

ANALYSIS OF MICRORNA-TARGET INTERACTIONS BY A TARGET STRUCTURE BASED HYBRIDIZATION MODEL

DANG LONG*, CHI YU CHAN*, YE DING#

Wadsworth Center, New York State Department of Health,
150 New Scotland Avenue, Albany, NY 12208

Email: dlong, yding@wadsworth.org, c.clarence@yahoo.com

MicroRNAs (miRNAs) are small non-coding RNAs that repress protein synthesis by binding to target messenger RNAs (mRNAs) in multicellular eukaryotes. The mechanism by which animal miRNAs specifically recognize their targets is not well understood. We recently developed a model for modeling the interaction between a miRNA and a target as a two-step hybridization reaction: nucleation at an accessible target site, followed by hybrid elongation to disrupt local target secondary structure and form the complete miRNA-target duplex. Nucleation potential and hybridization energy are two key energetic characteristics of the model. In this model, the role of target secondary structure on the efficacy of repression by miRNAs is considered, by employing the Sfold program to address the likelihood of a population of structures that co-exist in dynamic equilibrium for a specific mRNA molecule. This model can accurately account for the sensitivity to repression by *let-7* of both published and rationally designed mutant forms of the *Caenorhabditis elegans lin-41* 3' UTR, and for the behavior of many other experimentally-tested miRNA-target interactions in *C. elegans* and *Drosophila melanogaster*. The model is particularly effective in accounting for certain false positive predictions obtained by other methods. In this study, we employed this model to analyze a set of miRNA-target interactions that were experimentally tested in mammalian models. These include targets for both mammalian miRNAs and viral miRNAs, and a viral target of a human miRNA. We found that our model can well account for both positive interactions and negative interactions. The model provides a unique explanation for the lack of function of a conserved seed site in the 3' UTR of the viral target, and predicts a strong interaction that cannot be predicted by conservation-based methods. Thus, the findings from this analysis and the previous analysis suggest that target structural accessibility is generally important for miRNA function in a broad class of eukaryotic systems. The model can be combined with other algorithms to improve the specificity of predictions by these algorithms. Because the model does not involve sequence conservation, it is readily applicable to target identification for microRNAs that lack conserved sites, non-conserved human miRNAs, and poorly conserved viral mRNAs. StarMir is a new Sfold application module developed for the implementation of the structure-based model, and is available through Sfold Web server at <http://sfold.wadsworth.org>.

* Joint first authors with equal contributions

Corresponding author

1. Introduction

MicroRNAs (miRNAs) are endogenous non-coding RNAs (ncRNAs) of ~22 nt, and are among the most abundant regulatory molecules in multicellular organisms. miRNAs typically negatively regulate specific mRNA targets through essentially two mechanisms: 1) when a miRNA is perfectly or nearly perfectly complementary to mRNA target sites, as is the case for most plant miRNAs, it causes mRNA target cleavage¹; and 2) a miRNA with incomplete complementarity to sequences in the 3' untranslated region (3' UTR) of its target (as is the case for most animal miRNAs) can cause translational repression, and/or some degree of mRNA turnover². miRNAs regulate diverse developmental and physiological processes in animals and plants²⁻⁶. Besides animals and plants, miRNAs have also been discovered in viruses⁷.

The targets and functions of plant miRNAs are relatively easy to identify due to the near-perfect complementarity¹. By contrast, the incomplete target complementarity typical of animal miRNAs implies a huge regulatory potential, but also presents a challenge for target identification. A number of algorithms have been developed for predicting animal miRNA targets. A common approach relies on a "seed" assumption, wherein the target site is assumed to form strictly Watson-Crick (WC) pairs with bases at positions 2 through 7 or 8 of the 5' end of the miRNA. In the stricter, "conserved seed" formulation of the model, perfect conservation of the 5' seed match in the target is required across multiple species^{8,9}. One well-known exception to the seed model is interaction between *let-7* on *lin-41*, for which G-U pair and unpaired base(s) are present in the seed regions of two binding sites with experimental support¹⁰. While the seed model is supported as a basis for identifying many well-conserved miRNA targets¹¹, two studies suggest that G-U or mismatches in the seed region can be well tolerated, and that conserved seed match does not guarantee repression^{12,13}. These suggest that the seed model may represent only a subset of functional target sites, and that additional factors are involved in further defining target specificity at least for some cases with conserved seed matches. Recently, a number of features of site context have been proposed for enhancing targeting specificity¹⁴.

For posttranscriptional gene modulation by mRNA-targeting nucleic acids, the importance of target structure and accessibility has long been established for antisense oligonucleotides and ribozymes^{15,16}, and evidence for this has also emerged for siRNAs^{17,18}; and more recently for miRNAs¹⁹⁻²². These suggest that target accessibility can be an important parameter for target specificity.

We recently developed a model for modeling the interaction between a miRNA and a target as a two-step hybridization reaction: nucleation at an

accessible target site, followed by hybrid elongation to disrupt local target secondary structure and form the complete miRNA-target duplex¹⁹. Nucleation potential and hybridization energy are two key energetic characteristics of the model. In this model, the role of target secondary structure on the efficacy of repression by miRNAs is taken into account, by employing the Sfold program to address the likelihood of a population of structures that co-exist in dynamic equilibrium for a specific mRNA molecule. This model can accurately account for the sensitivity to repression by *let-7* of both published and rationally designed mutant forms of the *Caenorhabditis elegans lin-41* 3' UTR, and for the behavior of many other experimentally-tested miRNA-target interactions in *C. elegans* and *Drosophila melanogaster*. The model is particularly effective in accounting for certain false positive predictions obtained by other methods. In this study, we employed this model to analyze a set of miRNA-target interactions that were experimentally tested in mammalian models. We here report the results of the analysis and discuss implications of the findings.

2. Methods

2.1 mRNA Secondary Structure Prediction

The secondary structure of an mRNA molecule can influence the accessibility of that mRNA to a nucleic acid molecule that can bind to the mRNA by complementary base-pairing. Determination of mRNA secondary structure presents theoretical and experimental challenges. One major impediment to the accurate prediction of mRNA structures stems from the likelihood that a particular mRNA may not exist as a single structure, but in a population of structures in thermodynamic equilibrium²³⁻²⁵. Thus, the computational prediction of secondary structure based on free energy minimization is not well suited to the task of providing a realistic representation of mRNA structures.

An alternative to free energy minimization for characterizing the ensemble of probable structures for a given RNA molecule has been developed²⁶. In this approach, a statistically representative sample is drawn from the Boltzmann-weighted ensemble of RNA secondary structures for the RNA. Such samples can faithfully and reproducibly characterize structure ensembles of enormous sizes. In particular, in comparison to energy minimization, this method has been shown to make better structural predictions²⁷ and to better represent the likely population of mRNA structures²⁸, and to yield a significant correlation between predictions and data for gene inhibition by antisense oligos²⁹, gene knockdown by RNAi³⁰ and target cleavage by hammerhead ribozymes (unpublished data), and translational repression by miRNAs¹⁹. A sample size of 1,000 structures is sufficient to guarantee statistical reproducibility in sampling statistics and

clustering features^{26,28}. The structure sampling method has been implemented in the Sfold software package³¹ and is used here for mRNA folding. The entire target transcript is used for folding if its length is under 7000 nts. For two targets in this study with transcript lengths over 9000 nt, we only used the UTRs (HCV and THRAP1, Table 1), so the folding could be efficiently managed.

2.2 Two-step Hybridization Model

We recently introduced a target-structure based hybridization model for prediction of miRNA-target interaction¹⁹. Here, we briefly describe this model and summarize its energetic characteristics. *In vitro* hybridization studies using antisense oligonucleotides suggested that hybridization of an oligonucleotide to a target RNA requires an accessible local target structure³². This requirement has been supported by various *in vivo* studies³³⁻³⁵. Such a local structure includes a site of unpaired bases for nucleation, and duplex formation progresses from the nucleation site and stops when it meets an energy barrier. In a kinetic study, it was suggested that the nucleation step is *rate-limiting*, and that it involves formation of four or five base pairs between the interacting nucleic acids³⁶. Based on these and other related studies^{37,38}, we model the miRNA-target hybridization as a two-step process: 1) nucleation, involving four consecutive complementary nucleotides in the two RNAs (Fig. 1A), and 2) the elongation of the hybrid to form a stable intermolecular duplex (Fig. 1B).

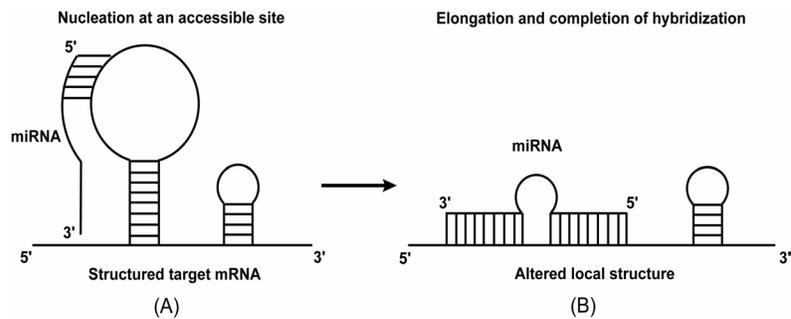


Figure 1. Two-step model of hybridization between a small (partially) complementary nucleic acid molecule and a structured mRNA: 1) nucleation at an accessible site of at least 4 or 5 unpaired bases (A); 2) elongation through “unzipping” of the nearby helix, resulting in altered local target structure (B).

The model is characterized by several energetic parameters. For a given predicted target structure, the nucleation potential, ΔG_N , is the stability of the particular single-stranded 4-bp block within the a potential mRNA binding site

that would form the most stable 4-bp duplex with the miRNA (In Fig. 1, there are two 4-bp blocks for the 5-bp helix formed between the miRNA and the target). For the sample of 1000 structures predicted by Sfold for the target mRNA, the final ΔG_N is the average over the sample. The initiation energy threshold, $\Delta G_{\text{initiation}}$, is the energy cost for initiation of the interaction between two nucleic acid molecule. For two published values of $\Delta G_{\text{initiation}}$ ^{36,39}, 4.09 kcal/mol appeared to perform somewhat better in our previous study¹⁹. Nucleation for a potential site is considered favorable if the nucleation potential can overcome the initiation energy threshold, i.e., $\Delta G_N + \Delta G_{\text{initiation}} < 0$ kcal/mol. For a site with favorable nucleation potential, we next compute ΔG_{total} , the total energy change for the hybridization, by $\Delta G_{\text{total}} = \Delta G_{\text{hybrid}} - \Delta G_{\text{disruption}}$, where ΔG_{hybrid} is the stability of the miRNA-target hybrid as computed by the RNAhybrid program⁴⁰, and $\Delta G_{\text{disruption}}$ is the energy cost for the disruption of the local target structure (Fig. 1B), and is computed using structure sample predicted by Sfold for the target mRNA. These calculations have been incorporated into STarMir, a new application model for the Sfold package. To model the cooperative effects of multiple sites on the same 3' UTR for either a single miRNA or multiple miRNAs, we assume energetic additivity and compute $\sum \Delta G_{\text{total}}$, where the sum is over multiple sites.

2.3 Dataset of MicroRNA-Target Interactions

We tried to assemble a set of high-quality and representative miRNA-target pairs in mammals. We selected reported miRNA-target interactions that were supported by at least two experimental testing using either human cells or mouse or rat models. These interactions play important roles in various biological processes. The targets also include a viral target for a cellular miRNA, and cellular targets for a viral miRNA family. The complete mRNA target sequences were typically retrieved from the Reference Sequence (RefSeq) database from the National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/RefSeq>). Information for these miRNA-target pairs and the references is given in Table 1. For a few reported interactions in these references, the complete transcripts were not available from the GenBank databases and thus these interactions were not included in this study.

3. Results

3.1 Analysis of Interaction between Mammalian miRNAs and Viral Genomes

An intriguing case worthy of particular note is the regulation of Hepatitis C virus (HCV) by miR -122⁴¹. In the viral RNA genome, there are a seed site in

the 5' non-coding region (NCR) and a seed site in the 3' NCR, both are conserved among the six HCV genotypes. However, the site in the 5' NCR was found to be essential for up-regulation of HCV replication by miR-122, whereas the site in the 3' NCR was not. Current miRNA prediction algorithms that based on seed site conservation, e.g., TargetScan⁸, PicTar⁴², cannot explain the lack of function of the 3' NCR seed site. Other algorithms that based only on the alignment and hybridization energy of miRNAs and potential binding sites, e.g., miRanda⁴³, RNAhybrid⁴⁰, cannot explain the difference between those two sites. We analyzed this miRNA-target pair using our interaction model that takes into account secondary structures of the target sequence. To classify an interaction as functional or nonfunctional, we previously used an empirical threshold of -10.0 kcal/mol for $\sum\Delta G_{\text{total}}$ ¹⁹. For this threshold, we predicted a functional interaction between miR-122 and the 5' NCR, but a lack of interaction between miR-122 and the 3' UTR, for which the $\sum\Delta G_{\text{total}}$ is merely -3.54 kcal/mol.

The energetic characteristics for potential binding sites that passed the nucleation threshold are listed below:

hsa-miR-122a:HCV 5' NCR interaction

Site 1: Target site position in 5' NCR: 21–44

		G	CUC	A	AU	C	
Target	21	ACA	CACCAU	G	CACUCC	44	
miRNA	23	UGU	GUGGUA	C	GUGAGG	1	
			UU	A	AGU	U	

$\Delta G_{\text{total}} = -16.70$ kcal/mol; $\Delta G_{\text{disruption}} = 6.40$ kcal/mol; $\Delta G_{\text{hybrid}} = -29.10$ kcal/mol;
 $\Delta G_{\text{N}} + \Delta G_{\text{initiation}} = -3.71$ kcal/mol.

Site 2: Target site position in 5' NCR: 55–70

		C	CU	A	
Target	55	UACUGU	UCACGC	70	
miRNA	23	GUGGUA	AGUGUG	1	
		UGUUU	AC	AGGU	

$\Delta G_{\text{total}} = -9.981$ kcal/mol; $\Delta G_{\text{disruption}} = 7.619$ kcal/mol; $\Delta G_{\text{hybrid}} = -17.60$ kcal/mol;
 $\Delta G_{\text{N}} + \Delta G_{\text{initiation}} = -2.61$ kcal/mol.

hsa-miR-122a: HCV 3' NCR

Site 1: Target site position in 3' NCR: 9–36

		C	UG	AG	GGGUAA	G
Target	9	CGA	A	GUUG	ACACUCCG	36
miRNA	23	GUU	U	UAAC	UGUGAGGU	1
		U	UG	GG	AG	

$\Delta G_{\text{total}} = -3.538$ kcal/mol; $\Delta G_{\text{disruption}} = 20.262$ kcal/mol; $\Delta G_{\text{hybrid}} = -23.80$ kcal/mol; $\Delta G_{\text{N}} + \Delta G_{\text{initiation}} = -3.71$ kcal/mol.

The result here suggests that the lack of function for some (conserved) seed sites can be explained by poor target accessibility. In addition, for each of two single-substitution mutations (p3, p6) and a double-substitution mutation (p3-4) of the proposed seed region in the 5' NCR⁴¹, the HCV RNAs failed to accumulate. Our predictions for the mutants are consistent with the experimental finding, with $\sum \Delta G_{\text{total}}$ of -2.057 kcal/mol, -2.013 kcal/mol, and -1.934 kcal/mol, respectively. We note that the more energetically favorable site 1 in the 5' NCR predicted by our model has some overlap with but is substantially different from the published binding site. This suggests an alternative binding conformation for further testing.

3.2 Analysis of Other MicroRNA-Target Interactions

We next analyzed 18 other validated interactions listed in Table 1. Our model accounted for 16 of the 18 (thus 17 for 19 including HCV 5' NCR, a sensitivity of 89.5%) positive interactions. Among the two positive cases unaccounted for by our model, the interaction between miR-133a and HCN4 has a $\sum \Delta G_{\text{total}}$ of -9.5 kcal/mol, which is close to the threshold, and thus could be effective for miRNA-target hybridization. Moreover, the sum of this energy and that for the interaction between miR-1 and HCN4 is -20.304 kcal/mol, which is consistent with the combined effect by miR-133a and miR-1 on HCN4 that was reported⁴⁴. Because miR-200c is not conserved across five vertebrate genomes, no target prediction can be made by TargetScan⁸.

The regulation of HMGA2 by the *let-7* family (all family members sharing the same seed sequence) has been reported by two studies, with *let-7a* used in one study⁴⁵, and *let-7b*, *let-7d* used in the other⁴⁶. Data from both studies suggested functionality of multiple target sites identified by conserved seed matches. The rather large value of $\sum \Delta G_{\text{total}}$ for the interaction between HMGA2 and any of three tested *let-7* members is consistent with the understanding that a target can be efficiently regulated through multiple sites for the same miRNA.

While convincingly validated mammalian miRNA targets are limited, the functions of viral miRNAs are even less understood. Recently the regulation of several cellular targets by the KSHV-encoded miRNAs has been reported⁴⁷. We found that our model supports the cooperativity of multiple miRNAs acting on the same target. In particular, for the well-validated target, THBS1, the $\sum \Delta G_{\text{total}}$ is rather large, a results of many binding sites on this target 3' UTR. The results for both *let-7* and KSHV miRNAs suggest that $\sum \Delta G_{\text{total}}$ presents a promising

measure for modeling the additive effects of multiple binding sites by either single or multiple mammalian or viral miRNAs.

Table 1. Target prediction based on miRNA-target interaction energy computed by $\sum\Delta G_{\text{total}}$ and local AU content of 19 positive interactions and four negative interactions (shaded)

miRNA	Target name, GenBank Accession, test system ^a and references			$\sum\Delta G_{\text{total}}$ (kcal/mol) and the prediction	Local AU content ^b and the prediction
miR-200c	TCF8	AL831979	H ^{48,49}	-109.979 +	0.742 +
miR-133	RhoA	NM_016802	M ⁵⁰	-17.098 +	NA ^c
miR-133	Cdc42	NM_009861	M ⁵⁰	-29.939 +	0.432 -
miR-133	WHSC2	NM_011914	M ⁵⁰	-18.240 +	0.461 -
miR-1	GJA1	NM_012567	R ⁵¹	-55.805 +	0.626 +
let-7a	HMGA2	NM_003483	H ⁴⁵	-185.195 +	0.628 +
let-7b	HMGA2	NM_003483	H ⁴⁶	-300.446 +	0.628 +
let-7e	HMGA2	NM_003483	H ⁴⁶	-218.989 +	0.628 +
miR-124a	Foxa2	NM_010446	M ⁵²	-7.662 -	0.672 +
miR-1	HCN4	NM_021658	R ⁴⁴	-10.798 +	NA ^c
miR-133a	HCN4	NM_021658	R ⁴⁴	-9.506 -	NA ^c
10 KSHV miRNAs	SPP1	NM_000582	H ⁴⁷	-75.231 +	NA ^d
10 KSHV miRNAs	SRGN	NM_002727	H ⁴⁷	-113.169 +	NA ^d
10 KSHV miRNAs	THBS1	NM_003246	H ⁴⁷	-325.078 +	NA ^d
miR-155	C-Maf	NM_001025577	M ⁵³	-78.858 +	0.710 +
miR-208	THRAP1	NM_005121	H ⁵	-45.032 +	0.895 +
miR-375	Mtpn	NM_145808	H ⁵⁴	-13.084 +	0.715 +
miR-122	CAT-1	NM_013111	R ⁵⁵	-136.82 +	0.530 -
miR-122	HCV 5' NCR	NC_004102	H ⁴¹	-26.681 +	0.478 -
miR-122	HCV 3' NCR	NC_004102	H ⁴¹	-3.538 -	0.412 -
miR-122	HCV 5' NCR p3	NC_004102	H ⁴¹	-2.057 -	NA ^c
miR-122	HCV 5' NCR p6	NC_004102	H ⁴¹	-2.013 -	NA ^c
miR-122	HCV 5' NCR p3-4	NC_004102	H ⁴¹	-1.934 -	NA ^c

^a H: human cells; M: mouse; R: rat. ^b as defined in Grimson *et al.*, 2007¹⁴; ^c no perfect (7- or 8-mer) seed sites; ^d not calculated due to multiple miRNAs; + : predicted effective target, - : predicted ineffective target

We also calculated local AU content of seed sites of the miRNAs and targets following a scoring scheme proposed by Grimson *et al.*¹⁴. When there are multiple seed sites in the same 3' UTR sequence, we report the best local AU content (Table 1). In order to correlate the local AU content to the qualitative information of miRNA activity in our dataset, we select a threshold of 0.6 for the local AU content. miRNA-target pairs having the local AU content is higher or equal 0.6 are predicted functional. This threshold is partly based on the experimental data in Grimson *et al.*¹⁴, where the local AU content of 0.6 correlated to the average fold change of 0.89 in the mRNA level from the

microarray experiment. The AU content of 0.6 is also just above the mean AU content of all possible 7-mer sites of the 3' UTR sequences being considered here (data not shown). For this threshold, the local AU content alone can explain the positive interactions for 9 of the 13 miRNA-target pairs. For each of these 13 pairs, there is at least one seed site and only the concerned miRNA is known to be involved in regulation of the target. In comparison, we predict effective interactions for 12 of the 13 cases (Table 1). Furthermore, both of the two conserved seed sites for miRNA-122 in HCV 5' NCR and 3' NCR have comparable low AU content (Table 1). Therefore, the local AU content cannot explain the functional difference between the two seed sites.

4. Conclusion

In this study, we employed a recently developed target-structure based hybridization model to analyze a set of miRNA-target interactions. These interactions were experimentally tested in human cells or in animal models (mouse or rat). These include mammalian targets for both cellular miRNAs and viral RNAs, and a viral target for a cellular miRNA. Our model can well account for positive interactions, as well as negative interactions. In particular, the model can explain the difference in the interactions of miR-122 to HCV 5' NCR and HCV3' NCR, which could not be explained by several popular miRNA target prediction programs. In our previous analysis of repression data for worm and fly¹⁹, we observed that the model can not only uniquely account for interactions between *let-7* and worm *lin-41* mutants that cannot be explained by other algorithms, but also explain negative experimental results for 11 of 12 targets with seed matches for *lsv-6*. These and the findings of this analysis here suggest that target structural accessibility is generally important for miRNA function in a broad class of eukaryotic systems, and that the model can be combined with other algorithms to improve the specificity of predictions by these algorithms. Our comparison of the predictions based on the interaction energies and the ones based on the local AU content suggests that the local AU content does not reflect accurately target sites' accessibility in many cases. Therefore, the interaction model considered here can more accurately account for miRNA activities. Because the model does not involve sequence conservation, it can be particularly valuable for target identification for microRNAs that lack conserved sites⁵⁶, non- or poorly-conserved human miRNAs⁵⁷ (e.g., the lack of prediction by TargetScan for miR-200c), and usually poorly conserved viral mRNAs.

Acknowledgments

The Computational Molecular Biology and Statistics Core at the Wadsworth Center is acknowledged for providing computing resources for this work. This work was supported in part by National Science Foundation grants DMS-0200970, DBI-0650991, and National Institutes of Health grant GM068726 (Y.D.).

References

1. Rhoades, M.W. et al. *Cell* **110**, 513-20 (2002).
2. Ambros, V. *Nature* **431**, 350-5 (2004).
3. Boehm, M. & Slack, F. *Science* **310**, 1954-7 (2005).
4. Dugas, D.V. & Bartel, B. *Curr Opin Plant Biol* **7**, 512-20 (2004).
5. van Rooij, E. et al. *Science* **316**, 575-9 (2007).
6. Calin, G.A. et al. *N Engl J Med* **353**, 1793-801 (2005).
7. Cullen, B.R. *Nat Genet* **38 Suppl**, S25-30 (2006).
8. Lewis, B.P., Burge, C.B. & Bartel, D.P. *Cell* **120**, 15-20 (2005).
9. Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P. & Burge, C.B. *Cell* **115**, 787-98 (2003).
10. Vella, M.C., Choi, E.Y., Lin, S.Y., Reinert, K. & Slack, F.J. *Genes Dev* **18**, 132-7 (2004).
11. Rajewsky, N. *Nat Genet* **38 Suppl**, S8-13 (2006).
12. Didiano, D. & Hobert, O. *Nat Struct Mol Biol* **13**, 849-51 (2006).
13. Miranda, K.C. et al. *Cell* **126**, 1203-17 (2006).
14. Grimson, A. et al. *Mol Cell* **27**, 91-105 (2007).
15. Vickers, T.A., Wyatt, J.R. & Freier, S.M. *Nucleic Acids Res* **28**, 1340-7 (2000).
16. Zhao, J.J. & Lemke, G. *Mol Cell Neurosci* **11**, 92-7 (1998).
17. Overhoff, M. et al. *J Mol Biol* **348**, 871-81 (2005).
18. Schubert, S., Grunweller, A., Erdmann, V.A. & Kurreck, J. *J Mol Biol* **348**, 883-93 (2005).
19. Long, D. et al. *Nat Struct Mol Biol* **14**, 287-294 (2007).
20. Zhao, Y. et al. *Cell* **129**, 303-17 (2007).
21. Zhao, Y., Samal, E. & Srivastava, D. *Nature* **436**, 214-20 (2005).
22. Robins, H., Li, Y. & Padgett, R.W. *Proc Natl Acad Sci U S A* **102**, 4006-9 (2005).
23. Christoffersen, R.E., McSwiggen, J.A. & Konings, D. *J. Mol. Structure (Theochem)* **311**, 273-284 (1994).
24. Altuvia, S., Kornitzer, D., Teff, D. & Oppenheim, A.B. *J Mol Biol* **210**, 265-80 (1989).
25. Betts, L. & Spremulli, L.L. *J Biol Chem* **269**, 26456-63 (1994).
26. Ding, Y. & Lawrence, C.E. *Nucleic Acids Res* **31**, 7280-301 (2003).
27. Ding, Y., Chan, C.Y. & Lawrence, C.E. *RNA* **11**, 1157-66 (2005).

28. Ding, Y., Chan, C.Y. & Lawrence, C.E. *J Mol Biol* **359**, 554-71 (2006).
29. Ding, Y. & Lawrence, C.E. *Nucleic Acids Res* **29**, 1034-46 (2001).
30. Shao, Y. et al. *RNA* (2007).
31. Ding, Y., Chan, C.Y. & Lawrence, C.E. *Nucleic Acids Res* **32**, W135-41 (2004).
32. Milner, N., Mir, K.U. & Southern, E.M. *Nat Biotechnol* **15**, 537-41 (1997).
33. Darnell, J.C. et al. *Genes Dev* **19**, 903-18 (2005).
34. Friebe, P., Boudet, J., Simorre, J.P. & Bartenschlager, R. *J Virol* **79**, 380-92 (2005).
35. Mikkelsen, J.G., Lund, A.H., Duch, M. & Pedersen, F.S. *J Virol* **74**, 600-10 (2000).
36. Hargittai, M.R., Gorelick, R.J., Rouzina, I. & Musier-Forsyth, K. *J Mol Biol* **337**, 951-68 (2004).
37. Paillart, J.C., Skripkin, E., Ehresmann, B., Ehresmann, C. & Marquet, R. *Proc Natl Acad Sci U S A* **93**, 5572-7 (1996).
38. Reynaldo, L.P., Vologodskii, A.V., Neri, B.P. & Lyamichev, V.I. *J Mol Biol* **297**, 511-20 (2000).
39. Xia, T. et al. *Biochemistry* **37**, 14719-35 (1998).
40. Rehmsmeier, M., Steffen, P., Hochsmann, M. & Giegerich, R. *RNA* **10**, 1507-17 (2004).
41. Jopling, C.L., Yi, M., Lancaster, A.M., Lemon, S.M. & Sarnow, P. *Science* **309**, 1577-81 (2005).
42. Krek, A. et al. *Nat Genet* **37**, 495-500 (2005).
43. Enright, A.J. et al. *Genome Biol* **5**, R1 (2003).
44. Xiao, J. et al. *J Cell Physiol* **212**, 285-92 (2007).
45. Mayr, C., Hemann, M.T. & Bartel, D.P. *Science* **315**, 1576-9 (2007).
46. Lee, Y.S. & Dutta, A. *Genes Dev* **21**, 1025-30 (2007).
47. Samols, M.A. et al. *PLoS Pathog* **3**, e65 (2007).
48. Hurteau, G.J., Carlson, J.A., Spivack, S.D. & Brock, G.J. *Cancer Res* **67**, 7972-6 (2007).
49. Hurteau, G.J., Spivack, S.D. & Brock, G.J. *Cell Cycle* **5**, 1951-6 (2006).
50. Care, A. et al. *Nat Med* **13**, 613-8 (2007).
51. Yang, B. et al. *Nat Med* **13**, 486-91 (2007).
52. Baroukh, N. et al. *J Biol Chem* **282**, 19575-88 (2007).
53. Rodriguez, A. et al. *Science* **316**, 608-11 (2007).
54. Poy, M.N. et al. *Nature* **432**, 226-30 (2004).
55. Jopling, C.L., Norman, K.L. & Sarnow, P. *Cold Spring Harb Symp Quant Biol* **71**, 369-76 (2006).
56. Farh, K.K. et al. *Science* **310**, 1817-21 (2005).
57. Bentwich, I. et al. *Nat Genet* **37**, 766-70 (2005).