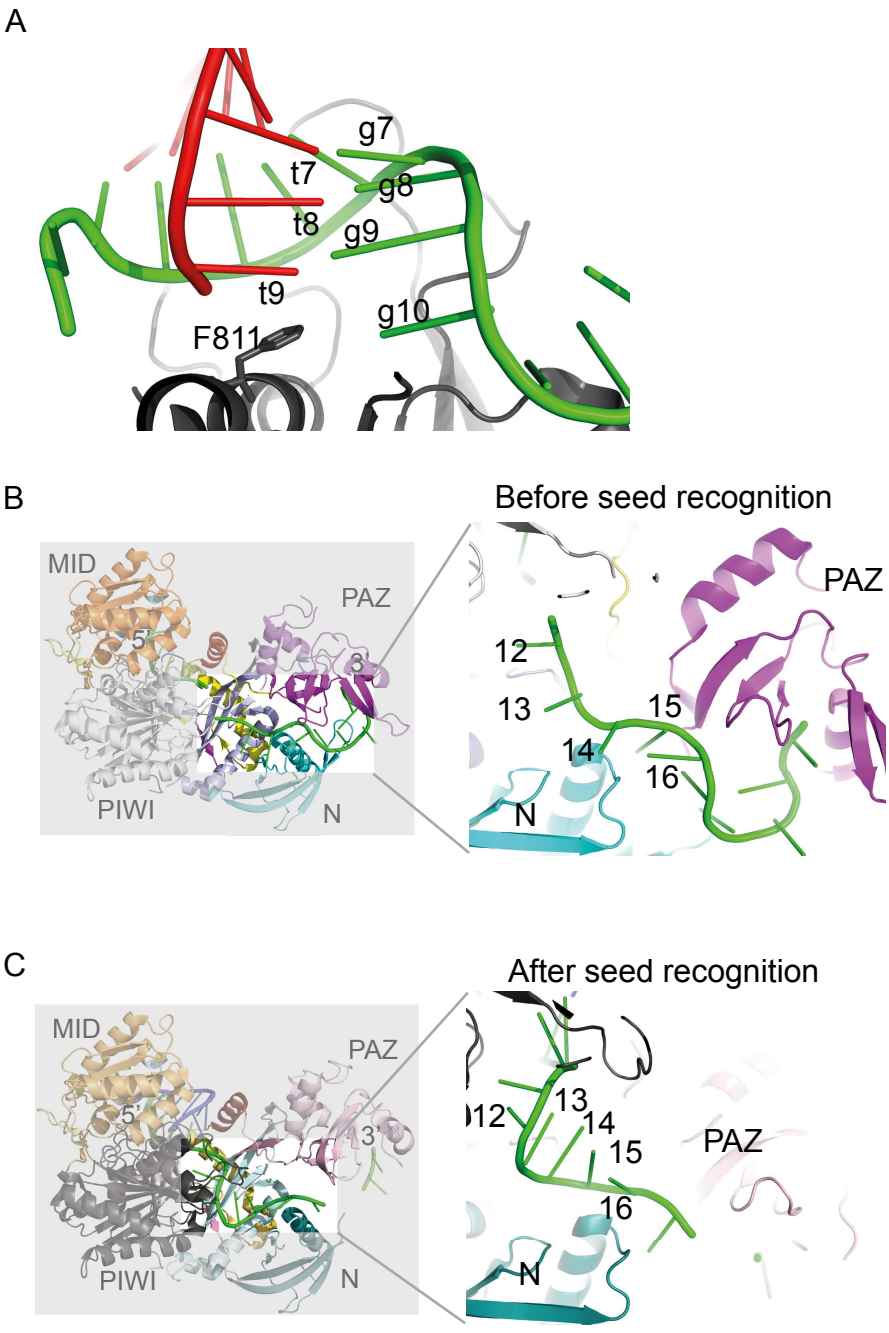

A
g7
t7
g8
t8
g9
t9
g10
F811
B
Before seed recognition
MID
PAZ
5'
3'
PIWI
N
12
13
15
14
16
N
PAZ
C
After seed recognition
MID
PAZ
5'
3'
PIWI
N
12
13
14
15
16
N
PAZ

Figure: 5

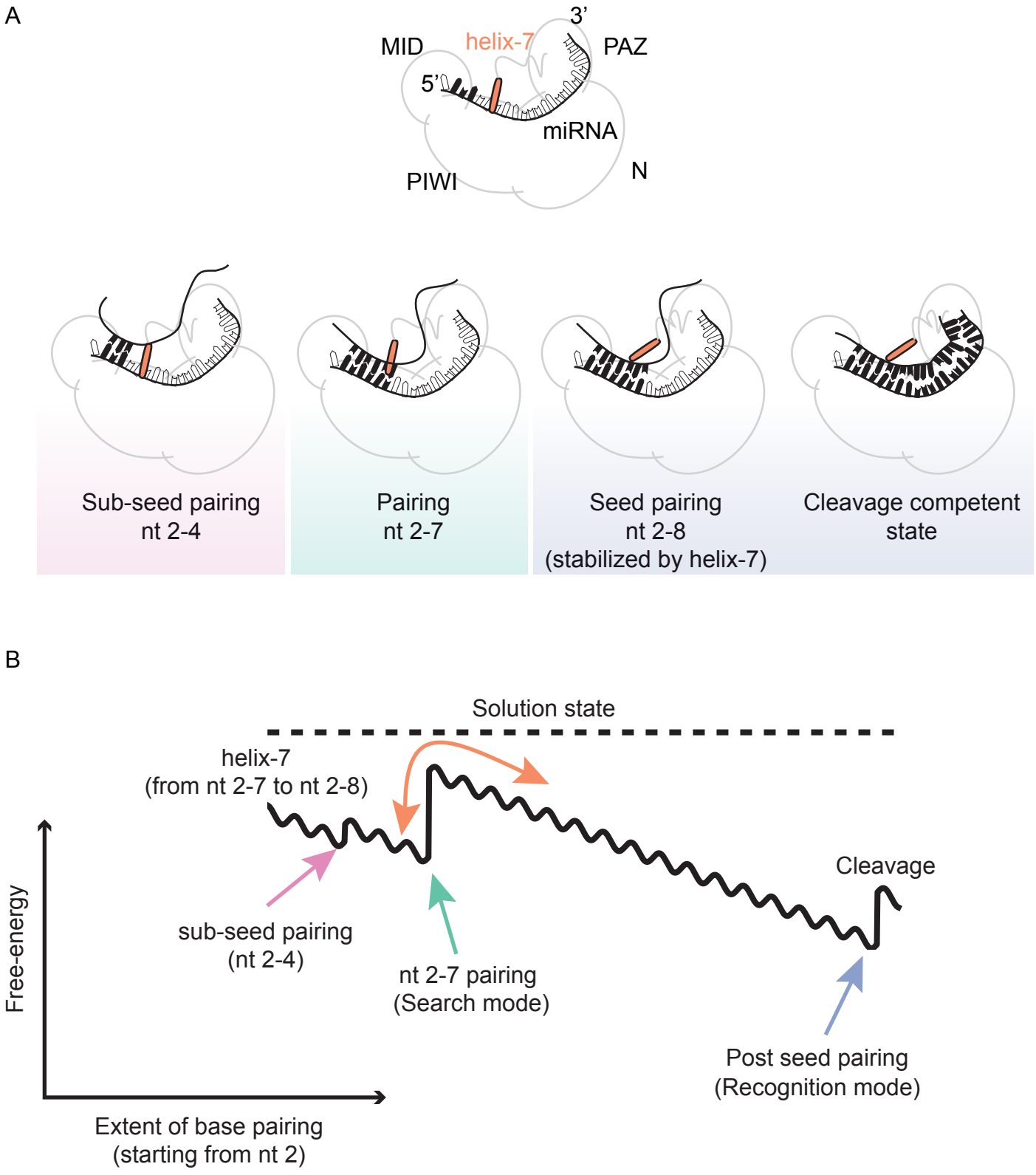

A
3'
MID
helix-7
PAZ
5'
miRNA
PIWI
N
Sub-seed pairing
nt 2-4
Pairing
nt 2-7
Seed pairing
nt 2-8
(stabilized by helix-7)
Cleavage competent
state
B
Solution state
helix-7
(from nt 2-7 to nt 2-8)
Cleavage
sub-seed pairing
(nt 2-4)
nt 2-7 pairing
(Search mode)
Post seed pairing
(Recognition mode)
Free-energy
Extent of base pairing
(starting from nt 2)

Figure: 6